Economics 1011A Textbook

Based on class by Prof. Edward L. Glaeser

KEVIN BI AND JIMMY LIN

September 7, 2020

Contents

Part I

Introduction

1 Economic Modeling Overview

1.1 What is economic modeling?

The goal of an economic model is to take some real world phenomena and represent it in a way that we can analyze it rigorously and tractably. At the most fundamental level, a **model** is just a set of assumptions that we make. Models can take many forms and you are likely already familiar with some models of real world phenomena. For example, you might construct a model of dining preferences by assuming that the number of people who eat in the dining halls is greater on days where the food is good. Or you might have seen toy models of molecules where the atoms are represented by marshmallows and the bonds between them as toothpicks. With the COVID-19 pandemic, models of infectious disease were used to predict the future number of infectious by assuming how often people interacted with each other,

Economic models try to answer economic questions. Some of these questions might include:

- How does the minimum wage affect the amount of labor hired?

- What is the tax rate that maximizes revenue for the government?

- How does the interest rate affect savings and investment?

Importantly, economic models usually focuses on **positive** economic questions, that is questions about how things in the world work. This is contrast to **normative** economics, which answers how things should be. However, that does not mean that economic models cannot tell us the optimal policy. While an economic model cannot answer what the 'best' policy is in a vaccum, it can tell us what the best policy is once we have defined what 'best' means. Normative economics examines what it means for a policy to be the 'best,' while positive economics tells us how we get there.

A model can take many forms, but in this class, we focus on mathematical models of economic behavior. That is, we try to represent the behavior of people, firms, and governments via mathematical functions and see what insights can be gained from such representation. The use of mathematics allows us to formalize our economic reasoning and make precise what conclusions must follow from certain assumptions.

What makes a good economic model?

As you will learn throughout this course, economic models can take many forms and there is almost no limit to the models that you can create. However, just because you can write down a certain model, does not mean that model is a good one. Most good economic models share a few key characteristics:

- A model should make clear what assumptions are being included, how these assumptions affect the model's conclusions, and the potential limitations of these assumptions.

- A model should be general enough to be a realistic representation of the real world but also simple enough to be easy to manipulate and interpret mathematically. Striking this balance will be a key theme of this course and will be important for doing well in each of the modeling projects.

- The conclusions of a model should tie back to the problem being asked. Deriving a mathematical expression that quantifies a particular behavior or result is often the majority of the battle, but it is important to understand how these results answer our initial question and whether our interpretation makes sense intuitively.

The role of assumptions

You may have heard that economists make unrealistic assumptions in their models, and to a certain degree this is true. Most economists do not think that individuals or even firms are actually able to perfectly optimize their decisions. However, since the real world is too complex to model perfectly, assumptions play a few crucial roles:

Tractability Perhaps the main role of assumptions is to make models tractable to solve analytically. It would be almost impossible to make concise models that generate useful predictions if we had to figure out how every person in the world makes decisions. Assumptions allow us to simplify the model so that they can actually be solved with current mathematical techniques.

Illustrate possibility Related to the tractability rationale for models, we may make assumptions to show that certain mechanics are at least possible under assumptions that are not too unreasonable. The assumptions allow us to simplify the problem so that the mechanisms are more clear, and helps us obtain a better understanding by removing some of the "noise" that might be present absent said assumptions.

Evaluate differences between models By specifying assumptions explicitly, economists are able to understand where two models differ and why they might reach different results. In particular, it tells us when one model might be more applicable than another. For example, if we assume that individuals drive at the fastest possible

speed, this might be an accurate assumption on an empty highway, while a model that says drivers try to minimize their risk of an accident would be more applicable to a crowded intersection.

Specify points of failure By specifying our assumptions, we also specify what must follow if you believe those assumptions to hold. Importantly, if we observe that the real world does not behave the way that our model predicts, it tells us exactly where we should look to see why the model is inaccurate.

Close enough While assumptions in economic models might seem very unrealistic on an individual level, they can often be close enough to the truth in aggregate that we can still derive useful and accurate predictions from said models. For example, while individual firms might not be perfectly optimal, it may be reasonable to say that on aggregate, they make decisions that are pretty close to optimal even if some firms deviate slightly.

Assumptions can make your life a lot easier when trying to model some economic phenomena. However, you will want to be careful. In particular, assumptions should help you reach conclusions, but you should avoid assuming the conclusion itself. While it is in general better for models to approximate reality, you should not feel pressure to make your model too close to reality or else it loses much predictive power and clarity.

Now that we have some understanding of how economists think about developing economic models, the rest of this chapter will build up a foundation of the different components of many models and introduce the mathematical tools relevant for analyzing them. Future chapters will apply and extend this foundation to various economic settings, which will help us characterize the way different pieces of the economy behave.

1.2 Optimization problems

Almost all economic models boil down to one component: agents making decisions. The question that we try to answer is how agents make those decisions, and what are the consequences when many agents are making decisions at the same time. Some canonical examples of questions we can ask include the following:

- How do workers decide how many hours to work?

- How do households decide what goods they consume, and how much?

- How do firms decide how much to produce? How do they decide how much labor to employ and capital to use?

In economics, it is usually assumed that the agents are trying to achieve the best possible outcome in some form. However, the term "best" can be unclear, so to formalize

the concept, economists assume that agents are trying to maximize (or minimize) an
objective function, which is a function $f : X \rightarrow \mathbb{R}$ from the set of possible choices X
to the real numbers $\mathbb{R}$. Examples of objective functions might include:

- A firm choosing how many people to hire to maximize profits

- A politician choosing which ads to buy to maximize votes

- A shopper choosing what food to buy to maximize health

The inputs of the objective function, from the perspective of the agent, are the choices
that the agent makes. We refer to these variables as **choice variables** – the choices that
the agent gets to make. An example of a choice variable might be how many workers
to hire. Choice variables are a part of a broader class of variables called **endogenous
variables**, which is any variable where the value of the variable is determined by the
choices the agent makes. The difference between endogenous variables and choice variables
is subtle. An example of an endogenous variable that might not be considered a choice
variable would be the profits that a firm makes. While the firm might not directly choose
the profits, the choices they make clearly affect the profits.

Endogenous variables are in contrast to **exogenous variables**, which are variables
that are determined outside of the model and are not affected by the agent's decisions.
Some examples of exogenous variables might include:

- The amount of land available

- The tax rate

- The productivity of workers

One important point is that exogenous variables are not exogenous in all models. A firm
might interpret the government's tax rate as exogenous in one model, but if our agent is
the government, then the tax rate would be endogenous. Throughout this book, we will
refer to agents *perceiving* certain quantities as exogenous, which means that the agents
optimize by assuming that their actions do not affect said quantities, even if they might
in the full model.

In this book, we distinguish endogenous and exogenous variables in a function's
arguments by writing the endogenous variables to the left of a semicolon (;), and the
exogenous variables to the right:

$$f(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m)$$

In the example above, $x_1, \ldots, x_n$ would be the endogenous variables, while $y_1, \ldots, y_n$
would be the exogenous variables.

In many cases, agents are not able to make any choice that they want to maximize
their objective function. For example, an individual who goes shopping cannot buy

unlimited goods, because they cannot spend more than their budgets. Such restrictions on the choices that an agent can make are called **constraints**. Constraints can either be equality or inequality constraints. Some examples include:

- The amount of spending s must be less than or equal to the budget b, $s \leq b$.

- The farm must produce exactly x bushels of corn c, $c = x$.

- The number of hours spent working h can be at most the number of hours in a day 24, $h \leq 24$

For the most part in economics, we deal with inequality constraints because we are mostly considering how agents behave when resources are scarce. A problem with constraints is called a **constrained maximization** problem, and conversely, a problem without constraints is called an **unconstrained maxizmization** problem.

Now that we have defined the vocabulary of optimization, we can proceed to setting up a general optimization problem. For the purposes of concision, we use vector notation. Let $\mathbf{x} = (x_1, \ldots, x_n) \in X$ be the choice variables, and let $\mathbf{y} = (y_1, \ldots, y_m) \in Y$ be exogenous variables. Let $f : X \times Y \to \mathbb{R}$ be the objective function.[1] If we do not have constraints, then we can write the unconstrained maximization problem as

$$\max_{\mathbf{x}} f(\mathbf{x}; \mathbf{y})$$

The above is mathematical for "maximize f from choices of $\mathbf{x}$". Now suppose we have constraints, $g_1(\mathbf{x}; \mathbf{y}) \leq c_1, \ldots, g_k(\mathbf{x}; \mathbf{y}) \leq c_k$. Then we can write the maximization problem as,

$$\max_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}) \text{ s.t. } g_1(\mathbf{x}; \mathbf{y}) \leq c_1, \ldots, g_k(\mathbf{x}; \mathbf{y}) \leq c_k$$

The above is mathematical notation for "maximize f with choice of $\mathbf{x}$ subject to (s.t.) the constraints."

Optimized quantities as functions

Now that we have set up the problem, we can consider the choice of $\mathbf{x}$ that maximizes our objective function. We use the argument maximum to refer to this maximized quantity, and we normally denote the maximized quantity with an asterisk:

$$\mathbf{x}^* = \operatorname*{argmax}_{\mathbf{x}} f(\mathbf{x}; \mathbf{y})$$

---

[1]As a review of notation, $X \times Y$ refers to the Cartesian product between sets X and Y. That is, if $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2, y_3\}$, then $X \times Y$ is the set of ordered pairs with the first element from X and the second element from Y, so $X \times Y = \{(x_1, y_1), (x_1, y_2), (x_1, y_3), (x_2, y_1), (x_2, y_2), (x_2, y_3)\}$. Do not worry if this notation (and other instances of mathematical notation) is new for you; it is much more important that you are able to grasp the meaning of the underlying concepts.

If we assume that the value of $\mathbf{x}^*$ is unique, then notice that $\mathbf{x}^*$ is a *function* of the exogenous variables, $\mathbf{y}$. That is, once we have specified the objective function f, as well as the exogenous quantities, the value of $\mathbf{x}^*$ is entirely determined by $\mathbf{y}$. This means that we can ask questions like, if the value of $\mathbf{y}$ changes, how would our optimal choice, $\mathbf{x}^*$ change? For example, if you decide to buy apples in order to maximize your happiness, you might ask how does a change in the price of apples affect the amount of apples that you buy. This is known as a **comparative static**, how some optimal choice or equilibrium quantity changes in response to a change in an exogenous variable. Taking comparative statics is one of the central goals of economic modeling. Questions related to comparative statics might include:

- If a firm chooses an optimal number of workers to maximize profits, how does this quantity depend on the minimum wage set by the government?

- If households choose an amount to save to maximize their long-term utility, how does this amount depend on the economy's interest rate?

- If a consumer chooses a quantity of goods to purchase based on their personal preferences and the goods' prices, how do consumption quantities depend on the sales tax rate?

As we see in the examples above, modeling comparative statics is often useful for studying the effects of different policy interventions. At a high level, solving an agent's optimization problem describes how individual agents behave, while solving for comparative statics describes how environmental changes affect these behaviors.

A common point of confusion among students is the distinction between $\mathbf{x}$, which is the name that we give to the choice variable, and $\mathbf{x}^*$, which is the value that optimizes the objective function. In principle, $\mathbf{x}$ can be any value. For example, let's say that $\mathbf{x} = (a, b)$, where a is the number of apples you buy and b is the number of bananas. You could in principle buy $a = 2$ apples and $b = 3$ bananas, even though you would be happier with $a = 4$ and $b = 4$. If we do not know your optimization function and how your choice of a and b depend on exogenous variables like prices, then we cannot study the comparative statics for a or b.

However, if you say that you are going to buy the number of apples and bananas that makes you the most happy, then this is $\mathbf{x}^*$ and will depend on the price as well as many other outside factors. In this case, it makes sense to ask how the number of apples you buy depends on price. Optimization is a bit like telling a robot, "go and make the choices that will make me most happy." Once you have done that, the choice is fully determined by the outside world, and so we say that the optimal choice, $\mathbf{x}^*$ is a function of the exogenous variables $\mathbf{y}$. Technically, we should write $\mathbf{x}^*(\mathbf{y})$, but often we will treat this as implicit and just write $\mathbf{x}^*$.

A common mistake when solving optimization problems is to write $\mathbf{x}^*$ as a function of one or more of the choice variables. This is a category error. It essentially says that our optimal choice depends on our choices. However, you can say that the optimal choice of one quantity depends on the optimal choice of another quantity. That is, you might have that x_1^* depends on x_2^* in some way. However, because x_2^* is a function of the exogenous variables, $\mathbf{y}$, then $x_1^*(x_2^*(\mathbf{y}))$ is also a function of only exogenous variables.

2 Math Review

Throughout economics, we use mathematics to formalize our thinking and to make sure that our chain of reasoning makes sense. In this chapter, we provide a review of the mathematics that will be necessary for this course.

Because economics focuses primarily on optimizing agents on the margin, we extensively use multivariable calculus, for both constrained and unconstrained optimization. In this chapter, we review the basic concepts of differentiation, constrained and unconstrained optimization, as well as some notation that will be used throughout the course.

2.1 Differentiation

Single variable differentiation

Perhaps the most important mathematical concept for this course is that of the derivative. Suppose we have some function, $f : \mathbb{R} \to \mathbb{R}$, where $\mathbb{R}$ is the set of real numbers and the above notation tells us that the function f takes a real number as an input and returns a real number. Formally, the **derivative** of f at a point x is defined as,

$$\frac{df}{dx}(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{dh}$$

Informally, the derivative, $\frac{df}{dx}(x)$, represents how much the value of f changes for a very small increase in the value of x. Graphically, the derivative is the slope of the line tangent to f at x. The derivative, $\frac{df}{dx}$, is a function of x, but we will often omit the arguments to the function and just write $\frac{df}{dx}$.

Notice that the definition of the derivative assumes that the limit exists. For the most part in this course, we assume that f is **smooth**, which means that we can differentiate f infinitely many times. $\frac{df}{dx}$ is also called the **first derivative** of f, because it is the result of differentiating f once. To get a higher order derivative, we simply differentiate the derivative. $\frac{d^2 f}{dx^2}$ is the **second derivative** of f, and is found by taking the derivative of $\frac{df}{dx}$, and higher order derivatives are found similarly. The notation for the nth derivative of f is given by $\frac{d^n f}{dx^n}(x)$. The second derivative, $\frac{d^2 f}{dx^2}$, is of particular importance in economics because it represents the concavity/convexity of a function. If $\frac{d^2 f}{dx^2} > 0$, then we say that f is **convex** at x, and if $\frac{d^2 f}{dx^2}(x) < 0$, then we say that f is **concave** at x. If $\frac{d^2 f}{dx^2}(x) < 0$ for all x, then f is **globally concave**, and if $\frac{df^2}{dx^2}(x) > 0$ for all x, then f is **globally**

[convex](#). We will very rarely need to deal with cases where the derivative is of an order higher than 2.

Differentiation rules

We assume knowledge of some basic rules and properties of differentation. We list some of the most important ones here:

Power rule For a constant α,

$$\frac{d}{dx}(x^\alpha) = \alpha x^{\alpha-1}$$

Linearity For $\alpha, \beta \in \mathbb{R}$, and functions f, g, we have

$$\frac{d}{dx}(\alpha f(x) + \beta g(x)) = \alpha \frac{df}{dx}(x) + \beta \frac{dg}{dx}(x) = \alpha f'(x) + \beta g'(x)$$

Product Rule For functions f, g,

$$\frac{d}{dx}(f(x) \cdot g(x)) = \frac{dg}{dx}(x)f(x) + \frac{df}{dx}(x)g(x) = g'(x)f(x) + f'(x)g(x)$$

Chain Rule For functions f, g,

$$\frac{d}{dx}(f(g(x))) = \frac{df}{dx}(g(x)) \cdot \frac{dg}{dx}(x) = f'(g(x))g'(x)$$

Log In this course we use log to refer to the natural logarithm (also commonly written as ln),

$$\frac{d}{dx}(\log(x)) = \frac{1}{x}$$

Exponential

$$\frac{d}{dx}(e^x) = e^x$$

We can generalize this using the chain rule, so that for any constant a, we have

$$\frac{d}{dx}a^x = \log(a)a^x$$

Inverse differentiation While the derivative answers how f changes for a small change in x, we can similarly ask how much does x change for a small change in f, which

is the inverse derivative,

$$\frac{dx}{df}(x) = \frac{1}{\frac{df}{dx}(x)}$$

Differentiation with respect to a function We can more generally ask, how does a function $f(x)$ change if we increase one component of f, say $g(x)$ by a small amount, this yields the derivative of $f(x)$ with respect to $g(x)$,

$$\frac{df}{dg}(x) = \frac{df(x)}{dx}\frac{dx}{dg(x)} = \frac{df}{dx}(x)\frac{1}{\frac{dg}{dx}(x)} = \frac{f'(x)}{g'(x)}$$

Multivariable differentiation

While single variable differentiation tells us how a function changes when there is a single input, we often have functions of multiple variables. Suppose we have a function $f(x_1, x_2, \ldots, x_n)$, where $x_1, x_2, \ldots, x_n$, are the different arguments that are taken as inputs to the function f. We can also write the input to f in **vector notation**, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, and the function as $f(\mathbf{x})$. Throughout this course, we use a bolded letter to denote a vector of values. Formally then, a multivariable function is a function $f : \mathbb{R}^n \to \mathbb{R}$ which takes an n dimensional vector as input, and returns a number as output.[1]

Now, we can examine how to differentiate such a multivariate function.

Partial Differentiation

While in the single variable case, the derivative tells us how f changes for small change in the input, x, in the multivariable, we consider how f changes for a small change to one of the inputs, say x_k, while holding all other inputs fixed. Formally, the **partial derivative** of f with respect to an input x_k at a point $\mathbf{x} = (x_1, \ldots, x_k, \ldots, x_n)$,

$$\frac{\partial f}{\partial x_k}(\mathbf{x}) = f_{x_k}(\mathbf{x}) = f_k(x) = \lim_{dx_k \to 0} \frac{f(x_1, \ldots, x_k + h, \ldots, x_n) - f(x_1, \ldots, x_k, \ldots, x_n)}{h}$$

You may notice that this is very similar to the single variable cases, and indeed partial differentiation is very similar to single variable differentiation, except you treat all other components as fixed. This means that all of the above single differentiation rules also hold for the multivariable case, except replacing the derivative with the partial derivative.

We can also take higher order derivatives. Similar to the single variable case, we can differentiate with respect to the same variable twice, which we denote as $\frac{\partial^2 f}{\partial x_k^2}$. We could also first differentiate with respect to x_k first, and then differentiate that result with

---

[1] As a review of notation, recall that $\mathbb{R} \times \mathbb{R}$ is the Cartesian product of $\mathbb{R}$ with itself, which is the set of ordered pairs of real numbers. We denote $\mathbb{R}^n := \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ (n times); that is, $\mathbb{R}^n$ is the set of ordered n-tuples of real numbers, or equivalently the set of real vectors in n-dimensional space.

respect to another variable, say x_j. This is known as the **cross-partial** of f with respect to x_k and x_j, and is written,

$$\frac{\partial^2 f}{\partial x_k \partial x_j}$$

Intuitively, the quantity $f_{x_k x_j}$ represents how an increase in x_j changes the marginal effect that x_k has on f.

One important result on cross-partials is **Young's Theorem**, which states the following:

Theorem (Young)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function with inputs $x_1, \ldots, x_n$, then* $\frac{\partial^2 f}{\partial x_k \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_k}$.

This tells us that for a well-behaved function (in this case we assume smooth with respect to all inputs), then the order we take derivatives in does not matter.

Total differentiation

While partial differentiation tells us how a function changes for a single input, keeping all other inputs fixed, it is important to remember that with a partial derivatives, all the inputs are really the *names* of inputs that will eventually take on values. In that case, we can consider the following scenario. Suppose we have some multivariable function, $f : \mathbb{R}^n \to \mathbb{R}$, but then we define the single variable function $g : \mathbb{R} \to \mathbb{R}$ as follows:

$$g(t) = f(t, t, t, \ldots, t)$$

That is, we are defining g to be the value of f when $x_1 = x_2 = \ldots x_n = t$, where t is some value. While we can take partial derivatives of f to see how f changes in response to a single input, in this case, we want to see how g changes in response to all inputs. One way to do this would be to find out the explicit form of f and just plug in t in all the appropropriate places, and then differentiate, but this would require us to know exactly what f is. However, it would be nice if we could see how g changes with respect to t without needing to know how t enters into f explicitly.

This is the value of the **total derivative**, which tells us how $g(t) = f(x_1(t), \ldots, x_n(t))$ changes with respect to t. In this case, we treat each x_k as a single variable function of t which then feeds into the kth input of f. To find how this changes with respect to t, we use the **multivariable chain rule**, which states,

$$\frac{dg(t)}{dt} = \frac{df(x_1(t), \ldots, x_n(t))}{dt} = \frac{\partial f(\mathbf{x}(t))}{\partial x_1} \frac{dx_1(t)}{dt} + \cdots + \frac{\partial f(\mathbf{x}(t))}{\partial x_n} \frac{dx_n(t)}{dt}$$

Intuitively, you can think of each term of the sum as how much a small change in x_k affects f, multiplied by how a small change in t affects x_k. The total effect of a small change to t is all of those individual changes added together.

Example. Define $f(x,y) = x^\alpha y^\beta$. Define $g(t) = f(t^2, 3t)$. So we are setting $x(t) = t^2$ and $y(t) = 3t$. The multivariable chain rule tells us,

$$\frac{dg(t)}{dt} = \frac{\partial f(x(t), y(t))}{\partial x} \frac{dx(t)}{dt} + \frac{\partial f(x(t), y(t))}{\partial y} \frac{dy(t)}{dt}$$

Computing each of the terms, we have

$$\frac{\partial f}{\partial x} = \alpha x^{\alpha-1} y^\beta$$
$$\frac{\partial f}{\partial y} = \beta x^\alpha y^{\beta-1}$$
$$\frac{dx(t)}{t} = 2t$$
$$\frac{dy(t)}{t} = 3$$

Plugging in all of our values yields,

$$\frac{dg(t)}{dt} = \alpha x(t)^{\alpha-1} y(t)^\beta \cdot 2t + \beta x(t)^\alpha y(t)^{\beta-1} \cdot 3$$
$$= \alpha 2t(t^2)^{\alpha-1}(3t)^\beta + 3\beta(t^2)^\alpha(3t)^{\beta-1}$$

Notice we could have solved this another way. We could have first written $g(t)$ explicitly in terms of t,

$$g(t) = f(x(t), y(t)) = x(t)^\alpha y(t)^\beta = (t^2)^\alpha (3t)^\beta$$

We leave it as an exercise to differentiate this using the power and product rules, and you will obtain the same as the above result.

Total vs Partial Derivative

One common point of confusion is the difference between the total derivative and the partial derivative. After all, the difference between $\frac{\partial f}{\partial x}$ and $\frac{df}{dx}$ seems to be just one of notation, but they are not in general the same for a multivariable function.

The partial derivative, $\frac{\partial f}{\partial x}$ tells you the *direct* effect of a small change in the argument *named* x on the value of the function f, while the total derivative tells you the *total* effect (including direct and indirect effects) of the variable x. That is, supposed we hvae a multivariate function, $f(x,y)$. However, suppose that y is a function of x, so $y = y(x)$. Then our function is really $f(x, y(x))$. To compute $\frac{\partial f}{\partial x}$, we would differentiate with respect to the first variable while holding y constant. That is, even though a change in x affects the value of y which would affect the value of f, we ignore this change when taking the partial derivative. Hence, we only find the direct, *partial*, effect of x on the function f. For the total derivative, on the other hand, we would differentiate with respect to x

while taking into account the fact that x enters into y, and we would need to use the chain rule. This tell us the *total* effect of x on the function f, including when x is an input to other variables.

To see the difference more explicitly, let's consider the example of the two variable function,

$$f(x, y) = x \cdot y^2$$

Where the name of the first input is x, and the name of the second input is y. Now we can consider evaluating f at some value x for both inputs, so $f(x, x)$. What is the partial derivative with respect to x and what is the total derivative?

In this case, the partial derivative with respect to x at the point x, $\frac{\partial f}{\partial x}$, is given by differenatiating with respect to the first variable, and then plugging in the values of x. To see this, we can first treat it as $f(x, y)$, and differentiate with respect to x holding y fixed,

$$\frac{\partial f}{\partial x}(x, y) = y^2$$

Next, we plug in the x for both the first and second inputs, so that

$$\frac{\partial f}{\partial x}(x, x) = x^2$$

Compare that to how we take the total derivative. In this case, we first plug in x for both the first and second inputs, so that $f(x, x) = x \cdot x^2 = x^3$, and then differentiate this totally with respect to x, so

$$\frac{df}{dx} = 3x^2$$

We can also use the multivariable chain rule.

$$\frac{df(x, x)}{dx} = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx}$$
$$= (x^2)(1) + (x)(2x)$$
$$= 3x^2$$

One useful way to think of the partial derivative, $\frac{\partial f}{\partial x}$, is as differentiating with respect to the variable named x first, and then plugging in a specific value of x, while the total derivative is first plugging in a specific value of x, and then differentiating with respect to that value.

Implicit Differentiation

So far, we have covered how to differentiate functions when we know their explicit form, but one reasonable question is whether or not we can differentiate functions without needing to solve for the function explicitly. This leads us to **implicit differentiation**, which allows us to differentiate a function while only knowing equality conditions that it satisfies. In some calculus classes, this may have come in the context of solving a related rates problem.

Consider the area of a circle: $A = \pi r^2$. Suppose we want to find out how the radius must change for a small change in the area of the circle. That is, we want to find $\frac{dr}{dA}$. One way to do this would be to rearrange and solve for r in terms of A, and then differentiate,

$$r = \sqrt{\frac{A}{\pi}}$$
$$\frac{dr}{dA} = \frac{1}{2\sqrt{\pi A}}$$

However, another approach we might take is to differentiate both sides of the equation $A = \pi r^2$ with respect to A, treating r as a function of A, and then solving for $\frac{dr}{dA}$ when it appears as a result of the chain rule,

$$\frac{d}{dA}(A) = \frac{d}{dA}(\pi r^2)$$
$$1 = \pi 2r \frac{dr}{dA}$$
$$\frac{1}{2\pi r} = \frac{dr}{dA}$$

This admits an easier solution, and more important, tells us how r changes with respect to A as a function of r, without as ever needing to solve for r in terms of A explicitly. Of course, you can verify that if we plug in our above expression of r in terms of A for the derivative, then we obtain the same result.

However, importantly in this case, only r was a function of the area A. Consider a similar problem of finding the volume of a cylinder, $V = \pi r^2 h$. In this case, we would not know how r changes with respect to V because h is also a function of V and we do not know how h changes with respect to V. Essentially, the problem is that we have two unknowns but only one equation. To solve explicitly, we would need another equation. However, using implicit differentiation, we can still obtain some useful information. Once again differentiating both sides with respect to V yields and solving for $\frac{dV}{dr}$ yields

$$\frac{dr}{dV} = \frac{\pi^{-1} - r^2 \frac{dh}{dV}}{h2r}$$

Notice here that we can tell how $\frac{dr}{dV}$ is related to each of the other terms. For example, if $\frac{dh}{dV}$ is large, then $\frac{dr}{dV}$ is smaller because the change in volume is mostly accounted for in

the height. Similar analyses on economic variables can help us obtain useful insights in terms of how two variables must be related to each other, even when we cannot solve explicitly for the functions or the derivatives.

Notation for derivatives

There are two commonly used notations for derivatives that are commonly used in economics in general, and that we will use throughout this book. The two are simply different ways of expressing the same mathematical concept, but occasionally one is used over the other for the sake of convenience or aesthetics. We will describe both notations for the derivative here.

Leibnitz Notation

So far, the notation we have used for the derivative is called **Leibnitz notation**. This notation treats the derivative as an operator, $\frac{d}{dx}$ or $\frac{\partial}{\partial x}$ applied to the function f. So in the single variable case, we write the derivative of a function f with respect to x as,

$$\frac{d}{dx}f = \frac{df}{dx}$$

Similarly, if we have a multivariable function, then the partial derivative of f with respect to the variable named x is,

$$\frac{\partial}{\partial x}f = \frac{\partial f}{\partial x}$$

Note however that derivatives written in this notation are *functions*, not values. If we want to obtain the derivative at a particular point, we need to plug in its argument. So to express the derivative of f with respect to x evaluated at some point a, we would write

$$\frac{df}{dx}(a) = \frac{df(a)}{dx}$$

To express higher order derivatives, we simply apply the operators multiple times. For example, the second derivative is really

$$\left(\frac{d}{dx}\right)\left(\frac{d}{dx}\right)f = \left(\frac{d}{dx}\right)^2 f = \frac{d^2}{dx^2}f = \frac{d^2 f}{dx^2}$$

In general, we write the nth derivative of f with respect to x as,

$$\frac{d^n f}{dx^n}$$

Similar notation applies for the partial derivative.

Leibnitz notation has a particularly nice interpretation as a fraction of the infinitesimal quantities $df(x)$ divided by dx. In this interpretation, we treat dx as some very small quantity, with $df(x) = f(x + dx) - f(x)$. So, $\frac{df(x)}{dx}$ is a very small change in f divided by a small change in x, which is the traditional slope interpretation of the derivative. It is important to note however that while this interpretation is useful and can lend to much intuition, it is not formally rigorous (at least without some non-standard analysis).

Lagrange Notation

You may be familiar with another notation for the derivative known as **Lagrange notation**. In the single variable case, the derivative of f with respect to x is written as

$$f'(x) = \frac{df}{dx}(x)$$

The second derivative is written as

$$f''(x) = \frac{d^2 f}{dx^2}$$

And in general, the nth derivative is written as

$$f^{(n)}(x) = \frac{d^n f}{dx^n}(x)$$

In the multivariable case, we use a subscript to denote which variable we are differentiating with respect to. Suppose we have a multivariable function, $f(x_1, \ldots, x_n)$. The derivative of f with respect to x_k is expressed as

$$f_{x_k} = \frac{\partial f}{\partial x_k}$$

Note however, that the subscript is merely the name of the variable. It is common to express the partial derivative of f with respect to the first, second, or kth argument, in which case we subscript with the index of the argument. For example, the partial derivative of f with respect to the first argument is written f_1, and the partial derivative of f with respect to the kth argument is written f_k.

If we wish to express higher order partial derivatives or cross-partials, then we simply add to the subscripts. So the partial derivative of f first with respect to x and then with respect to y would be written,

$$f_{xy} = \frac{\partial^2 f}{\partial y \partial x}$$

While Lagrange notation is often more concise and cleaner than Leibnitz notation, it is less expressive. In particular, when we use Lagrange notation we are always expressing the partial derivative rather than the total derivative. However, Lagrange notation is

often preferable with more complicated expressions involving many partial derivatives.

When to use each

In general there are no set rules on when you should prefer Leibnitz notation or Lagrange's notation, and is purely a question of convenience. In this text, we will use a mix of both in order to present expressions in what we believe to be the clearest way possible. However, a good rule of thumb is to prefer Lagrange notation when you are differentiating an explicit function with respect to one of its inputs, while using Leibnitz notation when expressing the relationship between two quantities that are implicit functions of each other.

2.2 Optimization

Agents in economics are generally assumed to be optimizing an objective function, and derivatives offer us a convenient mathematical tool for optimization of differentiable functions. There are generally two types of optimization functions: unconstrained optimization and constrained optimization, both of which can be solved with differentiation given the approproriate conditions.

Unconstrained Optimization

The most basic type of optimization is **unconstrained optimization**, which seeks to optimize some objective function f without any restrictions on what values its arguments can take. We will assume that f is a function of n variables, $x_1, \ldots, x_n$, which in vector notation is $\mathbf{x}$, and that it returns a real number. We will also assume that f is twice continuously differentiable (has two continuous derivatives).

We want to find some way of characterizing the value $\mathbf{x}^*$ that maximizes f. The problem we want to solve is therefore,

$$\max_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

In order for $\mathbf{x}^*$ to characterize an optimum, there are two conditions that must be satisfied: the first and second order conditions.

First order conditions

The **first order conditions** state that in order for $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$ to be a local optimum, we require that the partial derivative of f with respect to each input x_k to be

equal to 0,

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^*) = 0$$

$$\vdots$$

$$\frac{\partial f}{\partial x_k}(\mathbf{x}^*) = 0$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n}(\mathbf{x}^*) = 0$$

This is also often known as the **first derivative test**. This is often written in terms of the **gradient** of f, $\nabla f = (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_k}, \ldots, \frac{\partial f}{\partial x_n}) = 0$.

To see why this must be the cast, consider the alternatives. Suppose that $\frac{\partial f}{\partial x_k}(\mathbf{x}^*) > 0$ for some x_k. In that case, we could "nudge" x_k^* to be slightly larger, and because the partial derivative is positive, the value of the function f will slightly increase, which means that $f(\mathbf{x}^*)$ is not an optimum. Similarly, if $\frac{\partial f}{\partial x_k}(\mathbf{x}^*) < 0$, we could decrease x_k^* by a small amount and increase the value of f. So, in order for $\mathbf{x}^*$ to achieve the optimal value of f, it must be that the partial derivatives are 0.

Note that this is a necessary condition, but not a sufficient condition. For example, in the single variable case, the function $f(x) = x^3$ has 0 derivative at $x = 0$, but it is clearly not an optimum.
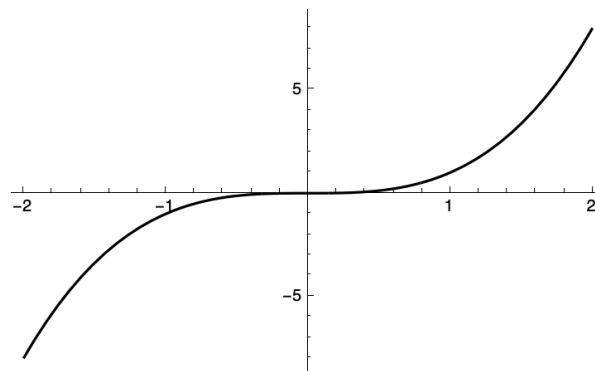
Moreover, the first order conditions do not distinguish between a maximum and a minimum, and are not sufficient to show that the optimum is global rather than local. In order to verify that $f(\mathbf{x}^*)$ is a global maximum, an additional condition must be satisfied.
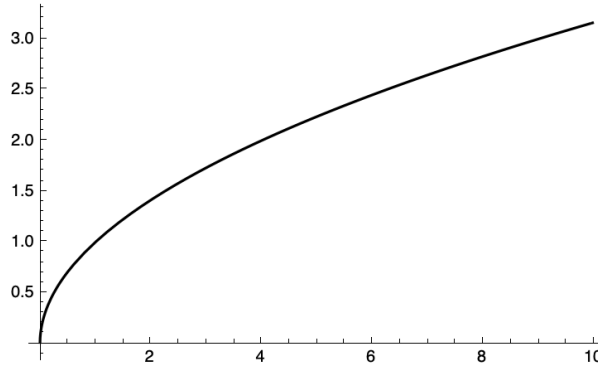
Second order conditions

The **second order conditions** (SOC) for a maximum are conditions on the second derivative of f. We will focus on the single variable and two-variable case, as higher dimensional second order conditions are beyond the scope of this course.

The **single variable second order conditions** for a global maximum are given by

$$\frac{d^2 f}{dx^2}(x) < 0 \text{ for all } x$$

In other words, the f must be a globally concave function.

To see why this is the case, we can think about what it means for $\frac{d^2 f}{dx^2}(x) < 0$. Since $\frac{d^2 f}{dx^2}$ is the derivative of $\frac{df}{dx}$, it means that $\frac{df}{dx}(x)$ is decreasing at this point. However, the first order conditions tell us that $\frac{df}{dx}(x) = 0$ at the maximum. Since $\frac{df}{dx}$ is decreasing at this point, the derivative at a slightly greater value of x must be negative, so the value would be lower. Similarly, the derivative at a slightly lower value of x must be positive, which menas that the value can increase. Moreover, because the $\frac{d^2 f}{dx^2}(x) < 0$ globally, then it must be that once $\frac{df}{dx}(x) < 0$, it must be negative for all greater values of x.

Similarly, if we are searching for a global minimum, then we require that $\frac{d^2 f}{dx^2}(x) > 0$ for all x.

However, the conditions are slightly more complicated for functions of more than one variable. The **two variable second order conditions** for a global maximum of a function $f(x, y)$ are given by

$$\frac{\partial^2 f}{\partial x^2}(x, y) < 0$$
$$\frac{\partial^2 f}{\partial x^2}(x, y)\frac{\partial^2 f}{\partial y^2}(x, y) - \left(\frac{\partial^2 f}{\partial x \partial y}(x, y)\right)^2 > 0 \text{ for all } x, y$$

One thing to notice is that the above inequalities also imply that $\frac{\partial^2 f}{\partial y^2} < 0$, so the function must be concave in both variables. However, $\frac{\partial^2 f}{\partial y^2} < 0$ and $\frac{\partial^2 f}{\partial x^2} < 0$ alone are not sufficient to achieve a local maximum. For a global minimum, we replace the first inequality with $\frac{\partial^2 f}{\partial x^2} > 0$.

For functions of more than 2 variables, we require that the **Hessian** matrix of f is negative-semidefinite. We generally will not need to deal with functions of more than 2 variables in this course, and so will not address these conditions here.

The first and second order conditions are sufficient and necessary conditions for the characterization of a global maximum in an unconstrained maximization problem.

However, our problem will often have additional constraints that must be satisfied, and so we can not maximize using any set of inputs.

Constrained Maximization

Many optimization problems require that you optimize an objective function f while satisfying some constraint, $g(\mathbf{x}) = c$, where g is a function with the same inputs as f and c is some constant. In economics, this might be a consumer who has to spend c dollars or a firm that is required to produce c units of good. Formally, we write a constrained optimization as

$$\max_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \text{ s.t. } g(\mathbf{x}) = c$$

Notice that in this case we only have a single constraint $g(\mathbf{x}) = c$. In general there are problems with multiple constraints, but we will not address those cases as in this course we only handle cases with a single constraint.

To solve the constrained optimization problem, we use the **Lagrangian**, which is defined as,

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda(g(\mathbf{x}) - c)$$

λ is known as the Lagrange Multiplier, and is an added variable to help us handle the constraint.

To find the constrained maximum, we must satisfy the **constrained first order conditions**, which require that the partial derivatives of $\mathcal{L}$ with respect to each input is 0,

$$\frac{\partial \mathcal{L}}{\partial x_1}(\mathbf{x}^*, \lambda^*) = 0$$

$$\vdots$$

$$\frac{\partial \mathcal{L}}{\partial x_n}(\mathbf{x}^*, \lambda^*) = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda}(\mathbf{x}^*, \lambda^*) = 0$$

Notice that we treat the Lagrangian almost as its own objetive function. We can compute the partial derivatives and obtain more direct first order conditions:

$$\frac{\partial \mathcal{L}}{\partial x_k} = \frac{\partial f}{\partial x_k} - \lambda \frac{\partial g}{\partial x_k} = 0 \iff \frac{\partial f}{\partial x_k} = \lambda \frac{\partial g}{\partial x_k} \text{ for each } k$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = g(x) - c = 0 \iff g(x) = c$$

For students who took a multivariable calculus course that did not use the Lagrangian,

this may be a more familiar set of conditions. Namely, $\frac{\partial \mathcal{L}}{\partial \lambda} = 0$ ensures that the constraint is satisfied, while the Lagrange multiplier ensures that $\nabla f = \lambda \nabla g$, which says that the partial derivatives of f and g are a constant proportion of each other.

While this is the first order condition, the full second order conditions for constrained optimization are beyond the scope of this course. Instead, it is sufficient to note that if f is a strictly increasing function and f is concave, then it is a sufficient condition (although not necessary) to find a global optimum.

2.3 Important Properties and Notation

The differentiation and optimization techniques above will be the foundation of mathematics necessary for this course, there are also some important mathematical properties and notations that we will use throughout this course and will be important to know. We list them here:

Convexity So far, we have described convex functions as a function with a positive second derivative. However, this is not the formal definition of a convex function.

Definition. Given two points, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and some $\lambda \in (0, 1)$, a **convex combination** of $\mathbf{x}$ and $\mathbf{y}$ is

$$\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}.$$

Graphically, a convex combination is any point on the line segment between points $\mathbf{x}$ and $\mathbf{y}$.
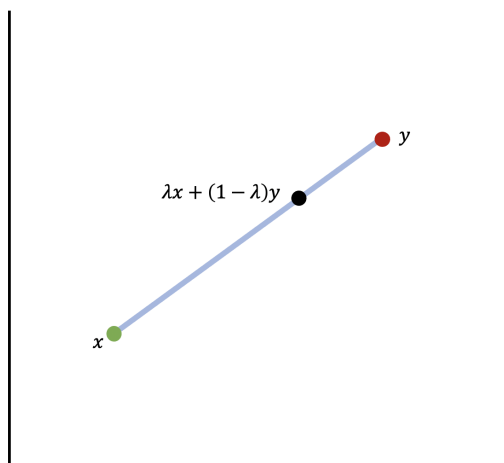
Then we can formally define a convex function,

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is **convex** if for any $\lambda \in [0, 1]$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda)\mathbf{y}$$

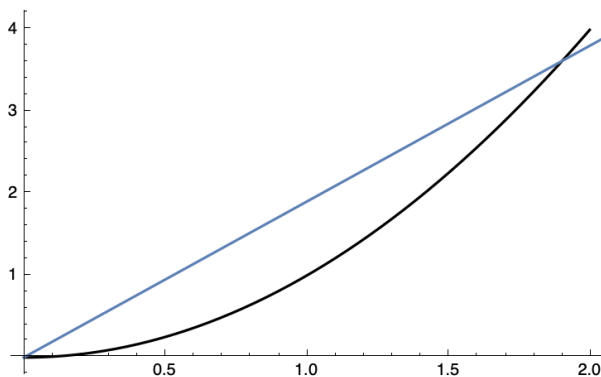It can be show that if f is twice differentiable, then f is convex if and only if $\frac{d^2 f}{dx^2}(x) \geq 0$ for any x. Graphically, this says that for a convex function, a line segment connecting two points on the function is greater at every point than the function itself. The inequality is reversed for concave functions. A function that is both concave and convex is linear.

Monotonic A monotonic function is any order preserving function. That is, f is a monotonically increasing function if $f(x) < f(y)$ for any $x < y$.

Sets A set is simply a collection (possibly infinite) of mathematical objects. For a set X, we write that an element x is in X by $x \in X$. We say that Y is a subset of X if every element of Y is also in X, and we write $Y \subset X$.

Summation Suppose we have a sum $x_1 + x_2 + \cdots + x_n$. We often write this as $\sum_{i=1}^{n} x_i$. For any sequences $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ and constants α, β, we have

$$\sum_{i=1}^{n}(\alpha x_i + \beta y_i) = \alpha \sum_{i=1}^{n} x_i + \beta \sum_{i=1}^{n} y_i$$

Exponents Here are some common properties of exponents that should be known,

- $a^{x+y} = a^x a^y$
- $(a^x)^y = a^{xy}$
- $a^{-x} = \frac{1}{a^x}$
- $(ab)^x = a^x b^x$
- $a^{x/y} = \sqrt[y]{a^x}$
- e^x is a monotonically increasing function

Logarithm We use log to refer to the natural logarithm, and here are some important properties:

- $e^{\log x} = x$
- $\log(e^x) = x$

- $\log(1) = 0$

- If $0 < x < 1$, $\log x < 0$

- If $x > 1$, $\log x > 0$

- If $x < 0$, $\log x$ is undefined

- $\log(xy) = \log(x) + \log(y)$

- $\log(x/y) = \log(x) - \log(y)$

- $\log(x^y) = y \log(x)$

- $\log(x)$ is a monotonically increasing function

2.4 A Guided Example

We can finally use the concepts and ideas developed in these first two chapters to work through an example of an economic model. You should not expect to understand every step of this model on a first read-through or all of the terms that we use. This example is simply to give you an idea for how the mathematical techniques we've described so far are used to model a problem, and to serve as a reference for later in the course for how to approach such problems generally.

Example. Model how a member of the workforce chooses how many hours to work. What factors influence this choice?

It should not be surprising that this question is one that real economists are often interested in. Understanding this problem might help governments think about policies like tax structures or minimum wages, or it might help firms decide how much to pay their workers. We will present a simple model of this problem and highlight the key modeling steps that apply to a much broader set of modeling problems.

1. Write down the optimization problem.   The first step is to determine who the decision-making agent is, and what their optimization function is. In this case, the agent is the worker, and we will say their goal is to maximize their utility.[2]

Next, in order to write down the optimization problem, we need to be able to represent utility as a function of some arguments. It is impossible to write down a model that encapsulates everything a worker might consider in this decision, but also a model that is too simple prevents meaningful insights (e.g. if workers only care about income, they would spend all their time working).

We will assume that utility is only a function of income (Y) and leisure (L), so utility is a function $U(Y, L)$. We could assign a functional form like $U(Y, L) = Y + L$, but

---

[2]We will discuss this concept more later, but for now, think of utility as a quantification of someone's level of satisfaction.

for the sake of generality, it often helps to start with a general function, think about the characteristics it should have, and only introduce additional structure when it is reasonable and needed.[3] In this case, we can assume that U is increasing in Y and L everywhere, so $\frac{\partial U}{\partial Y}, \frac{\partial U}{\partial L} > 0$ everywhere. This has the interpretation that, all else equal, you would always prefer more income or more leisure over what you currently have. We also assume that $\frac{\partial^2 U}{\partial Y^2}, \frac{\partial^2 U}{\partial L^2} < 0$; this has the interpretation that the marginal utility of income and leisure are both diminishing.

We now need to consider what the choice variables are in this optimization problem. A worker probably does not choose their income directly; rather, we can assume that they choose the number of hours they work per day (H), and the rest of the day is devoted to leisure. One approach is to say that H and L are our choice variables. If we assume that $Y = wH$ for some hourly wage w (which is exogenous), we could write down our utility function as $U(wH, L)$. Since we are constrained by the number of hours in a day, we can now write down our constrained optimization problem as

$$\max_{H,L \geq 0} U(wH, L) \quad \text{s.t.} \quad H + L = 24.$$

It is a useful exercise to solve this maximization problem using the Lagrangian, but a useful trick that will often come in handy in this class is to substitute the constraint directly into the optimization function. That is, we note that $L = 24 - H$, so we can write our utility function as $U(wH, 24 - H)$. Now, we only have one choice variable, H, and our constrained optimization problem is now the unconstrained optimization problem

$$\max_{H \in [0,24]} U(wH, 24 - H).$$

2. Solve the optimization problem. We initially wrote our optimization function as $U(Y, L)$, but we can equivalently think of this as a function of our choice variable H, where $V(H) \equiv U(wH, 24 - H)$.[4] We assume that $U(Y, L)$ is twice continuously differentiable, which implies $V(H)$ is twice continuously differentiable as well.

We want to solve for the optimal solution H^* for the choice variable (as a function of the exogenous variable w). If there is an **interior solution** $H^* \in (0, 24)$ that maximizes the optimization function (as opposed to the solution being $H^* = 0$ or $H^* = 24$), then it must satisfy the first-order condition

$$\frac{d}{dH} V(H^*) = 0.$$

---

[3]Exercise: What functional forms might make sense? We will discuss a few common examples later in this course.

[4]Notation: the $\equiv$ sign means 'is equivalent to' and is used when two names refer to the same fundamental quantity by definition, which we might want to differentiate from the cases where the $=$ sign simply refers to two different fundamental quantities that are equal.

To solve, we substitute our definition of $V(H)$ to get

$$\frac{d}{dH}U(wH^*, 24 - H^*) = 0.$$

We now apply the multivariate chain rule to get

$$w\frac{\partial U}{\partial Y} - \frac{\partial U}{\partial L} = 0,$$

or if you prefer,

$$\boxed{w\frac{\partial U}{\partial Y} = \frac{\partial U}{\partial L}}.\text{[5]}$$

    This is our solution! You might have been expecting something of the form

$$H^* = \text{stuff.}$$

This type of solution would be called an **explicit solution** for H^*. The solution we gave instead is called an **implicit solution.** It is very important to note that this solution is just as acceptable. Recall that $\frac{\partial U}{\partial Y}$ and $\frac{\partial U}{\partial L}$ are functions of H; thus, the equation we gave pins down the value of H^* that gives us our optimum. Intuitively, this equation tells us that you should work until your marginal utility of leisure and income are equal; this is the point where if you had an extra second in your day, you would be indifferent between spending it on work and spending it on leisure.

3. Check your solution.   There is some value of H^* that satisfies the implicit solution we gave, but how do we know that it actually maximizes the optimization problem we initially gave? There are two important checks we need to perform.

    First, how do we know that the solution for the FOC is a maximum and not a minimum? We need to check the second-order condition (SOC)

$$\frac{d^2}{dH^2}V(H^*) < 0.$$

If we assume that V is concave in H, then this condition is automatically satisfied. If instead we needed to compute this, we would write down what we had before

$$\begin{aligned}
V'(H) &= \frac{d}{dH}U(wH, 24 - H)\\
&= w\frac{\partial U}{\partial Y}(wH, 24 - H) - \frac{\partial U}{\partial L}(wH, 24 - H)\\
&= wU_Y(wH, 24 - H) - U_L(wH, 24 - H).
\end{aligned}$$

---

[5]We suppress the arguments of $\frac{\partial U}{\partial Y}$ and $\frac{\partial U}{\partial L}$ for clarity, but remember that they are functions of H.

We would then differentiate again to get

$$
\begin{aligned}
V''(H) &= w\frac{d}{dH}U_Y(wH, 24 - H) - \frac{d}{dH}U_L(wH, 24 - H) \\
&= w^2 U_{YY} - wU_{YL} - wU_{LY} + U_{LL} \\
&= w^2 U_{YY} - 2wU_{YL} + U_{LL}.
\end{aligned}
$$

Our SOC is satisfied when this value is negative at H^*.

Second, we need to check for **corner solutions**. These are potential solutions where the optimal value of the choice variable is on the boundary of the constraints we set for it. In our example, our implicit solution pins down some real value for H^*, but what if it were negative or greater than 24? Alternatively, what if our solution for H^* was only a local maximum, but the global maximum in our interval were actually on the boundaries of the interval? Thus, we should check that $H^* \in [0, 24]$, and we should plug in $H = 0$ and $H = 24$ into our optimization function $V(H)$ to check that their values are less than $V(H^*)$.

4. Take comparative statics. Recall that comparative statics describe how the optimal values of the choice variables change when the exogenous variables change; at the end of the day, this quantity is what we are really interested in and is useful in policy discussions. Our only exogenous variable is w, and we are interested in finding $\frac{\partial H^*}{\partial w}$ (recall that H^* is a function of the exogenous variables). Our solution for H^*, which we can think of as $H^*(w)$, from our FOCs was

$$
w\frac{\partial U}{\partial Y} = \frac{\partial U}{\partial L}.
$$

If we include the arguments of $\frac{\partial U}{\partial Y}$ and $\frac{\partial U}{\partial L}$, we can write this as

$$
wU_Y\left(wH^*(w), 24 - H^*(w)\right) = U_L\left(wH^*(w), 24 - H^*(w)\right).
$$

In order to get $\frac{\partial H^*}{\partial w}$ out of this, we use the **Implicit Function Theorem** to take the derivative of both sides of the equation, yielding

$$
\frac{d}{dw}\left[wU_Y\left(wH^*(w), 24 - H^*(w)\right)\right] = \frac{d}{dw}U_L\left(wH^*(w), 24 - H^*(w)\right).
$$

We use the chain rule and the product rule to get (note that we suppress arguments of functions again for clarity)

$$
U_Y + wU_{YY}H^* + w\left(U_{YY}w - U_{YL}\right)\frac{\partial H^*}{\partial w} = U_{LY}H^* + \left(U_{LY}w - U_{LL}\right)\frac{\partial H^*}{\partial w}.
$$

We can rearrange to solve for

$$\frac{\partial H^*}{\partial w} = \frac{U_Y + wH^*U_{YY} - H^*U_{YL}}{-\left(w^2 U_{YY} - 2wU_{YL} + U_{LL}\right)}.$$

Note that we got this without solving for $H^*(w)$ explicitly! The naive approach would have been to find the explicit solution and then manually take its derivative. Here, we see that using an implicit solution and applying the Implicit Function Theorem leads to a much faster solution.

5. Interpret your results. What we really care about is the sign of $\frac{\partial H^*}{\partial w}$. If wages increase, do people work more or less? The denominator of our solution must be positive, since it is exactly the expression from our SOC. The sign of the numerator depends on the magnitudes of the different terms. We said in our initial assumptions that U_Y is positive and that U_{YY} is negative. What about the sign of U_{YL}? Maybe it's positive: if you are richer, then maybe the marginal utility of leisure is higher, since you can afford more expensive hobbies. Maybe it's negative: if you are richer, maybe you lose appreciation for the everyday moments in life, so the marginal utility of leisure is actually lower. How might we resolve this question with data?

Part II

Competitive Firms

3 Firms with a Single Input

We start with one of the simplest economic models: a firm in a perfectly competitive market with a single input. Firms are one of the most basic components of the economy. Firms purchase inputs, convert those inputs into outputs via a production function, and then sell those goods to make a profit. This also gives firms a clear objective function: profit. To make things simple, we assume that the firm only requires one input to produce their good, so the only choice that the firm makes is how much input to purchase.

3.1 Model setup

We will formalize a mathematical model of the firm. We assume that the firm can choose to hire L units of labor at a wage w, which is the price per unit of labor. The firm has a production function, $f(L)$, which takes the units of labor as an input, and returns some unit of product. We assume that the production function is continuous and twice differentiable, increasing, $f'(L) > 0$, and concave $f''(L) < 0$, for any of L. We will first establish this production function as The firm can then sell each unit of product at a price p. The firm takes both p and w as exogenous variables. We can then define the firm's profit, π, as follows:

$$\pi(L; p, w) = pf(L) - wL \tag{3.1}$$

Notice that our profit function has one endogenous input and two exogenous inputs. L, the units of labor hired, is the choice variable for the firm. p, the price of the product, and w, the wage cost of labor, are both exogenous variables. While we write them explicitly here, we will often only write π and the arguments to it are implicit.

While this seems like a fairly simple model, there are some pretty important assumptions underlying it.

Assumptions

Perfectly competitive market for output Notice that the firm treats the price p as exogenous. That is, no matter how much the firm produces, they can always sell goods at price p, and *only* at price p. This means, first, that the firm is a **price-taker**, which means that they cannot set a price p that differs from the market price p. The underlying assumption here is that there are enough other

firms that if this firm were to raise its price, all of the customers would buy from other firms and our firm would sell 0. This assumption also entails that the amount our firm produces does not affect the market price, which can be taken to mean that there are many other firms producing a lot of the same good, so our firm's decisions do not have a noticeable effect on p.

Perfectly competitive labor market Similar to the above market for goods, we also assume that the market for labor is perfectly competitive. That is, the firm can only hire at the wage w, and that no matter how much labor the firm hires, the wage will not change.

No liquidity constraints We assume that the firm has the ability to hire as much labor as they want, and all that matters is the final profit. That is, the firm does not have some fixed budget for labor at the beginning. This can be thought of as a firm's ability to borrow at zero interest to finance labor so long as the loan is paid back. This assumption is key to the firm problem, as it allows us to deal with an unconstrained maximization problem rather than requiring a budget constraint for the firm.

Diminishing marginal returns to consumption This was expressed mathematically as $f''(L) < 0$. In real terms, this says that each additional unit of labor contributes less additional production than the previous unit of labor did, and represents a sort of "too many cooks in the kitchen" effect. Notice however that we will assume $f'(L) > 0$, so even if each additional unit of labor contributes less additional output than the previous unit, adding more units of labor can never make us produce less output.

One reasonable question to ask with all of these assumptions in place is whether they are realistic assumptions. The answer is that they probably are not all perfectly realistic. However, there are cases where these assumptions might be close enough. Consider the market for corn, for example. Each individual farmer's corn production has a negligible effect on the market as a whole, and they have enough money every year to grow as much corn as is profitable. However, we will see that even if these assumptions are not all realistic, they help simplify the model so that we can solve it and gain some useful insights about the mechanics of this economy.

3.2 Solving the model

Now that we have setup the model and established the underlying assumptions, we can begin to solve the model. Our objective function is π, and our choice variable is L, so we

can write our maximization problem as follows:

$$\max_L \pi(L; p, w) = \max_L pf(L) - wL \tag{3.2}$$

First order conditions

Our first order condition is that the derivative of the objective function with respect to our choice variable is 0, which is

$$\frac{\partial \pi}{\partial L} = 0 \tag{3.3}$$

We can plug in our function for π using our definition in 3.1 and differentiate with respect to L to obtain,

$$\frac{\partial \pi}{\partial L} = pf'(L) - w = 0 \tag{3.4}$$

This is our first order condition for an optimal profit. With some rearranging, we obtain that

$$pf'(L) = w \tag{3.5}$$

Notice that this yields a very useful interpretation. The left hand side is the marginal revenue from an additional unit of labor. That is, for a small amount more labor, we produce $f'(L)$ more goods, which are sold at a price p. On the right hand side is the marginal cost for an additional unit of labor, because for a small amount more labor, we pay that labor a wage w. What this tells us is that when a firm is profit maximizing, **marginal revenue equals marginal cost**.

Notice further that this implicitly defines our optimal choice of labor, which we will denote L^*. Because we do not know the functional form of f, we cannot yet write an explicit definition of L^*, but it is implicitly defined by the first order condition,

$$pf'(L^*) = w \tag{3.6}$$

Observe further that the value of L^* depends on the values of p and w, so we can express it as a function of the exogenous variables, $L^*(p, w)$.

Second order conditions

In order to verify that L^* is indeed a global maximum, and not a local maximum or even a minimum, we need to verify the second order conditions. Namely, that

$$\frac{\partial^2 \pi}{\partial L^2} < 0$$

To do so, we can just differentiate the first derivative from 3.3 with respect to L again to obtain,

$$\frac{\partial^2 \pi}{\partial L^2} = p f''(L) \tag{3.7}$$

We need the above derivative to hold with respect to any L. However, we assumed that $f''(L) < 0$, and since $p > 0$, then we have that $p f''(L)(L) < 0$, which satisfies our second order conditions.

This tells us that L^* defines a global maximum, which means that the firm is indeed profit maximizing!

Some intuition

While the above provides the mathematical technique for how to maximize, it is useful to get some intuition about what is actually happening here. To do so, we can think about the firm's decision process.

Suppose that you are running the firm but do not know any calculus. However, you do know how much you will produce for a given amount of labor and how much money you will make. One way you might decide how much labor to hire is to think what will happen if you hire one additional worker. At each point, you ask yourself whether you will make more money by hiring an additional worker or less money by hiring an additional worker. What determines this? Whether the contribution of that additional worker is greater than the additional cost of hiring that worker. Mathematically, you would hire an additional worker so long as, $p f'(L) > w$.

However, you also know that because each additional worker's contribution is less than the last worker hired, at some point you will stop hiring workers. This is the second order condition at work. So, if the increments are small enough, you will eventually reach a point where hiring an additional worker makes no difference. That is, that $p f'(L) = w$, which is the first order condition. At this point, you should stop hiring workers because any more workers that you hire will cost more than they produce.

At some level, all economic optimization methods boil down to this process. Thinking about what will happen if increase some quantity by an infinitesimal amount, and requiring that doing so makes no difference.

3.3 Specific production functions

In the above section we showed how to set up and solve for a general production function, f. However, because of this generality, we were unable to obtain an explicit formulation of L^*. To do so, we need to specify the functional form of f, and we offer some examples here of specific functional forms that f could take.

Single-variable Cobb-Douglas

One of the most common production functions used in economics is the Cobb-Douglas production function. Although typically the production function is in multiple variables, we present a single variable version of that production function here.

$$f(L) = L^\alpha,\ 0 < \alpha < 1 \tag{3.8}$$

In this expression, we can interpret α as a constant measuring the productivity of labor. We will see that $0 < \alpha < 1$ is necessary to ensure that the first and second order conditions hold. We can write the profit function,

$$\pi(L) = pL^\alpha - wL \tag{3.9}$$

We will now solve for the optimum.

First order condition The first order condition for a maximum is

$$\frac{\partial \pi}{\partial L} = p\alpha L^{\alpha-1} - w = 0 \tag{3.10}$$

Rearranging to solve yields

$$L^* = \left(\frac{w}{p\alpha}\right)^{\frac{1}{\alpha-1}} = \left(\frac{p\alpha}{w}\right)^{\frac{1}{1-\alpha}} \tag{3.11}$$

Where the second equality holds by taking the reciprocal of the inside term and negating the exponent.

Second order conditions We should verify that L^* is indeed a maximum by checking the second order conditions. We can take the second derivative of π with respect to L to obtain,

$$\frac{\partial^2 \pi}{\partial L^2} = p\alpha(\alpha - 1)L^{\alpha-2} \tag{3.12}$$

Now we can determine the sign of the above expression by examining each of the terms. $p > 0$ and $\alpha > 0$ by assumption. $L^{\alpha-2} > 0$ if we assume that we hire at least some labor. And $\alpha - 1 < 0$ by assumption that $\alpha < 1$. So, we multiply three positive terms and a negative term, which means that the entire expression is negative. This satisfies our second order conditions.

Log production

Another relatively simple production function is the log production function, which takes the form,

$$f(L) = \log(L) \tag{3.13}$$

We will solve this optimization problem in a slightly different way than we did for the Cobb-Douglas case. Because we have already solved for the general first and second order conditions in 3.2, we can just plug in our derivations.

First order conditions The only expression we need to plug into the generalized first order condition is $f'(L)$, which is,

$$f'(L) = \frac{d \log(L)}{dL} = \frac{1}{L}$$

Plugging this into 3.6 yields,

$$\frac{p}{L^*} = w \iff L^* = \frac{p}{w}$$

Second order conditions To test the second order conditions, all we need is to solve for $f''(L)$, which is,

$$f''(L) = \frac{d^2 \log(L)}{dL^2} = -\frac{1}{L^2}$$

Then plugging into the generalized second order conditions, 3.7, yields

$$\frac{\partial^2 \pi}{\partial L^2} = -\frac{p}{L^2} \tag{3.14}$$

Since $L^2 > 0, p > 0$, this must be negative, which satisfies the second order conditions.

3.4 Comparative statics

Now that we have solved for the optimal quantity of labor used for a given price and wage, we can see how that optimal quantity changes for a given change in exogenous variables. That is, we will take a **comparative static**. We will first do so in the specific case of a Cobb-Douglas production function, and then we will see how we can do so for a general production function.

Cobb-Douglas

Because we can solve for the optimal quantity of labor explicitly as a function of price and wage, we can consider what happens with a small change in either. Before we start, it will be useful for us to rewrite 3.11 by writing each of the terms as a product of individual exponents,

$$L^* = \left(\frac{p\alpha}{w}\right)^{\frac{1}{1-\alpha}} = \left(p^{\frac{1}{1-\alpha}}\right)\left(\alpha^{\frac{1}{1-\alpha}}\right)\left(w^{\frac{1}{\alpha-1}}\right) \tag{3.15}$$

This will make it easier for us to take derivatives with respect to each of the variables.

Price To determine the effect of a small change in the price of the good, we can take the derivative of L^* with respect to p. This yields,

$$\frac{dL^*}{dp} = \left(\frac{p^{\frac{\alpha}{1-\alpha}}}{1-\alpha}\right)\left(\alpha^{\frac{1}{1-\alpha}}\right)\left(w^{\frac{1}{\alpha-1}}\right)$$

You can check that each of the terms above is positive to obtain that the entire expression is positive. So, all else equal, an increase in price will lead to the firm using more labor. This makes sense because with a higher price, the marginal revenue for the good increases, which means that the firm should hire more labor to produce more output.

Wage We can follow a similar procedure as above, differentiating L^* with respect to the wage, w,

$$\frac{dL^*}{dw} = \left(\frac{p\alpha}{w}\right)^{\frac{1}{1-\alpha}} = \left(p^{\frac{1}{1-\alpha}}\right)\left(\alpha^{\frac{1}{1-\alpha}}\right)\left(\frac{w^{\frac{2-\alpha}{\alpha-1}}}{\alpha-1}\right)$$

Notice that once again, every term is positive with the exception of $\alpha - 1$, which is negative. So, multiplying a series of positive terms with a single negative term means that the entire expression is negative.

Implicit Function Theorem

While the above example tells us how price and wages affect the quantity of labor demanded in the specific case of the Cobb-Douglas production function, it seems that the specific form of the production function should not matter. After all, it would be very strange if there were a production function where prices increased and the quantity of labor decreased. But how can we determine this change if we do not know the form of f?

To gain some intuition for how we might approach this problem, let's look back at the first order condition,

$$pf'(L^*) = w$$

Now consider what happens if we increase w by a little bit. The right hand side has increased, so the left hand side must also increase by a small amount to ensure the equality holds. However, p is fixed, so it must be that the value of L^* has changed. This makes sense because recall that with the optimized choice of L, L^* is an implicit function of p and w, so the value of L^* changes. In particular, it must have decreased, because we know that $f'' < 0$, which means that f' decreases as L increases, so to make f' smaller, the value of L^* must decrease. Notice that we were able determine how L^* would change without having to know anything about the actual value of L^* or the form of f. This is the idea behind using implicit differentiation to calculate the change.

While this is an informal argument, we can formalize it with the **implicit function theorem**. While we do not offer a formal state of the theorem, it essentially states that if we have an equality of the form $f(x) = c$ for some constant c, and f is sufficiently "well-behaved", then we can express the value of x as a function of c. This also means that we can use implicit differentiation to solve for the relatioship between the two variables.

In the context of our problem, it tells us that by implicity differentiating both sides of the first order condition, we can determine how L^* changes with respect to wage and price, without even knowing the actual value of L^*!

Wage We will first formalize the above example of the wage change. To do so, we implicitly differentiate both sides of the FOC with respect to w.

$$\frac{d}{dw}\left(pf'(L^*(w))\right) = \frac{d}{dw}w$$

We can start by doing both sides separately. The derivative on the right hand side is simple,

$$\frac{d}{dw}(w) = 1$$

The left hand side requires us to use the chain rule. We explicitly write L^* as a function of w in this case to make things clear.

$$\frac{d}{dw}\left(pf'(L^*) = pf''(L^*)\frac{dL^*}{dw}\right.$$

Because the FOC must hold with equality, a small change in w must lead to both sides staying the same. So, we can set the above expressions equal to each other,

$$pf''(L^*)\frac{dL^*}{dw} = 1$$

Now, we can solve for $\frac{dL^*}{dw}$. Rearranging yields,

$$\frac{dL^*}{dw} = \frac{1}{pf''(L^*)}$$

We can also determine the sign of $\frac{dL^*}{dw}$. Observe that $f''(L^*) < 0$ by assumption on our production function, and $p > 0$, so we have that overall the expression must be negative. That is,

$$\frac{dL^*}{dw} < 0$$

This tells us that holding everything else constant, a small increase in the wage decreases the amount of labor that the firm demands.

Price We can perform a similar calculation with the price of the output good. Differentiate both sides of the FOC with respect to p,

$$\frac{d}{dp}\left(pf'(L^*)\right) = \frac{d}{dp}w$$
$$\implies pf''(L^*)\frac{dL^*}{dp} + f'(L^*) = 0$$

Since w is an exogenous variable, it does not depend on p so the right hand side is clearly 0. The left hand side has a product, so we use the product role to implicitly differentiate. So we can solve for, $\frac{dL^*}{dp}$,

$$\frac{dL^*}{dp} = -\frac{f'(L^*)}{pf''(L^*)}$$

Before reading further, try figuring out the sign of $\frac{dL^*}{dp}$ yourself. It should be a fairly straightforward exercise in examining the assumptions we made about f and your result should match with your intuitions.

By assumption, we have that $f'(L^*) > 0, p > 0$ and we know that $f''(L^*) < 0$, so the entire term must be positive. To check that this accords with our intutions, if the price of the good increases, the firm should produce more of that good because the marginal revenue has increased.

If you observe both of our comparative statics above, you may notice some similarities between them. For instance, the denominator in both is the same. It turns out that we can say generally how to take a comparative static for a given objective function. We show this below.

Proposition. *Given a twice-differentiable increasing and concave objective function $F(x; z)$, with choice variable x and exogenous variable z, and $x^*(z) = \text{argmax}_x F(x; z)$, then $\frac{dx^*}{dz} = -\frac{F_{xz}}{F_{xx}}(x^*(z); z)$.*

The above yields a general way of finding the comparative static of an optimized choice variable x^* with respect to an exogenous variable z for any given single variable optimization problem. We will prove the above result.

Proof. This proof follows straightforwardly from taking first order conditions and then using the multivariate chain rule to implicitly differentiate.

The first order condition for an optimal x is

$$F_x(x^*, z) = 0$$

Now, we totally differentiate both sides with respect to z. The right hand side is clearly 0. The left hand side uses the multivariate chain rule.

$$F_{xx}(x^*, z)\frac{dx^*}{dz} + F_{xz}(x^*, z) = 0$$

Rearranging for $\frac{dx^*}{dz}$ completes the proof,

$$\frac{dx^*}{dz} = -\frac{F_{xz}(x^*, z)}{F_{xx}(x^*, z)}$$

$\square$

One useful component to notice is that because we have assumed the second order conditions to hold, $F_{xx} < 0$, then the sign of the comparative static is the same as the sign of F_{xz}. This is a fairly powerful result that tells us how to very straightforwardly find the comparative statics for a single variable optimization problem. One important thing to note however, is that above, F would be akin to profit, π, in the firm's problem because it is the objective function. This does *not* hold for the production function f.

3.5 Value function

In all of the above, we find that we can write the optimal amount of labor demanded as a function of the exogenous variables. However, we could do the same for all of the other functions. For example, we could define the optimal production function in terms of the exogenous variables as well,

$$f^*(w, p) = f(L^*(w, p))$$

And we can do the same for the profit function to determine how much profit the firm makes in terms of the exogenous parameters,

$$\pi^*(w, p) = pf(L^*(w, p)) - wL^*(w, p)$$

Since these are all functions of the exogenous variables, we can take comparative statics of each of them by straightforward differentiation and the chain rule. For example, to

see how the quantity produced changes for an increase in price, we have

$$\frac{df^*}{dp} = \frac{df(L^*)}{dp} = f'(L^*)\frac{dL^*}{dp} = -\frac{f'(L^*)^2}{pf''(L^*)}$$

Where the last equality holds by plugging in the value of $\frac{dL^*}{dp}$ that we obtained previously.

Now we can do the same for the optimized profit function. This optimized version of the profit function, which we can denote as $\pi^*(w, p)$, is referred to as the value function. In general, the **value function** refers to the value of the objective function when making the optimal choice. The value function is always a function of the exogenous variables.

To find the comparative statics for the value function, in this case profit, we can use a result known as the **envelope theorem**. Informally, the envelope theorem says that when calculating how the value function is affected by a change in an exogenous variable, we only need to look at the direct effect of the exogenous variable, and can treat our choices as fixed. We can write this more formally for the case of a single variable (although note that it generalizes to multiple variables),

Theorem (Envelope). *Let $F(x; z)$ be an objective function with choice variable x and exogenous variable z. Let $V(z) = \max_x F(x; z)$ be the value function and $x^* = \text{argmax}_x F(x; , z)$ be the optimizing choice of x. Then,*

$$\frac{dV}{dz} = \frac{\partial F}{\partial z}(x^*, z)$$

Notice that in the last line, we are taking the partial derivative of F with respect to the argument x, and then evaluating that derivative at the point (x^*, z).

We will apply this to examine how profit changes with respect to price. Let π^* be the value function. First, we compute the partial derivative of profit π with respect to the price,

$$\frac{d\pi(L; p, w)}{dp} = \frac{d(pf(L) - wL)}{dp} = f(L)$$

The envelope theorem tells us that the comparative static on the value function is equal to the above evaluated at L^*,

$$\frac{d\pi^*}{dp} = f(L^*)$$

In other words, if prices increase by a small amount, the additional profit that we earn will be equal to how much we are producing.

We will now prove the envelope theorem in the single variable case.

Proof. First, observe that we can write $V(z) = F(x^*(z), z)$ because x^* is the maximizing

choice. We can use the multivariate chain rule to totally differentiate V with respect to z

$$\frac{dV}{dz} = \frac{\partial F}{\partial x}(x^*(z), z)\frac{dx^*}{dz} + \frac{\partial F}{\partial z}(x^*(z), z)$$

However, note that the first order condition for optimization is

$$\frac{\partial F}{\partial x}(x^*(z), z) = 0$$

So, plugging into above yields,

$$\frac{dV}{dz} = 0\frac{dx^*}{dz} + \frac{\partial F}{\partial z}(x^*(z), z) = \frac{\partial F}{\partial z}(x^*(z), z) = \frac{\partial F}{\partial z}(x^*(z), z)$$

$\square$

Essentially the envelope theorem says that because we are at an optimum, a small change in our choice variable does not change the value of the overall objective function because of the first order condition. In the context of profit, this tells us that the effect of a change in price is purely the increase in price from the units that we are already producing, $f(L^*)$. Note that this does *not* say that our choice of labor or that our production amount is not changing. Rather, it says that becuase we are at an optimum, the derivative of the profit with respect to L must be 0. This means that a small change in L will not affect the overall profit, so the only effect is the mechanical effect of increasing the price of the current units sold.

3.6 Policy applications

Now that we have a basic model of the firm setup, we can augment it slightly to add some common policy questions and see how they affect the firm's output. In particular, we can examine fixed costs and taxes on firms. For each addition to the basic model, we can ask whether it changes the firm's decisions or its profits.

Fixed costs

Suppose that firms have some fixed cost, b, that must be paid. This might be the cost of purchasing land, a factory, or filling out paperwork. First, we will assume that the firm must operate. In this case, the profit function is given by,

$$\pi = pf(L) - wL - b$$

Notice however, that this is a monotonic transformation of the profit function, which means that the optimal choice remains the same. We can verify this by taking first order

conditions,

$$\frac{\partial \pi}{\partial L} = pf'(L^*) - w = 0 \implies pf'(L^*) = w$$

This is the same as the first order conditions in the basic firm model, which tells us that the optimal choices are the same. In particular, this means that profits are simply b dollars lower than in the standard case.

However, so far we have assumed that the firm is forced to operate. In reality, if profits are negative, the firm would likely shut down. So, if the choice of L^* implicitly defined above yields a positive profit, then the firm will choose L^*. However, if it yields a negative profit, then the firm will shut down and hire 0 labor. So, the choice of labor is instead given by,

$$\begin{cases} L^* & \text{if } pf(L^*) - wL^* - b \geq 0 \\ 0 & \text{o/w} \end{cases}$$

Tax per unit sold

In both the United States and Europe, taxes are levied on many goods at the point of sale. In the United States, this takes the form of a sales tax, while the European Union uses a Value-Added Tax. Here, we model the effect of a tax on sales on firm production.

Suppose that the government levies a tax of τ on each unit of production sold. If the market price is still p, then this means that the firm really receives $p - \tau$ dollars per unit sold. Notice that we could also use this to model a per-unit subsidy, which would simply be the case where $\tau < 0$. We assume that the market price does not adjust to the imposition of this tax, and that $\tau < p$. This means that the firm's profit function is now,

$$\pi = (p - \tau)f(L) - wL$$

We can once again take first order conditions to obtain,

$$\frac{\partial \pi}{\partial L} = (p - \tau)f'(L^*) - w = 0 \implies (p - \tau)f'(L^*) = w$$

Notice that this is almost exactly the same as the original first order conditions, except that we've replaced p with $p - \tau$. This is because the firm is facing essentially the same problem, except instead of facing a price p for each good, they face $\tilde{p} = p - \tau$ as the price.

One question we might have is whether this leads to higher or lower labor consumption that in the normal case. If we had an explicit functional form, we could directly compare for the resulting values of L^*. However, this is more difficult without an explicit functional form. A nice trick though is take comparative statics. Notice that the original case with no tax is just the special case of $\tau = 0$. So, if we take comparative statics with respect to

τ, we could determine the direction of the change in labor usage. To do so, we implicitly differentiate the first order condition with respect to τ,

$$\frac{d}{d\tau}\left((p-\tau)f'(L^*)\right) = \frac{d}{d\tau}w$$

$$\implies (p-\tau)f''(L^*)\frac{dL^*}{d\tau} - f'(L^*) = 0$$

$$\implies \frac{dL^*}{d\tau} = \frac{f'(L^*)}{(p-\tau)f''(L^*)}$$

There are a few things to notice. The first is that since we assumed $\tau < p$, then $\frac{dL^*}{d\tau}$ is negative. That is, the higher the tax, the less labor the firm hires. Since the production function is monotonically increasing in L, then this also tells us that as taxes are higher, the firm produces less output.

The second thing to notice is that if we were to differentiate with respect to the market price, p, instead, we would get,

$$\frac{dL^*}{dp} = -\frac{f'(L^*)}{(p-\tau)f''(L^*)} = -\frac{dL^*}{d\tau}$$

That is, the effect of a change in the tax is exactly the negative of the effect of a change in price. This is because the firm only cares about the after-tax price, in which case a tax increase is essentially the same as a price increase.

Finally, a key point to notice is that if you were to compare $\frac{dL^*}{d\tau}$ here to the expression for $\frac{dL^*}{dp}$ in the model without taxes, you would get the same negative relationship except with p replaced with $p - \tau$. This is to emphasize the point that if you know how the firm would respond to p, you can essentially replace p with $p - \tau$ everywhere if you impose a tax.

Tax on labor

Rather than a tax on the end product, we can instead consider what would happen with a tax (or subsidy) on labor. This is modeled in essentially the same way as the tax on production, except now we impose a tax τ on each unit of labor, so that the total cost of one unit of labor is given by $w + \tau$. To avoid degenerate cases, we assume that $w + \tau > 0$ (this could be negative if we instead subsidized labor with $\tau < 0$). Then our profit function is given by

$$\pi = pf'(L^*) - (w+\tau)L$$

Much like the case of a tax on output however, all we have done is change the effective price of labor. This means that rather than re-calculating everything, we can just plug in $w + \tau$ in place of w everywhere in our original model. This immediately yields the

first order condition as,

$$pf'(L^*) = w + \tau$$

Furthermore, notice that τ is essentially the same as w from a mathematical perspective. That is, we could swap the symbols and the math would remain the smae. This means that the comparative statics are the same, so we can just use our original expression for $\frac{dL^*}{dw}$ (which does not actually depend on w),

$$\frac{dL^*}{d\tau} = \frac{1}{pf''(L^*)}$$

These calculations are useful shortcuts, and can be replicated by our normal approach.

Tax on profits

Finally, we will consider a proportional tax on profits. That is, we impose a tax, τ, on each dollar of profits earned. We assume that $\tau < 1$. This means that the firm's after tax profits are given by,

$$\pi = (1 - \tau)\left(pf(L) - wL\right)$$

However, notice that if $\tau < 1$, this is just a monotonic transformation of the original profit function, so the optimal choice of L remains the same. This tells us that *firm behavior is unchanged by a tax on profits*. However, you may have seen that economists are often opposed to high corporate income taxes. Why would this be the case if we can impose large taxes with no distortionary effects? This is because our model only has one period, and ignores important factors like investment in future productivity, and the risk of starting a firm in the first place.

Recap

We can now calculate how profit-maximizing firms will choose how many workers to hire, and we have also examined how the firm's labor demand, production, and profit will change in response to variations in price and wages. We have observed these changes not only with specific production functions, but also obtained more general results for how firms should respond to changes. These are already some very powerful economic modeling tools that can formalize existing economic intutions and suggest how we might capture more complex results. In the next chapter, we will develop the model slightly by considering firms who not only choose labor, but multiple inputs.

4 Firms with Multiple Inputs

In the previous chapter, we developed the basic concepts for understanding firms' behavior by setting up a model for firms with a single input (labor) and solving for comparative statics as wages and prices change. This chapter will expand the model to include multiple inputs, which will demonstrate more generally the tools used to study models with more than one choice variable.

4.1 Model Setup

As before, firms produce according to some production function f, and they seek to maximize their profits π. Now, we consider the case where the firm optimizes over two inputs, capital (K) and labor (L), so we can express our production function as $f(K, L)$. In addition to these two choice variables, we have three exogenous variables: price of the product (p), wage (w), and cost of renting capital (r). Making the same assumptions about a perfectly competitive market for labor and the firm, we can express our profit function as

$$\pi(K, L; p, r, w) = pf(K, L) - rK - wL.$$

The firm thus solves the optimization problem

$$\max_{K,L} \pi(K, L; p, r, w) = \max_{K,L} pf(K, L) - rK - wL.$$

4.2 Solving the Model

We want to solve the optimization problem

$$\max_{K,L} pf(K, L) - rK - wL$$

to find the optimal values of K^* and L^* for the choice variables. We follow the same steps as the single variable case.

First Order Conditions   Since we have two choice variables, we now have two first order conditions that must be simultaneously satisfied

$$\begin{cases} p\frac{\partial}{\partial K}f(K^*, L^*) - r = 0 \\ p\frac{\partial}{\partial L}f(K^*, L^*) - w = 0. \end{cases}$$

Rearranging gives us an implicit definition for $K^*(p, r, w)$ and $L^*(p, r, w)$:

$$\boxed{\begin{cases} p\frac{\partial}{\partial K}f(K^*, L^*) = r \\ p\frac{\partial}{\partial L}f(K^*, L^*) = w. \end{cases}}$$

This result carries the same intuition as the univariate case: firms will purchase an input (e.g. labor, capital) until the point where its marginal revenue product is equal to its marginal cost.

Second Order Conditions   We again check second order conditions to verify that our optimum is indeed a maximum. These conditions are trickier when we have multiple inputs: we need to make sure that our function π at the point (K^*, L^*) is not increasing in *any* direction, not just the two directions along K and L. Formally, the second order condition is satisfied in the multivariate case if and only if the Hessian matrix is negative definite (see Math Review). In the case of two variables, this condition is equivalent to checking that

$$\begin{cases} \frac{\partial^2 \pi}{\partial K^2} < 0 \\ \frac{\partial^2 \pi}{\partial L^2} < 0 \\ \frac{\partial^2 \pi}{\partial K^2}\frac{\partial^2 \pi}{\partial L^2} - \left(\frac{\partial^2 \pi}{\partial K \partial L}\right)^2 > 0. \end{cases}$$

If we were considering a general function π, then these conditions are automatically true if π is concave. We can check these more explicitly for our expression for profit. Like in the previous chapter, the first two conditions give

$$\begin{cases} \frac{\partial^2 \pi}{\partial K^2} = p\frac{\partial^2 f}{\partial K^2}(K^*, L^*) < 0 \\ \frac{\partial^2 \pi}{\partial L^2} = p\frac{\partial^2 f}{\partial L^2}(K^*, L^*) < 0, \end{cases}$$

which have the interpretation that the marginal revenue product of capital and labor are diminishing, as assumed. The third condition gives

$$p^2 \frac{\partial^2 f}{\partial K^2}\frac{\partial^2 f}{\partial L^2} > \left(p\frac{\partial f^2}{\partial K \partial L}\right)^2,$$

or equivalently,

$$\frac{\partial^2 f}{\partial K^2}\frac{\partial^2 f}{\partial L^2} > \left(\frac{\partial f^2}{\partial K \partial L}\right)^2.$$

Intuitively, an example where this condition might not hold would be if capital and labor were very strong complements. Then, even though the marginal benefit of capital decreases with more capital and the marginal benefit of labor decreases with more labor, the marginal benefit of capital increases with more labor, enough to an extent that there still exists a direction where the production function is upward sloping. However, if we assume that the production function is concave, then this second order condition is automatically met.

4.3 Comparative statics

Now that we have implicit definitions for $K^*(p, r, w)$ and $L^*(p, r, w)$, we can take the comparative statics with respect to the exogenous variables p, r, and w. Here, we will take the comparative statics with respect to w, which means we are interested in finding $\frac{\partial K^*}{\partial w}$ and $\frac{\partial L^*}{\partial w}$ and interpreting their signs. Switching notation for our differentiation, recall that we have the first order conditions

$$\begin{cases} pf_K\left(K^*, L^*\right) = r \\ pf_L\left(K^*, L^*\right) = w. \end{cases}$$

that implicitly define K^* and L^*. We can thus apply the Implicit Function Theorem and totally differentiate both of the above conditions, yielding

$$\begin{cases} \frac{d}{dw}\left[pf_K\left(K^*(p, r, w), L^*(p, r, w)\right)\right] = \frac{dr}{dw} \\ \frac{d}{dw}\left[pf_L\left(K^*(p, r, w), L^*(p, r, w)\right)\right] = \frac{dw}{dw}. \end{cases}$$

We can suppress the arguments to K^* and L^* (but do not forget that these are functions!) and simplify to get

$$\begin{cases} p\left(f_{KK}\frac{\partial K^*}{\partial w} + f_{KL}\frac{\partial L^*}{\partial w}\right) = 0 \\ p\left(f_{LK}\frac{\partial K^*}{\partial w} + f_{LL}\frac{\partial L^*}{\partial w}\right) = 1. \end{cases}$$

We have a system of linear equations and are interested in obtaining $\frac{\partial K^*}{\partial w}$ and $\frac{\partial L^*}{\partial w}$, so we can solve our system with Gaussian elimination, substitution, Cramer's rule, or any method you prefer. Using Gaussian elimination, we can rearrange to get

$$\begin{cases} f_{LL}p\left(f_{KK}\frac{\partial K^*}{\partial w} + f_{KL}\frac{\partial L^*}{\partial w}\right) - f_{KL}\left[p\left(f_{LK}\frac{\partial K^*}{\partial w} + f_{LL}\frac{\partial L^*}{\partial w}\right) - 1\right] = 0 \\ g_{KK}\left[p\left(f_{LK}\frac{\partial K^*}{\partial w} + f_{LL}\frac{\partial L^*}{\partial w}\right) - 1\right] - f_{LK}p\left(f_{KK}\frac{\partial K^*}{\partial w} + f_{KL}\frac{\partial L^*}{\partial w}\right) = 0. \end{cases}$$

Simplifying yields

$$\begin{cases} pf_{KK}f_{LL}\frac{\partial K^*}{\partial w} - pf_{KL}^2\frac{\partial K^*}{\partial w} + g_{KL} = 0 \\ pf_{KK}f_{LL}\frac{\partial L^*}{\partial w} - f_{KK} - pf_{KL}^2\frac{\partial L^*}{\partial w} = 0. \end{cases}$$

This allows us to solve for

$$\begin{cases} \frac{\partial K^*}{\partial w} = -\frac{1}{p} \frac{f_{KL}}{f_{KK} f_{LL} - f_{KL}^2} \\ \frac{\partial L^*}{\partial w} = \frac{1}{p} \frac{f_{KK}}{f_{KK} f_{LL} - f_{KL}^2}. \end{cases}$$

Remember that we are ultimately interested in the signs of these two terms. Notice that the $f_{KK} f_{LL} - f_{KL}^2$ term in each of the denominators must be positive, since this is exactly the third of our SOCs! We know that $f_{KK} < 0$ from our first SOC, so we know that

$$\frac{\partial L^*}{\partial w} < 0.$$

This has the unsurprising interpretation that as wages increases, the amount of labor hired decreases.

Notice that that the sign of $\frac{\partial K^*}{\partial w}$ depends on the sign of f_{KL}; the former is positive if and only if the latter is negative. Whether f_{KL} is positive or negative depends on the specific production function f. If $f_{KL} > 0$, we say that capital and labor are **complements**. That is, when labor increases, the marginal product of capital increases. Thus, when wages rise and the firm hires less labor, the marginal product of capital falls, so the optimal quantity of capital rented also falls. Intuitively, when inputs are complements, if we want less of one input, then we also want less of the other.

Alternatively, if $f_{KL} < 0$, we say that capital and labor are **substitutes**. That is, when labor increases, the marginal product of capital decreases. Then, when wages rise and the firm hires less labor, the marginal product of capital increases, so the optimal quantity of capital rented also increases. This effect explains why we call the inputs substitutes: when we want less of one input, we now want more of the other.

4.4 Returns to Scale

As we saw in the previous section, the nature of how firms hire labor and rent capital depends a lot on the structure of the actual production function $f(K, L)$. We often care about two key questions.

1. Are capital and labor substitutes or complements?

2. What are the returns to scale? That is, does the per-unit cost of production increase or decrease as production scales up?

We discussed the first question in the previous section; we now turn our attention to the second question. Given a production function $f(K, L)$, there are three cases:

- $f(K, L)$ has **constant returns to scale** if

$$f(\lambda K, \lambda L) = \lambda f(K, L)$$

for all $K, L, \lambda > 0$.

- $f(K, L)$ has **increasing returns to scale** if

$$f(\lambda K, \lambda L) > \lambda f(K, L)$$

for all $K, L, \lambda > 1$.

- $f(K, L)$ has **decreasing returns to scale** if

$$f(\lambda K, \lambda L) < \lambda f(K, L)$$

for all $K, L, \lambda > 1$.

Intuitively, the returns to scale tell us whether a big factory is more or less efficient than a small one. Doubling all of the inputs will always result in more production, but by how much? If doubling all of the inputs doubles output, then there are constant returns to scale. If doubling the inputs creates more than double the output, then there are increasing returns to scale. If doubling the inputs creates less than double the output, then there are decreasing returns to scale.

The returns to scale of a production function are fundamentally important to how the firm maximizes profits. If returns to scale are always increasing, then the firm will always want to produce an infinite quantity. If returns to scale are constant, then it will either (1) want to produce an infinite amount, (2) produce nothing, or (3) be indifferent across all output quantities.

4.5 Long Run vs. Short Run Comparative Statics

When we previously computed comparative statics with respect to wage, we assumed that the firm could adjust both capital and labor. However, we can also consider the case where the firm can only adjust labor in the short run, with capital fixed. For example, this might be relevant for a firm that would need a few years to build a new factory but can choose to higher more workers in the meantime to increase production.

Thus, our previous comparative static stays the same in the long run, when both capital and labor are adjustable, That is, we have

$$\left. \frac{\partial L^*}{\partial w} \right|_{LR} = \frac{1}{p} \frac{f_{KK}}{f_{KK} f_{LL} - f_{KL}^2} = \frac{1}{p \left(f_{LL} - f_{KL}^2 / f_{KK} \right)}$$

from before.[1]

---

[1]The $\frac{\partial L^*}{\partial w}|_{LR}$ and $\frac{\partial L^*}{\partial w}|_{SR}$ notation is a shorthand to represent the fact that the $L^*(p, r, w)$ functions are different between the two cases, since they come from solving the long run and short run optimization problems, respectively.

However, in the case where K is held at a constant $\bar{K}$ in the short term, then our optimization problem becomes

$$\max_{L} pf(\bar{K}, L) - r\bar{K} - wL.$$

Then our first order condition becomes

$$p\frac{\partial}{\partial L}f(\bar{K}, L^*) - w = 0,$$

which gives

$$pf_L(\bar{K}, L^*) = w.$$

We want the new comparative static with respect to w in the short term, so using the Implicit Function Theorem and differentiating both sides yields

$$\left.\frac{\partial L^*}{\partial w}\right|_{SR} = \frac{1}{pf_{LL}}.$$

Note that both the short run and long run response to a wage increase is negative, but the long run response is larger in magnitude. Intuitively, if capital and labor are substitutes, then in the long run, the increase in wages would increase capital, driving labor down even more than in the short run. If capital and labor are complements, then in the long run, the increase in wages would decrease capital, which would also result in even less labor hired. In either case, when capital is able to adjust to the wage hike, the marginal revenue product of labor goes down, so labor goes down even more than it does in the short term.

4.6 Profit function

Consider again the value function of the firm's optimization problem, which is the profit function $\pi(p, r, w)$. Remember that this function only depends on the exogenous variables p, r, and w, since we assume that K and L are already at their optimal values $K^*(p, r, w)$ and $L^*(p, r, w)$. If we wanted to write out the profit function fully, it would then look like

$$\pi(p, r, w) = pf(K^*(p, r, w), L^*(p, r, w)) - rK^*(p, r, w) - wL^*(p, r, w).$$

Effects of parameter changes

What happens to profits when price changes? Taking the partial derivative with respect to p gives

$$\frac{\partial \pi}{\partial p} = f(K^*, L^*),$$

which is a results known as **Hotelling's Lemma.** But wait! Where are all the $\frac{\partial K^*}{\partial p}$ and $\frac{\partial L^*}{\partial p}$ terms? Recall that by the Envelope Theorem, since we are differentiating the value function with respect to the parameters, we can treat the original choice variables as constants. Intuitively, this is because the choice variables are already optimized with respect to the parameters, which means that their derivatives must be 0.

We can similarly differentiate with respect to r and w to get that

$$\frac{\partial \pi}{\partial r} = -K^*$$

and

$$\frac{\partial \pi}{\partial w} = -L^*.$$

Intuitively, this means that when input costs suddenly change by a little, the immediate impact on production is small, so the primary effect on profit comes from the additional input cost that must be paid.

Homogeneity

We first introduce the concept of homogeneity.

Definition. A function $f(x_1, x_2, \ldots, x_m)$ is **homogeneous of degree** n if for all $\lambda > 0$, we have

$$f(\lambda x_1, \lambda x_2, \ldots, \lambda x_m) = \lambda^n f(x_1, x_2, \ldots, x_m).$$

Notice from our definition that homogeneity of degree 1 is equivalent to constant returns to scale. We also introduce an important theorem.

Theorem (Euler). *If $f(x_1, x_2, \ldots, x_m)$ is homogeneous of degree n, then*

$$x_1 f_{x_1} + x_2 f_{x_2} + \ldots + x_m f_{x_m} = n f(x_1, x_2, \ldots, x_m).$$

Notice that this theorem means that when we have constant returns to scale (i.e. f is homogeneous of degree 1), then the production is equal to the sum of all the input amounts multiplied by their marginal product.

We can now introduce an important result:

Proposition. *The profit function $\pi(p, r, w)$ is homogeneous of degree 1. That is, for all $\lambda > 0$,*

$$\pi(\lambda p, \lambda r, \lambda w) = \lambda \pi(p, r, w).$$

Intuitively, this just means that if we convert all of our price units for p, r, and w, nothing fundamental would change—our profits would simply scale to the same amount in the new units.

Proof. Observe

$$
\begin{aligned}
\pi(\lambda p, \lambda r, \lambda w) &= \max_{K,L}\{\lambda p f(K,L) - \lambda r K - \lambda w L\} \\
&= \max_{K,L}\{\lambda[p f(K,L) - r K - w L]\} \\
&= \lambda \max_{K,L}\{p f(K,L) - r K - w L\} \\
&= \lambda \pi(p, r, w).
\end{aligned}
$$

$\square$

Convexity

Recall the mathematically formal definition of a convex functions from 2.3, which is a function $f : \mathbb{R}^n \to \mathbb{R}$ such that for any $\lambda \in [0,1], \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$
f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \le \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}).
$$

We then have the following fact:

Proposition. *The profit function is convex. That is, for any sets of prices (p_0, r_0, w_0) and (p_1, r_1, w_1) and any $\lambda \in (0,1)$, let*

$$
(p_\lambda, r_\lambda, w_\lambda) = (\lambda p_0 + (1-\lambda)p_1, \lambda r_0 + (1-\lambda)r_1, \lambda w_0 + (1-\lambda)w_1)
$$

be their convex combination. Then

$$
\pi(p_\lambda, r_\lambda, w_\lambda) \le \lambda \pi(p_0, r_0, w_0) + (1-\lambda)\pi(p_1, r_1, w_1).
$$

Proof. Let $\mathbf{w} = (r, w)$ and let $\mathbf{Z} = (K, L)$. That is, we are pretending there is a single inpute $\mathbf{Z}$ with price $\mathbf{w}$. For any $(p_0, \mathbf{w}_0)$ and $(p_1, \mathbf{w}_1)$ and $\lambda \in (0,1)$, let

$$
(p_\lambda, \mathbf{w}_\lambda) = (\lambda p_0 + (1-\lambda)p_1, \lambda \mathbf{w}_0 + (1-\lambda)\mathbf{w}_1).
$$

We want to show

$$
\pi(p_\lambda, \mathbf{w}_\lambda) \le \lambda \pi(p_0, \mathbf{w}_0) + (1-\lambda)\pi(p_1, \mathbf{w}_1).
$$

We know that at $(p_\lambda, \mathbf{w}_\lambda)$, the profits are

$$
\pi(p_\lambda, \mathbf{w}_\lambda) = p_\lambda f(\mathbf{Z}^*(p_\lambda, \mathbf{w}_\lambda)) - \mathbf{w}_\lambda \cdot \mathbf{Z}^*(p_\lambda, \mathbf{w}_\lambda),
$$

where $\mathbf{Z}^*(p_\lambda, \mathbf{w}_\lambda)$ has been optimized by the firm at prices $(p_\lambda, \mathbf{w}_\lambda)$. If we tried to use the input quantity $\mathbf{Z}^*(p_\lambda, \mathbf{w}_\lambda)$ when prices were at $(p_0, \mathbf{w}_0)$ or $(p_1, \mathbf{w}_1)$, the resulting profits

would be weakly less than the profits $\pi(p_0, \mathbf{w}_0)$ and $\pi(p_1, \mathbf{w}_1)$ that have been optimized at those price points by definition. This gives us

$$\pi\left(p_0, \mathbf{w}_0\right) \geq p_0 f\left(\mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)\right) - \mathbf{w}_0 \cdot \mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)$$
$$\pi\left(p_1, \mathbf{w}_1\right) \geq p_1 f\left(\mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)\right) - \mathbf{w}_1 \cdot \mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right).$$

We can substitute these inequalities into our convex combination expression, yielding

$$
\begin{aligned}
\lambda \pi\left(p_0, \mathbf{w}_0\right) + (1 - \lambda)\pi\left(p_1, \mathbf{w}_1\right) &\geq \lambda\left[p_0 f\left(\mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)\right) - \mathbf{w}_0 \cdot \mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)\right] \\
&\quad + (1 - \lambda)\left[p_1 f\left(\mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)\right) - \mathbf{w}_1 \cdot \mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)\right] \\
&= \left[\lambda p_0 + (1 - \lambda)p_1\right] f\left(\mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)\right) \\
&\quad - \left[\lambda \mathbf{w}_0 + (1 - \lambda)\mathbf{w}_1\right] \cdot \mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right) \\
&= p_\lambda f\left(\mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right)\right) - \mathbf{w}_\lambda \cdot \mathbf{Z}^*\left(p_\lambda, \mathbf{w}_\lambda\right) \\
&= \pi\left(p_\lambda, \mathbf{w}_\lambda\right).
\end{aligned}
$$

$\square$

Notice that we did not even need for the profit function to be twice differentiable! The intuitive reasoning behind this proof is that average of the profits of the extreme prices would be the profits at the average prices if we chose the same inputs at the extreme prices as we did at the average price. However, we can always do weakly better at each of the extreme prices by reoptimizing our inputs compared to what we chose at the average, which means that the average profits of these extremes must be weakly greater.

How do we interpret this result? This means that given the options between (1) prices constantly fluctuating between (p_0, w_0, r_0) and (p_1, w_1, r_1) with some probability and (2) prices staying at their expected value always, the firm would prefer scenario (1). This is because under price fluctuations, the firm has more opportunities to reoptimize its input quantities compared to the case where prices are fixed.

4.7 Specific production functions

Now that we have the tools to study general production functions, we consider two common and useful examples.

Leontief (fixed-proportion) technology

Here, we have the functional form

$$f(K, L) = \min\left\{\frac{K}{a_K}, \frac{L}{a_L}\right\}.$$

The interpretation of this form is that the firm needs capital and labor in specific proportions; any additional increase in one without an increase in the other will yield no extra production. For example, if it takes exactly one machine to produce one unit of product, and there must be exactly one worker operating each machine, then we would have $a_K = a_L = 1$. Notice that the returns to scale are constant; the firm will produce at infinity if the price of the product is greater than the unit cost, it will produce at 0 if the price is less than the unit cost, and it will be indifferent if the price and the cost are equal.

Mathematically, a firm with this production function solves the problem

$$\max_{K,L} \left[p \left(\min \left\{ \frac{K}{a_K}, \frac{L}{a_L} \right\} \right) - rK - wL \right].$$

Observe that at the maximum, we must have

$$\frac{K^*}{a_K} = \frac{L^*}{a_L},$$

since increasing either capital or labor contributes nothing to production unless its counterpart is also increased proportionally. We can substitute this relationship into our maximization problem to get

$$\max_{K} \left[p\frac{K}{a_K} - rK - w\frac{a_L}{a_K}K \right].$$

Rearranging, we have

$$\max_{K} K \left(\frac{p}{a_K} - w\frac{a_L}{a_K} - r \right).$$

Notice that we do not even need first order conditions here! If $\frac{p}{a_K} - w\frac{a_L}{a_K} - r > 0$, then K^* and L^* go to infinity; otherwise, they stay at 0. This result is exactly what we concluded in the previous paragraph.

Cobb-Douglas technology

In this case, we have the production function

$$f(K, L) = AK^{\alpha}L^{\beta},$$

where $0 < \alpha, \beta < 1$ to ensure diminishing returns to capital and labor, and $A > 0$. Notice that the returns to scale depend on the value of $\alpha + \beta$. The firm has increasing returns to scale if this value is greater than 1, constant returns to scale if it is equal to 1, and decreasing returns to scale if it is less than 1. We can also see that

$$\frac{\partial f^2}{\partial K \partial L} = \alpha\beta AK^{\alpha-1}L^{\beta-1} > 0,$$

so capital and labor are complements in this model.

We could solve this model using the general solutions we developed in the previous sections, but it is an informative example to solve this model directly, since we will be revisiting this functional form several times in this class. We assume that $\alpha + \beta < 1$ so that we have decreasing returns to scale and thus an interior solution, and we assume that $A = 1$ for simplicity. We have the maximization problem

$$\max_{K,L} pAK^\alpha L^\beta - rK - wL.$$

Our first order conditions give

$$\begin{cases} \alpha p(K^*)^{\alpha-1}(L^*)^\beta - r = 0 \\ \beta p(K^*)^\alpha (L^*)^{\beta-1} - w = 0. \end{cases}$$

There are many ways to solve this system for K^* and L^*, but a useful and insightful trick is to rearrange the system to yield

$$\begin{cases} p(K^*)^\alpha (L^*)^\beta = \frac{rK^*}{\alpha} \\ p(K^*)^\alpha (L^*)^\beta = \frac{wL^*}{\beta}. \end{cases}$$

Notice that the term on the left in both equations is firm revenue, and each term on the right is the cost from either capital or labor, divided by the exponent of that term in the production function. Equating these terms gives

$$\frac{rK^*}{\alpha} = \frac{wL^*}{\beta},$$

which tells us that the costs spent on capital (rK^*) and labor (wL^*) are proportional to each other according to their corresponding exponents. We now solve for K^* and L^* by first rearranging our system into

$$\begin{cases} L^* = \frac{\beta r K^*}{\alpha w} \\ p(K^*)^\alpha \left(\frac{\beta r K^*}{\alpha w}\right)^\beta = \frac{rK^*}{\alpha}. \end{cases}$$

Expanding and rearranging the second term in this system yields the solution

$$\boxed{K^* = \left[p \left(\frac{\beta}{w}\right)^\beta \left(\frac{\alpha}{r}\right)^{1-\beta} \right]^{\frac{1}{1-\alpha-\beta}}.}$$

In order to find comparative statics for K^*, we would differentiate this expression with respect to p, w, and r. If we only care about the signs of these derivatives, a useful trick is to first take the log of this expression, which preserves sign because it is a monotonic

transformation. We leave these derivatives as an exercise to the reader, since it is possible to determine the signs of the derivatives by simply observing the expression at hand, yielding

$$\frac{\partial K^*}{\partial p} > 0, \frac{\partial K^*}{\partial r} < 0, \frac{\partial K^*}{\partial w} < 0.$$

The first two inequalities are expected: the first says that we rent more capital if our product can be sold more expensively, and the second says that we rent less capital if capital is more expensive. The third inequality tells us that we hire less capital when wages rise, so capital and labor are complements, as we noted before.

We can substitute this value of K^* into our earlier system to solve

$$L^* = \left[p \left(\frac{\alpha}{r} \right)^\alpha \left(\frac{\beta}{w} \right)^{1-\alpha} \right]^{\frac{1}{1-\alpha-\beta}},$$

which is analogous to our solution for K^*. We also have a similar set of comparative statics

$$\frac{\partial L^*}{\partial p} > 0, \frac{\partial L^*}{\partial r} < 0, \frac{\partial L^*}{\partial w} < 0$$

that also carry the analogous interpretation.

Finally, we can collect these expressions to solve for the profit function

$$
\begin{aligned}
\pi(p, r, w) &= p(K^*)^\alpha (L^*)^\beta - rK^* - wL^* \\
&= \frac{rK^*}{\alpha} - rK^* - \frac{\beta r K^*}{\alpha} \\
&= \frac{r(1 - \alpha - \beta)}{\alpha} K^* \\
&= \frac{r(1 - \alpha - \beta)}{\alpha} \left[p \left(\frac{\beta}{w} \right)^\beta \left(\frac{\alpha}{r} \right)^{1-\beta} \right]^{\frac{1}{1-\alpha-\beta}} \\
&= (1 - \alpha - \beta) \left[p \left(\frac{\beta}{w} \right)^\beta \left(\frac{\alpha}{r} \right)^\alpha \right]^{\frac{1}{1-\alpha-\beta}}.
\end{aligned}
$$

We can then take comparative statics to see that

$$\frac{\partial \pi^*}{\partial p} > 0, \frac{\partial \pi^*}{\partial r} < 0, \frac{\partial \pi^*}{\partial w} < 0.$$

These results should be intuitive: profits rise when price of the product rises, and profits fall when input costs rise.

Recap

5 Cost Minimization

So far, we have dealt with firms choosing the inputs that will maximize the profit that is earned. There have been no restrictions on how much of the good needs to be produced other than that some quantities will yield higher profits than others. However, often times firms cannot produce as much as they want, and must produce a certain quantity. For example, a farmer may sign a contract to produce $1,000$ bushels of wheat by the end of the year for some fixed price. In these cases, the firm is not solving an unconstrained maximization problem, but instead they face a constraint of producing a fixed amount of good. The way for the firm to maximize profits if they must produce a fixed quantity of product is to minimize the cost of producing that quantity, which is known as a **cost minimization** problem

In this chapter, we will go over how to perform cost minimization, as well as why cost minimization can be useful in solving general profit maximization problems.

5.1 Problem setup

To set up the cost minimization problem, we need to first establish our production function. For simplicity, we will assume that the firm has production function $f(K, L)$ where K is capital and L is labor. We assume that f is increasing and concave with respect to both K and L. That is,

$$f_K > 0$$
$$f_L > 0$$
$$\frac{\partial^2 f}{\partial K^2} < 0$$
$$\frac{\partial^2 f}{\partial L^2} < 0$$

The cost of labor is r, and the cost of labor is w, with both exogenous. We also have an exogenous quantity, Q, of goods that must be produced. The total cost of inputs is given by $wL + rK$. So we can write our minimization problem as,

$$\min_{K,L} rK + wL \text{ s.t. } f(K, L) = Q$$

This says that we are choose K and L to minimize $rK + wL$ subject to the constraint

that the amount we produce, $f(K, L)$, is equal to Q. To do so, we use constrained optimization. The Lagrangian is given by

$$\mathcal{L}(K, L, \lambda) = rK + wL - \lambda(f(K, L) - Q)$$

We can solve this via our standard constrained optimization methods.

First order conditions

We take the first order conditions on the Lagrangian, differentiating with respect to each variable, to obtain necessary conditions for a minimum,

$$\frac{\partial \mathcal{L}}{\partial K} = r - \lambda f_K(K, L) = 0$$
$$\frac{\partial \mathcal{L}}{\partial L} = w - \lambda f_L(K, L) = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = f(K, L) - Q = 0$$

Let L^*, K^*, and λ^* denote the values that satisfy the above conditions. Note that the third condition is simply the constraint, $f(K^*, L^*) = Q$. However, we can also rearrange and divde the first and second constraints to obtain,

$$\frac{r}{w} = \frac{f_K(K^*, L^*)}{f_L(K^*, L^*)} \implies \frac{f_K(K^*, L^*)}{r} = \frac{f_L(K^*, L^*)}{w} \tag{5.1}$$

f_K and f_L tell us how much additional good is produced per unit of capital and labor respective, while r and w tell us how much an additional unit of each costs. The above equality tells us that, at an optimum, the additional good produced per dollar spent must be equal for capital and for labor.

Now, consider $rK^* + wL^*$, where K^* and L^* are both functions of r, w, and Q. This tells us the total cost of producing Q units of good. We can then define

$$C(Q; r, w) = rK^* + wL*$$

This is known as the **cost function**, and it tells us the minimum cost to produce Q units of good. In the next section, we will prove some important properties of the cost function.

5.2 Cost function

Now that we have defined the cost function, we can examine some properties that it must exhibit. To do so, it will be useful to use the **constrained envelope theorem**.

Theorem (Constrained Envelope). *Let $F(x, y; z)$ be an objective function with choice variables x, y and exogenous variable z, and let $g(x, y; z) = c$ be the constraint. Denote the optimal choices of x and y by $x^*(z)$ and $y^*(z)$, respectively. Let $v(z) = F(x^*(z), y^*(z); z)$. Then,*

$$\frac{dv}{dz}(z) = \frac{\partial \mathcal{L}}{\partial z}(x^*, y^*, \lambda^*; z) = F_z(x^*, y^*; z) - \lambda^* g_z(x^*, y^*; z)$$

Where λ^* is the value of the Lagrange multiplier that satisfies the first order conditions. [1]

Proof. First, it will be useful to recall the first order conditions for the Lagrangian,

$$\frac{\partial \mathcal{L}}{\partial x} = 0 \implies F_x(x^*, y^*; z) = \lambda^* g_x(x^*, y^*; z)$$

$$\frac{\partial \mathcal{L}}{\partial y} = 0 \implies F_y(x^*, y^*; z) = \lambda^* g_y(x^*, y^*; z)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \implies g(x^*, y^*; z) = c$$

Next, the value function is given by

$$v(z) = F(x^*(z), y^*(z); z)$$

Totally differentiating v with respect to z yields,

$$\frac{dv}{dz}(z) = F_x(x^*, y^*; z)\frac{dx^*}{dz} + F_y(x^*, y^*; z)\frac{dy^*}{dz}(z) + F_z(x^*, y^*; z)$$

Now, notice that we can replace F_x and F_y using the first two equations in the FOC,

$$\frac{dv}{dz}(z) = \lambda^* g_x(x^*, y^*; z)\frac{dx^*}{dz} + \lambda^* g_y(x^*, y^*; z)\frac{dy^*}{dz} + F_z(x^*, y^*; z)$$

$$= \lambda^* \left(g_x(x^*, y^*; z)\frac{dx^*}{dz} + g_y(x^*, y^*; z)\frac{dy^*}{dz} \right) + F_z(x^*, y^*; z)$$

Now, we totally differentiate the third equation in the FOC with respect to z to obtain,

$$\frac{d}{dz} g(x^*(z), y^*(z); z) = \frac{d}{dz} c$$

$$\implies g_x(x^*, y^*; z)\frac{dx^*}{dz} + g_y(x^*, y^*; z)\frac{dy^*}{dz} + g_z(x^*, y^*; z) = 0$$

$$\implies g_x(x^*, y^*; z)\frac{dx^*}{dz} + g_y(x^*, y^*; z)\frac{dy^*}{dz} = -g_z(x^*, y^*; z)$$

---

[1]This formulation of the theorem is dependent on how you write the Lagrangian. We write the Lagrangian in this text as, $\mathcal{L}(x, y; z) = F(x, y; z) - \lambda(g(x, y; z) - c)$. However, it is also sometimes written as $\mathcal{L}(x, y; z) = F(x, y; z) + \lambda(g(x, y; z) - c)$ (with addition instead of subtraction). These are equivalent except for the fact that the sign of λ^* will flipped between them. So, for the latter formulation, we would have, $\frac{dv}{dz}(z) = F_z(x^*, y^*; z) - \lambda^* g_z(x^*, y^*; z)$

Plugging into the expression for $\frac{dv}{dz}$ yields,

$$\frac{dv}{dz}(z) = F_z(x^*, y^*; z) - \lambda^* g_z(x^*, y^*; z)$$

$\square$

With the constrained envelope theorem at hand, we can now examine some useful properties of the cost function.

Properties of the cost function

First notice that the cost function $C(Q; r, w)$ is a value function, so the equivalent of v in the statement of the constrained envelope theorem. The statement and the intuition of these properties will be the most important to remember, although the proofs may be helpful in better understanding methods of economic reasoning.

Shephard's Lemma $\frac{dC}{dr} = K^*(Q, r, w), \frac{dC}{dw} = L^*(Q, r, w)$. This is similar to Hotelling's Lemma, but tells us that as the price of an input increases, the cost increases by the amount that input is used.

Proof. Shephard's lemma is a straightforward application of the constrained envelope theorem,

$$\frac{dC}{dr} = \frac{d\left(rK^* + wL^*\right)}{dr} - \lambda^*\frac{\partial f}{\partial r} = K^*$$
$$\frac{dC}{dw} = \frac{d\left(rK^* + wL^*\right)}{dw} - \lambda^*\frac{\partial f}{\partial w} = L^*$$

Where $\frac{\partial f}{\partial r} = \frac{\partial f}{\partial w} = 0$ since the production function does not directly depend on r or w. $\square$

Homogeneous of degree 1 in input prices $C(Q; \alpha r, \alpha w) = \alpha C(Q; r, w)$ for $\alpha \geq 0$. The intuition is that we are merely changing the unit of currency with which we are calculating costs.

Proof. The first order conditions from 5.1 requires that,

$$\frac{\alpha r}{\alpha w} = \frac{r}{w} = \frac{f_K(K^*, L^*)}{f_L(K^*, L^*)}$$

The constraint does not depend on r or w, so since the first order conditions are the same, we must have the optimized quantities are the same,

$$K^*(Q, \alpha r, \alpha w) = K^*(Q, r, w), L^*(Q, \alpha r, \alpha w) = L^*(Q, r, w)$$

Plugging into the cost function yields,

$$
\begin{aligned}
C(Q; \alpha r, \alpha w) &= \alpha r K^*(Q, \alpha r, \alpha w) + \alpha w L^*(Q, \alpha r, \alpha w) \\
&= \alpha r K^*(Q, r, w) + \alpha w L^*(Q, r, w) \\
&= \alpha(r K^*(Q, r, w) + w L^*(Q, r, w)) \\
&= \alpha C(Q; r, w)
\end{aligned}
$$

$\square$

Concave in input prices Using the more mathematically formal definition of a concave function, this states that

$$
C(Q, \alpha r_1 + (1 - \alpha)r_2, \alpha w_1 + (1 - \alpha)w_2) \geq \alpha C(Q, r_1, w_1) + (1 - \alpha)C(Q, r_2, w_2)
$$

Where $\alpha \in [0, 1]$. This tells us that the cost of the average of two prices is greater than the average of the costs at each price individually. The logic here is the exact same as the logic for the convexity of the profit function. Because firms can reoptimize, they have lower costs at any two prices of inputs than if they had any weighted average of the two prices.

Proof. This is essentially equivalent to the profit function being convex in prices, and the proof is also basically the same. Let $\mathbf{w}_1 = (r_1, w_1)$ and $\mathbf{w}_2 = (r_2, w_2)$ be vectors of the input prices and let $\alpha \in [0, 1]$. Denote $\mathbf{w} = \alpha \mathbf{w}_1 + (1 - \alpha)\mathbf{w}_2$. Let $X^*(Q, r, w) = (K^*(Q, r, w), L^*(Q, r, w)$ be the vector of optimal choices. Notice that we can then write the cost function as a dot product, $C(Q, \mathbf{w}) = X^*(Q, \mathbf{w}) \cdot \mathbf{w}$. Then we have,

$$
\begin{aligned}
C(Q, \mathbf{w}) &= X^*(Q, \alpha \mathbf{w}_1 + (1 - \alpha)\mathbf{w}_2) \cdot (\alpha \mathbf{w}_1 + (1 - \alpha)\mathbf{w}_2) \\
&= X^*(Q, \mathbf{w}) \cdot \alpha \mathbf{w}_1 + X^*(Q, \mathbf{w}) \cdot (1 - \alpha)\mathbf{w}_2
\end{aligned}
$$

Now, note that $X^*(Q, \alpha \mathbf{w}_1) \cdot \alpha \mathbf{w}_1$ is the cost function when we have input prices $\alpha \mathbf{w}_1$ and must be, by definition of the cost function, the minimum possible amount we spend to produce Q at prices $\alpha \mathbf{w}_1$. This means that $X^*(Q, \mathbf{w})$ must not be the best choice of inputs at prices $\alpha \mathbf{w}_1$, so the cost must be higher. That is,

$$
X^*(Q, \mathbf{w}) \cdot \alpha \mathbf{w}_1 \geq X^*(Q, \alpha \mathbf{w}_1) \cdot \alpha \mathbf{w}_1 = C(Q, \alpha \mathbf{w}_1)
$$

The same must also hold for $(1 - \alpha)\mathbf{w}_2$. So,

$$
\begin{aligned}
C(Q, \mathbf{w}) &= X^*(Q, \mathbf{w}) \cdot \alpha \mathbf{w}_1 + X^*(Q, \mathbf{w}) \cdot (1 - \alpha)\mathbf{w}_2 \\
&\geq X^*(Q, \alpha \mathbf{w}_1) \cdot \alpha \mathbf{w}_1 + X^*(Q, (1 - \alpha)\mathbf{w}_2) \cdot (1 - \alpha)\mathbf{w}_2 \\
&= C(Q, \alpha \mathbf{w}_1) + C(Q, (1 - \alpha)\mathbf{w}_2) \\
&= \alpha C(Q, \mathbf{w}_1) + (1 - \alpha)C(Q, \mathbf{w}_2) \text{ because homogeneous degree 1}
\end{aligned}
$$

$\square$

Inputs decrease with price increase $\frac{dK^*}{dr} \leq 0, \frac{dL^*}{dw} \leq 0$. That is, as the price of an input increases, we must use weakly less of that input.

Proof. The easiest way to see this is using the fact that the cost function is concave. Notice that using the envelope theorem, we have that

$$
\frac{dC}{dr} = K^*
$$

Then, differentiating again, we get the second derivative as,

$$
\frac{d^2 C}{dr^2} = \frac{dK^*}{dr}
$$

Because C is concave with respect to r, we have that $\frac{d^2 C}{dr^2} = \frac{dK^*}{dr} < 0$. And the same logic applies for $\frac{dL^*}{dw}$. $\square$

Costs increasing in quantity $\frac{dC}{dQ} > 0$, and in particular, $\frac{dC}{dQ} = \lambda^*$ where λ^* is the value of the Lagrange multiplier that satisfies the first order conditions. This is also known as the **shadow cost** of the constraint, which tells us how much costs increase for a small increase in the constraint Q.

Proof. First, we show that $\frac{dC}{dQ} = \lambda^*$. This follows from the constrained envelope theorem,

$$
\begin{aligned}
\frac{dC}{dQ} &= \frac{\partial \mathcal{L}}{\partial Q}(K^*, L^*, \lambda^*) \\
&= \frac{\partial (rK + wL - \lambda(f(K, L) - Q))}{\partial Q}(K^*, L^*, \lambda^*) \\
&= \lambda^*
\end{aligned}
$$

To show that $\lambda^* > 0$, we can look at the first order conditions of the Lagrangian:

$$
r = \lambda^* f_K \implies \lambda^* = \frac{r}{f_K}
$$

By assumption, $r > 0$ and $f_K > 0$, so we know that $\lambda^* > 0$. $\square$

Costs convex in quantity $\frac{d^2C}{dQ^2} > 0$. That is, as the amount of goods we must produce increases, so does the marginal cost, assuming that the production function is concave. Intuitively, a concave production function means that as we need produce more, we need more of the inputs to produce each additional unit of the output. This is the same as the cost of each additional unit increasing.

As a warning, the proof for this statement is longer than some of the other proofs and a little more confusing, so it is not important that you fully understand it. It is far more important to understand the intuition behind why convex cost functions and concave production functions are really the same thing. However, the proof may be useful to better understanding this intuition is formalized and common approaches to proving statements in mathematical economics.

Proof. To prove this fact, we will again use the more mathematically formal definition of convexity. Let Q_1 and Q_2 be two quantities of the good, and let $0 \leq \alpha \leq 1$. To prove convexity, we must show that

$$C(\alpha Q_1 + (1 - \alpha)Q_2) \leq \alpha C(Q_1) + (1 - \alpha)C(Q_2)$$

We will use vector notation to make the proof fully general and save some space on notation. Let $\mathbf{w}$ be the vector of the prices of the inputs, and let $\mathbf{X}^*$ be the vector of inputs that achieves the minimum cost $C(\alpha Q_1 + (1 - \alpha)Q_2)$. Let $\mathbf{X}_1^*$ and $\mathbf{X}_2^*$ be the cost minimizing inputs to produce Q_1 and Q_2, respectively.

We can then write our cost function as,

$$C(\alpha Q_1 + (1 - \alpha)Q_2) = \mathbf{w} \cdot \mathbf{X}^*$$

First, suppose we knew that if we used $\alpha \mathbf{X}_1^* + (1 - \alpha)\mathbf{X}_2^*$ inputs, then we could produce at least $\alpha Q_1 + (1 - \alpha)Q_2$ output. This is not obvious, and we will show it soon. However if we did know this, then we know that using $\alpha \mathbf{X}_1^* + (1 - \alpha)\mathbf{X}_2^*$ must have a higher cost than using $\mathbf{X}^*$, because $\mathbf{X}^*$ is the set of inputs with minimum cost to produce at least $\alpha Q_1 + (1 - \alpha)Q_2$ of the good. Mathematically, this says,

$$\begin{aligned} C(\alpha Q_1 + (1 - \alpha)Q_2) &= \mathbf{w} \cdot \mathbf{X}^* \\ &\leq \mathbf{w} \cdot (\alpha \mathbf{X}_1^* + (1 - \alpha)\mathbf{X}_2^*) \\ &= \alpha(\mathbf{w} \cdot \mathbf{X}_1^*) + (1 - \alpha)(\mathbf{w} \cdot \mathbf{X}_2^*) \end{aligned}$$

However, notice that $\mathbf{w} \cdot \mathbf{X}_1^*$ is simply the minimum cost to produce Q_1, and same for $\mathbf{X}_2^*$ and Q_2. So, we can rewrite the inequality,

$$\begin{aligned} C(\alpha Q_1 + (1 - \alpha)Q_2) &\leq \alpha(\mathbf{w} \cdot \mathbf{X}_1^*) + (1 - \alpha)(\mathbf{w} \cdot \mathbf{X}_2^*) \\ &= \alpha C(Q_1) + (1 - \alpha)C(Q_2) \end{aligned}$$

Which is precisely the definition of a convex function. Now, we simply need to show that by using inputs $\alpha\mathbf{X}_1^* + (1-\alpha)\mathbf{X}_2^*$, we could produce at least $\alpha Q_1 + (1-\alpha)Q_2$ output. To do so, we simply use the fact that the profit function is concave (reverse the inequality in the convexity definition), to obtain,

$$f(\alpha\mathbf{X}_1^* + (1-\alpha)\mathbf{X}_2^*) \geq \alpha f(\mathbf{X}_1^*) + (1-\alpha)f(\mathbf{X}_2^*)$$

However, recall how we defined $\mathbf{X}_1^*$ and $\mathbf{X}_2^*$. They are the inputs that would produce Q_1 and Q_2, respectively. So, plugging in that fact to the above inequality yields

$$f(\alpha\mathbf{X}_1^* + (1-\alpha)\mathbf{X}_2^*) \geq \alpha f(\mathbf{X}_1^*) + (1-\alpha)f(\mathbf{X}_2^*)$$
$$= \alpha Q_1 + (1-\alpha)Q_2$$

This is exactly what we wanted to prove, so we are finished. $\square$

5.3 Duality of Profit Maximization

So far, we have dealt with the cost function in a circumstance where the firm may be required to produce some amount Q of the output for outside reasons. However, the cost function is also useful in a general profit maximization problem. In fact, the problem of cost minimization is the *exact same* as the problem of unconstrained profit maximization. This is known as **duality**, which is when a maximization problem can be converted into an equivalent minimization problem, and vice versa.

How are they dual problems?

To understand duality, let's first take a closer look at the problem of profit maximizatoin. Typically when we set up a profit maximization problem, we choose how much of each input to use, say how many workers to hire or how much capital to buy. However, through this choice of inputs, we are also implicitly choosing how much output to produce. This means that there is some profit maximizing quantity. Let's call this quantity Q^*.

Now the relationship to cost minimization becomes clearer. Suppose we knew that to maximize profits, we would have to produce Q^* of the good. Then the inputs that we choose must achieve the minimum cost to produce Q^*, which is $C(Q^*)$. Why is this the case? Well suppose we chose some other set of inputs to produce Q^*. Since we sell Q^* for the same price no matter what, we could strictly increase profits by switching to the cost minimizing set of inputs to produce Q^*.

In fact, the same logic tells us that no matter how much we produce, we would make the most profit by producing with the cost minimizing set of inputs. So, we could instead see the profit maximization problem not as choosing the inputs, but choosing the quantity. We can treat the cost function as a machine that essentially tells us how much it costs

to produce some quantity, and we would maximize with the cost minimization as given. That is, we could write the profit maximization problem as,

$$\max_Q pQ - C(Q)$$

While we will not provide a formal proof of the duality of the problems, the above intuition and reasoning should give you a good idea for why cost minimization and choosing the optimal quantity is the same as profit maximization by choosing precisely which inputs.

Why use cost minimization?

You may ask, if cost minimization and profit maximization solve the same problem, why do we need to cost minimize at all? After all, profit maximization seems simpler. Indeed, for most problems, it is a good bet that profit maximization will be easier to perform. After all, in profit maximization all you have is an unconstrained maximization problem rather than a constrained minimization problem followed by an unconstrained maximization problem.

However, there are still good reasons to care about the fact that these are dual problems. The first is that there are in fact some cases where cost minimization will be easier than profit maximization. This is because cost minimization only depends on the production function itself, not on how the good is sold. While in the perfectly competitive case that we have been dealing with so far this is not an issue, it can become more complicated once we reach a situation where the quantity produced also affects the price. In those cases, it may be easier to choose a quantity for a given cost, rather than having to choose the specific inputs when maximizing. In essence, cost minimization allows us to separate the problems of choosing inputs and choosing quantities.

The second reason that duality is useful is not necessarily as a problem solving mechanism in itself, but for choosing good models. For example, if you are writing a simple model of production, you may not care precisely how a good is produced, only that it has some associated cost. Many economic models will simply assume a cost function. The properties in the previous section tell us what a cost function would have to satisfy in order to represent a concave production function. We can therefore abstract away from the specific production process, and assume merely that each good produced has an associated cost defined by the cost function.

5.4 More complex constraints

So far, we have dealt with the case of a single equality constraint. However, we could have more complicated constraints, which we will discuss here.

Single inequality constraint

So far, we have said that we need to produce *exactly* Q quantity of a good. However, the problem might be reformulated as needing produce *at least* Q of a good, in which case our problem would be

$$\min_{K,L} rK + wL \text{ s.t. } f(K, L) \geq Q$$

However, this does not in fact change our optimization problem at all. In fact, this is the exact same optimization problem as

$$\min_{K,L} rK + wL \text{ s.t. } f(K, L) = Q$$

Why is this the case? Well suppose that if at the optimum, we produced not exactly Q, but some quantity $Q' > Q$. Could this ever be cost minimizing? The answer is no, because the cost function for producing exact quantities is an increasing function in Q. That is, we know that $C(Q') > C(Q)$. So, we could lower costs by producing slightly less than Q', but still at a level above Q. This means that whatever arrangement we have of producing Q' cannot be cost minimizing. Thus, when cost minimizing subject to an inequality constraint, the optimal choice must produce exactly the amount of quantity required.

Multiple equality constraints

We dealt with the case that we had to produce exactly Q of a single good. However, what if we had to produce multiple goods and produce an exact quantity of each? That is, suppose we instead had production functions $f_1, f_2, \ldots, f_n$ for each good, and had to produce $Q_1, Q_2, \ldots, Q_n$ of each good. An example where this may be the case is a firm needing to produce a fixed amount of good in each period, for n periods. Let $X_1, \ldots, X_k$ be the set of inputs, and let $w_1, \ldots, w_k$ be the price of each input.

To solve this problem, we once again use the Lagrangian, which in the multiple constraints case is given by

$$\mathcal{L} = \sum_{i=1}^{k} X_i w_i - \sum_{j=1}^{n} \lambda_j (f_j(\mathbf{X}) - Q_j)$$

All that we have done is add additional Lagrange multipliers, λ_j, and then you solve by taking first order conditions the same way you would normally, except for each λ_j.

Notice one concern with the multiple constraint case that is not present in the single constraint case: we are not guaranteed that a solution exists. That is, it is not clear that in general there will be a choice of inputs that produces *exactly* the required amount for each good. We may therefore require some restrictions, such as requiring that the inputs

used on good 1 do not affect the production of good 2, and vice versa.

For the most part in this class, we will only deal with single equality constraints, but it may be useful to know how to handle more equality constraints.

Multiple inequality constraints

A final complication is the case of multiple inequality constraints. This is considerably more difficult than either the multiple equality constraints or the single inequality constraint, because we may not satisfy the constraints with equality and we do not know how much extra we would produce. The way to solve a problem with multiple constraints is given by the **Karush-Kuhn-Tucker** (KKT) conditions, which are a set of necessary conditions for an optimum.

The KKT conditions are far beyond the scope of this course, and so we will not even describe them here. However, it is useful to know the problem that they solve so that you can know where to look in case you may find them helpful when developing your own model or when you read about the KKT conditions in an academic publication.

Recap

In this chapter, we detailed how to setup and solve a cost minimization problem to obtain the cost function $C(Q)$, for producing exactly Q output. We also described and proved some of the key properties that the cost function must satisfy. Finally, we showed how cost minimization is the dual problem of profit maximization. The techniques of constrained optimization that have been introduced in this section will be useful beyond cost minimization, and especially in the upcoming chapters on individual utility maximization. If you do not fully understand constrained optimization, it would be worth re-reading parts of this chapter, the Math Review , or to obtain additional practice from other resources.

Part III

Consumer Theory

6 Consumer Utility

So far, we have dealt primarily with firms and how they decide what to produce. Now, we address the other side of the market: consumers and how they decide what to buy and consume. However, we need to find a way to convert this to a maximization problem. With firms, we made the very reasonable assumption that they would try to maximize profits. However, with individuals it is less clear what they would be maximizing. In economics, we assume that individuals are maximizing a **utility function**, which, in a somewhat tautological definition, is simply whatever an individual maximizes when they are making choices.

6.1 What is utility?

You may have seen utility in previous economics courses described as a quantification of the "happiness" of individuals, and the utility function describes how many "utils" that an individual receives from consuming certain goods. This may be a useful way of thinking about the utility function and can add some valuable insights, but we want a more formal treatment of utility functions that does not rely on something as abstract and non-specific as representing "happiness." However, this leaves us with a series of problems. Can this happiness be measured and observed? Is it the same across people? Can different types of happiness be compared? In this section, we will explore the formal treatment of utility in economics that defines utility in a way that handles some of these issues and avoids others. It will not be important for you to understand every aspect in this approach, and we will avoid delving into the fully formal technicalities of utility, but it may be useful for you understand where utility comes from to know what you can and cannot do with utility functions.

Utility as preference relations

We start by approaching the problem of quantifying an individual's preferences by considering a simpler problem: determining whether an individual prefers one outcome over another. Suppose we have two outcomes, A and B, for a given individual. These could be any set of outcomes. A might represent receiving 4 apples and B might represent receiving 6 bananas, or A might represent going to Harvard while B is going to Yale.

We have a fairly reasonable to judge whether an individual prefers outcome A or outcome B, by observing which they choose when presented with a choice. This means

that we can denote a **preference relation** on outcomes, which expresses which outcome an individual prefers between two outcomes. The notation is as follows:

- $A \prec B$ means that the agent strictly prefers B to A. That is, given the choice between A and B, the agent would choose B.

- $A \succ B$ means that the agent strictly prefers A to B.

- $A \sim B$ means that the agent is indifferent between A and B.

- $A \precsim B$ means that the agent weakly prefers B to A. That is, either $A \prec B$ or $A \sim B$.

- $A \succsim B$ means that the agent weakly prefers A to B.

This allows us to rigorously define an individual's preferences by a preference relation on the possible outcomes. However, this alone leaves us with a bit too much room. In order to have useful preferences, we need to assume that agents have **rational preference**. However, what economists mean by rational preferences is not a normative description of rationality. An economist makes no judgement, for example, on whether prefering chocolate to vanilla ice cream is "rational." Instead, we define rationality by the following two axioms:

Completeness For any two outcomes A and B, exactly one of the following holds: $A \prec B$, $A \succ B$, or $A \sim B$. This axiom tells us two things. The first is that the agent always has some preference between any two outcomes, even if that preference is to be indifferent. The second is that an agent cannot simultaneously prefer A to B and prefer B to A.

Transitivity For any outcomes A, B, and C, $A \precsim B$ and $B \precsim C$ imply that $A \precsim C$. That is, if we prefer B to A and C to B, then we must prefer C to A as well.

Whether you think that these axioms are required to be considered rational is up to you, but for the purposes of microeconomics, we impose these requirements and assume our agents to have rational preferences.

From preference relations to utility functions

Now that we can define an individual's preferences, we can try to convert them into a utility function. A utility function is simply a way of expressing these preference relationships over outcomes by mapping each outcome to a real number, and outcomes that are more preferred have a higher value. Formally:

Definition (Utility function)**.** A function $u : X \to R$ is a utility function for a preference relation $\precsim$ if for $A, B \in X$, $A \precsim B \iff u(A) \leq u(B)$.

Notice that the utility function for a given set of preferences is not unique. To make the idea of a utilty function more concrete, let's consider a simple example with a finite set of outcomes.

Example. Let $X = \{A, B, C\}$ be the set of outcomes. Maybe A is getting an apple, B is getting a banana, and C is getting a coconut. Suppose we have a preference relation $\precsim$ where $A \precsim B \precsim C$. We want to construct a utility function u that expresses this preference relation. We might define u as follows:

$$u(A) = 1, u(B) = 2, u(C) = 3$$

Notice that because C is preferred to B, $u(C)$ is greater than $u(B)$, and the same is true for all pairs of preference relations. However, this is not the unique representation of the preference relations. Define $\tilde{u}$ as the same as u except with the output doubled:

$$\tilde{u}(A) = 2, \tilde{u}(B) = 4, \tilde{u}(C) = 6$$

Notice that this still represents the preference relation $\precsim$, but has different values than u does.

The above example illustrates an important point. Utility functions are ordinal, not cardinal. That is, the magnitude of the difference between $u(A)$ and $u(B)$ does not matter, but the sign does. We can state this more formally:

Proposition. *Let $u : X \to \mathbb{R}$ be a utility function representing a preference relation $\precsim$. Let $f : \mathbb{R} \to \mathbb{R}$ be a monotonically increasing function. Then $f \circ u : X \to \mathbb{R}$ is also a utility function representing $\precsim$.*

Proof. Let $A, B \in X$ where $A \precsim B$. Then $u(A) \leq u(B)$. By monotonicity of f, we also have that $f(u(A)) \leq f(u(B))$. Since A and B were arbitrary, this holds for all $A \precsim B$. So, $f \circ u$ is a utility function for $\precsim$. $\square$

This tells us that we can add, multiply, apply a positive exponent, take logarithms, or apply any monotonic function to a utility function and keep the same underlying preferences.

However, this also tells us that you **cannot compare utilities across individuals**. That is, we can not decide that one person is happier than another because they receive more utility, nor can we say that maximizing utility is in general a desirable goal. Those are cases of normative utility functions, but in our case we only deal with the formally defined utility function. Throughout this text and in the course, we may say that higher utility corresponds to an agent being "happier," but this is merely shorthand and to acheive intuition, and should not be interpreted as a claim on utility actually mapping to happiness.

There is one last wrinkle in our construction of the utility function. In the finite case, or even in the countably infinite case, the above rationality axioms are sufficient to construct a utility function from a preference relation. However, we might have cases where the set of outcomes is uncountably infinitely large. For example, if you have a utility function over how much money you receive, in which case the outcome space is all real numbers. The rationality axioms alone are insufficient to guarantee the existence of a well-defined utilty function for a preference relation over uncountably infinite axioms in this case. So, we need an additional axiom.

Contiunity of preferences For any sequence of outcome pairs, $\{(x^n, y^n)\}_{n=1}^{\infty}$ where $x^n \succsim y^n$ for all n, and $x = \lim_{n \to \infty} x^n, y = \lim_{n \to \infty} y^n$, then $x \succsim y$.

The above is a bit more mathematically formal than required in this course, and you do not need to know the continuity property. It basically says that our preference relations are preserved under limits. However, the key is that if $\precsim$ is a continuous preference relation, then we have a *continuous* utility function $u : X \to \mathbb{R}$ representing $\precsim$.

While this guarantees that there is a continuous utility function, it does not say anything about the differentiability or other properties of the utility function. However, now that we have established the formal mathematical foundations of utility, we can impose more structure to handle the consumer problem specifically. We will do so in the following section.

6.2 The consumer's problem

The consumer's problem is in some sense the foundation of all economics. It has to do with individuals trying to achieve the best outcome that they can. That is, they are maximizing utility. In this section, we describe the basic setup of the model, the assumptions in the model, and some basic properties from solving the model.

Model setup

We consider a set of n goods that a consumer can consume, and that the consumer chooses real quantities of each good. We denote the choice for amount of these goods $\mathbf{x} = (x_1, \ldots, x_n)$. This means that our space of "outcomes" is $X = \mathbb{R}^n$. We assume our agent has a continuously differentiable utility function $u : \mathbb{R}^n \to \mathbb{R}$. We make a few additional asssumptions on the utility function.

Increasing in goods We assume that u is increasing in each good. Mathematically, this is $\frac{\partial u}{\partial x_i} > 0$ for all i. A key assumption here is that the consumer wants each good, and that there are no "bads." There will be cases where this assumption no longer holds for a general utility maximization problem (pollution or garbage for example), but in this case we assume the agent can only be happier with there allocation. This also assumes non-satiation, so that agents always want more of the good.

Concavity We assume that $u(\mathbf{x})$ is concave in $\mathbf{x}$. Since u is differentiable, this tells us that $\frac{\partial^2 u}{\partial x_i^2} < 0$ for all x_i. The intuition here is that agents tend to have diminishing marginal returns. The 10th chocolate bar adds less additional happiness than the first chocolate bar does.

However, there is a slight problem here, which is that clearly the optimal action for an agent given these assumptions is just to consume an infinite amount of everything. In the real world this does not occur because we have a limited amount of money. So we assume that agents have an exogenous fixed income y that can be spent on purchasing goods. Each good i also has a positive price $p_i > 0$, which we assume the agent takes as exogenous, yielding the price vector $\mathbf{p}$. This implicitly assumes that the agent is a price-taker, so that the amount of good that the agent purchases has no effect on the price, which is the case if the agent is a relatively small spender in an economy with many other consumers and firms to buy from. This gives us the agent's **budget constraint**,

$$\mathbf{p} \cdot \mathbf{x} = \sum_{i=1}^{n} p_i x_i \leq y$$

So, the consumer's problem can be summarized as follows:

$$\max_{\mathbf{x} \in \mathbb{R}^n} u(\mathbf{x}) \text{ s.t. } \mathbf{p} \cdot \mathbf{x} \leq y$$

Notice that the budget constraint in this case is an inequality constraint. These are typically more difficult to deal with, but we can simplify the problem by replacing it with an equality constraint. To see why this is the case, consider what would happen if the budget constraint held with strict inequality so that $\mathbf{p} \cdot \mathbf{x} < y$. Since buying more of each good increases utility, we could buy some very small additional amount of the first good, say dx_1, which would increase utility, while still satisfying the budget constraint. This means that when the agent is optimizing, they must spend their entire budget. We can therefore rewrite the problem,

$$\max_{\mathbf{x} \in \mathbb{R}^n} u(\mathbf{x}) \text{ s.t. } \mathbf{p} \cdot \mathbf{x} = y$$

Solving the model

Now that we have set up the model, we will solve it to determine the optimal quantities of goods that consumers should consume. For the sake of simplicity and clarity, we will consider the case of only two goods, a and b, with prices p_a, p_b and utility function $u(a, b)$. The consumer's budget is still given as y. The problem the consumer solves is therefore

$$\max_{a,b} u(a, b) \text{ s.t. } p_a a + p_b b = y$$

Lagrangian method

Since this is a constrained maximization problem, we can solve it via the standard Lagrangian method. The Lagriangian is

$$\mathcal{L}(a, b, \lambda) = u(a, b) - \lambda(p_a a + p_b b - y)$$

The first order conditions of an optimum are therefore,

$$\frac{\partial u}{\partial a}(a^*, b^*) = \lambda^* p_a$$
$$\frac{\partial u}{\partial b}(a^*, b^*) = \lambda^* p_b$$
$$p_a a + p_b b = y$$

The third equation is simply the budget constraint. If we divide the first and second equations and rearrange slightly, we obtain,

$$\frac{\frac{\partial u}{\partial a}}{p_a} = \frac{\frac{\partial u}{\partial b}}{p_b}$$

This tells us that an optimum, the marginal utility per dollar from consuming a must be the same as the marginal utility per dollar from consuming b.

Since we have assumed that u is concave, these conditions define the optimal quantities $a^*(p_a, p_b, y)$ and $b^*(p_a, p_b, y)$. These are called the **Marshallian demand functions** for a and b, which tells us how much an optimizing consumer will buy of a good for given prices and income. Often when writing Marshallian demand, we omit the asterisk to have $a(p_a, p_b, y)$ and $b(p_a, p_b, y)$. Notice that each Marshallian demand is a function of both of the prices as well as the income.

Substituting the constraint

Before moving on, it is worth noting that the Lagrangian method is not the only way to solve a constrained optimization problem. An alternative and often easier method is known as "substituting in the constraint." This relies on the fact that we know that at an optimum, we must have $p_a a + p_b b = y$. One thing we can do then is simply solve for b in terms of the other variables:

$$b = \frac{y - p_a a}{p_b}$$

Since this is a constraint, it must always hold even after optimizing. So, we can plug in this constraint into the utility function to obtain $u\left(a, \frac{y - p_a a}{p_b}\right)$. We can then write our

new optimization problem as

$$\max_a u\left(a, \frac{y - p_a a}{p_b}\right)$$

This is exactly the same as our original problem, except instead of a constrained maximization problem with two variables, we have an unconstrained maximization problem of a single variable. This can dramatically simplify our computations. As always, we take first order conditions by totally differentiating with respect to a,

$$\frac{\partial u}{\partial a}\left(a^*, \frac{y - p_a a^*}{p_b}\right) - \frac{\partial u}{\partial b}\left(a^*, \frac{y - p_a a^*}{p_b}\right)\frac{p_a}{p_b} = 0$$

Rearranging slightly yields,

$$\frac{\frac{\partial u}{\partial a}\left(a^*, \frac{y - p_a a^*}{p_b}\right)}{p_a} = \frac{\frac{\partial u}{\partial b}\left(a^*, \frac{y - p_a a^*}{p_b}\right)}{p_b}$$

Notice that if we were to plug in the fact that $b^* = \frac{y - p_a a^*}{p_b}$, this is the exact same first order condition as we obtained through the Lagrangian method. When solving with an explicit functional form, plugging in the constraint can often make it significantly easier to find the optimal choices.

6.3 Value function

Now that we know how to find the demand functions of optimizing consumers, we can the **value function**, which is the value of the utility function under optimal consumption. Mathematically, the value function is defined as,

$$v(\mathbf{p}, y) = u(\mathbf{x}^*(\mathbf{p}, y))$$

Where $\mathbf{x}^*$ is the vector of optimal choice of goods. Notice that the value function is only a function of the exogenous variables, and essentially tells us what is the most utility a consumer can obtain given prices and income. The value function is also known as the **indirect utility function**. This is analogous to the indirect profit function in the case of firms. Note however that the value function, like the utility function, is ordinal rather than cardinal.

Properties of indirect utility

Marginal benefit of income The marginal benefit of income is given by the value of the Lagrange multiplier, λ^*. That is, $\frac{dv}{dy} = \lambda^*$. In particular, $\frac{dv}{dy} > 0$, which tells us that the more income an individual has, the greater their utility.

The intuition here should be clear. If you have more money than you did before, you can afford the old consumption bundle, and you will have money left over to consume more and make you better off.

Proof. The proof of this fact is analogous to that of the cost function increasing with respect to quantity in 5.2. We use the constrained envelope theorem (5.2).

$$\frac{dv}{dy} = \frac{\partial \mathcal{L}^*}{\partial y} = \frac{\partial}{\partial y}\left(u(\mathbf{x}^*) - \lambda^*(\mathbf{p} \cdot \mathbf{x}^* - y)\right) = \lambda^*$$

That this is positive follows from the first order conditions. Recall that at an optimum, we require

$$\frac{\partial u}{\partial x_i} = \lambda^* p_i$$

By assumption, we know that $\frac{\partial u}{\partial x_i}, p_i > 0$. So we must have that $\lambda^* = \frac{\partial u}{\partial x_i}/p_i > 0$. $\quad\square$

Decreasing in item prices As the price of a good i increases, the value function increases, $\frac{dv}{dp_i} > 0$. In particular, $\frac{dv}{dp_i} = -\lambda^* p_i$.

The intuition here is similar to the intuition about incomes. Suppose you had the same or higher utility with the optimal consumption bundle at new prices. Since prices were lower before, you could have purchased the same bundle at the old prices. But since this bundle is different from the old bundle, and the old bundle was optimizing, the new bundle must be worse.

Proof. This proof is also a straightforward application of the constrained envelope theorem (5.2),

$$\frac{dv}{dp_i} = \frac{\partial \mathcal{L}^*}{\partial p_i} = \frac{\partial}{\partial p_i}\left(u(\mathbf{x}^*) - \lambda^*(\mathbf{p} \cdot \mathbf{x}^* - y)\right) = -\lambda^* p_i$$

Since $\lambda^*, p_i > 0$, then $\frac{dv}{dp_i} < 0$. $\quad\square$

Continuous The value function v is continuous in prices and income. We will not provide a formal mathematical proof of this result. However, this follows from the fact that u is continuous, and the Marshallian demand functions $\mathbf{x}^*$ are also continuous, so $v = u(\mathbf{x}^*)$ must be continuous.

Quasi-Convexity At any convex combinations of prices and incomes, the utility obtain must be weakly less than the value function at at least one of the individual prices and incomes. Mathematically, let $\lambda \in [0, 1]$. Fix prices $\mathbf{p}_0, \mathbf{p}_1$ and incomes y_0, y_1. Define $\mathbf{p}_\lambda = \lambda p_0 + (1 - \lambda)p_1$ and $y_\lambda = \lambda y_0 + (1 - \lambda)y_1$. Then,

$$v(\mathbf{p}_\lambda, y_\lambda) \le \max\left\{v(\mathbf{p}_0, y_0), v(\mathbf{p}_1, y_1)\right\}$$

Intuitively, the reason this holds is the same as the reason convexity holds for the profit function. If prices are varied, you are able to optimize and achieve a better outcome than if you were at the average of the varied prices.

Proof. The simplest way to prove this is by contradiction. Assume that quasi-convexity does not hold. Then we must have that $v(\mathbf{p}_\lambda, y_\lambda) > v(\mathbf{p}_0, y_0)$ and $v(\mathbf{p}_\lambda, y_\lambda) > v(\mathbf{p}_1, y_1)$. What must be true for the budget constraints for this to be the case? Because $v(\mathbf{p}_0, y_0)$ and $v(\mathbf{p}_1, y_1)$ are optimizing at given prices and incomes, then the allocation chosen with prices $\mathbf{p}_\lambda$ and income y_λ, call it $\mathbf{x}_\lambda^*$, must not satisfy the budget constraints for case 0 or case 1. That is, we must have that $\mathbf{x}_\lambda^* \cdot \mathbf{p}_0 > y_0$ and $\mathbf{x}_\lambda^* \cdot \mathbf{p}_1 > y_1$. To see why this is the case, if it did satisfy one of the budget constraints, say $\mathbf{x}_\lambda^* \cdot \mathbf{p}_0 \leq y_0$, then $v(\mathbf{p}_0, y_0)$ is optimal, and hence would achieve a weakly greater utility than the utility achieved by $\mathbf{x}_\lambda^*$.

Next, we can consider how much we spend at prices $\mathbf{p}_\lambda$ to purchase the allocation, $\mathbf{x}_\lambda^*$,

$$
\begin{aligned}
\mathbf{p}_\lambda \cdot \mathbf{x}_\lambda^* &= (\lambda \mathbf{p}_0 + (1-\lambda)\mathbf{p}_1) \cdot \mathbf{x}_\lambda^* \\
&= \lambda(\mathbf{p}_0 \cdot \mathbf{x}_\lambda^*) + (1-\lambda)(\mathbf{p}_1 \cdot \mathbf{x}_\lambda^*) \\
&> \lambda y_0 + (1-\lambda)y_1
\end{aligned}
$$

Where the last line is using the fact that the allocation cannot satisfy either of the original budget constraints.

However, recall that $y_\lambda = \lambda y_0 + (1-\lambda)y_1$. But then plugging into the above inequality yields,

$$
\mathbf{p}_\lambda \cdot \mathbf{x}_\lambda^* > y_\lambda
$$

This violates the budget constraint, and so we know that $\mathbf{x}^*$ is not in fact a feasible allocation, a contradiction. $\qquad\square$

Indirect utility as utility　While so far we have treated utility functions as representing preferences over certain bundles of goods, recall that in general a utility function represents preferences over any set of outcomes. In this sense, the indirect utility function is itself a utility function, representing preferences over prices and incomes.

6.4 Functional forms and properties

So far we have tried to find methodologies and conditions that must hold for optimizing consumers with general utility functions. However, in this section we will discuss some common functional forms for utility functions, as well as some of their key properties.

For each of these, we will deal with the two variable case, but most can be extended to more choices.

Cobb-Douglas

The Cobb-Douglas utility function is a very useful utility function because it tells us the share of the budget that will be spent on each good.

Functional form $u(a, b) = a^\alpha b^\beta$ where $\alpha + \beta = 1$

Solving The easiest way to solve this is by first applying a log transformation, which is monotonic, to obtain $\log(u(a, b)) = \alpha \log(a) + \beta \log(b)$. Next, we set up the Lagrangian,

$$\mathcal{L} = \alpha \log(a) + \beta \log(b) - \lambda(p_a a + p_b b - y)$$

The first order conditions then yield,

$$\frac{\alpha}{a^*} = \lambda^* p_a$$

$$\frac{\beta}{b^*} = \lambda^* p_b$$

$$p_a a^* + p_b b^* = y$$

Rearranging the first two equations yields

$$a^* = \frac{\alpha}{p_a \lambda^*}, b^* = \frac{\beta}{p_b \lambda^*}$$

Plugging these into the third condition,

$$\frac{\alpha}{\lambda^*} + \frac{\beta}{\lambda^*} = \frac{\alpha + \beta}{\lambda^*}$$
$$= \frac{1}{\lambda^*}$$
$$= y$$

Rearranging yields $\lambda^* = \frac{1}{y}$. Plugging this fact into the first two conditions,

$$\frac{\alpha}{a^*} = \frac{p_a}{y} \implies a^* = \frac{\alpha y}{p_a}$$

$$\frac{\beta}{b^*} = \frac{p_b}{y} \implies b^* = \frac{\beta y}{p_b}$$

Constant budget shares One of the key properties of the Cobb-Douglas model is that the budget share of good a is α and the budget share of good b is β. This can be

seen by rearranging $a^* = \frac{\alpha y}{p_a}$ to obtain,

$$\frac{a^* p_a}{y} = \alpha$$

$a^* p_a$ is how much is spent on good a, and its ratio with respect to income is α. Similar analysis applies for good b. Notice this does not say that the quantities of the goods themselves will be a constant proportion, but only that amount spent on each good is a constant proportion of the income. This is a key shortcut that can help you with solving problems involving Cobb-Douglas utility much more quickly.

Quasi-linear utility

Quasi-linear utility is used when we want utility to be linear in one of our goods and concave in another. In our case, we will take a to be our linear good, and b to be our concave good. The intuition behind quasi-linear utility is that the concave good is an "essential" good where the marginal utility is very high when you have very little of it, but that once you have a sufficient amount you would rather buy the linear good.

Functional form $u(a, b) = a + \tilde{u}(b)$ where $\tilde{u}$ is a single variable concave function.

Solving As always, we set up the Lagrangian first,

$$\mathcal{L} = a + \tilde{u}(b) - \lambda(p_a a + p_b b - y)$$

The first order conditions are then given by,

$$1 = \lambda^* p_a$$
$$\frac{\partial \tilde{u}}{\partial b}(b^*) = \lambda^* p_b$$
$$p_a a^* + p_b b^* = y$$

Using the first equation, we obtain that $\lambda^* = \frac{1}{p_a}$. Plugging into the second equation yields,

$$\frac{\partial \tilde{u}}{\partial b}(b^*) = \frac{p_b}{p_a}$$

This equation implicitly defines the optimal value of b^* as a function of p_b and p_a. Notably however, it does *not* depend on the budget, y. Since we have implicitly defined b^*, we can define a^* in terms of b^*,

$$a^* = \frac{y - p_b b^*}{p_a}$$

That is, we take whatever money we have leftover from buying b^* and use it all to buy good a. Hence the linearity.

However, there is a slight problem here. Because b^* does not depend on y, we could in theory have a situation where a^* is negative. This is a problem because we are not allowing our consumer to sell a without owning it first. Technically then, we have two possible consumption bundles:

1. If the implicitly defined b^* is such that $p_b b^* \leq y$, then the above analysis works and we have $a^* = \frac{y - p_b b^*}{p_a}$.

2. If $p_b b^* > y$, then we cannot afford b^*. This means that we will put all of our budget into buying b, which means that we will buy $\frac{y}{p_b}$ of good b, and we will have $a^* = 0$.

Case 2 captures the case where we are very poor, and we would spend all of our money on essentials like food and water. However, once we have enough for survival, then we are in case 1 where would rather spend money on the linear goods, which are often luxuries.

Case 2 is also known as a corner solution, where the normal first order conditions are insufficient and we would spend all our budget on a particular good. These types of corners are possible whenever you have an optimal choice that does not depend on income, because it is then possible to overspend your income.

Assumption on budget   Most of the typical analysis we have done depends on having a sufficient budget such that $y > p_b b^*$. If this is the case, then our marginal utility of income is given by $\lambda^* = \frac{1}{p_a}$. That is, with an additional dollar of income, we buy $\frac{1}{p_a}$ of good a, which has linear utility.

7 Expenditure Minimization

Recall that in our discussion of firm theory, we noted that the firm's problem of profit maximization is the dual problem of cost minimization. That is, if we were to determine the set of inputs that maximized profits, we would get the same result if we knew this maximum profit level and then determined the set of inputs that minimized the costs in order to achieve that level of profit.

Analogously, in the case of consumer theory, the dual problem of utility maximization is known as **expenditure minimization.** While utility maximization asks for the set of goods needed to maximize utility under some constraint (e.g. what is the most utility I can achieve with \$20?), expenditure minimization asks for the set of goods that most cheaply attains a given level of utility (e.g. what is the cheapest way to achieve 100 utils of utility?). This chapter will study the expenditure minimization problem and explore how examining the utility maximization and expenditure minimization problems together gives us a fuller picture of how consumer behavior reacts to prices.

7.1 Problem Setup

Formally, suppose we have goods $1, \ldots, n$ with prices $p_1, \ldots, p_n$, respectively, and our choice variables $x_1, \ldots, x_n$ correspond to how much we purchase of each good, respectively. Denote our utility function from these goods as $u(x_1, \ldots, x_n)$. Suppose we must attain a utility level of at least $\bar{u}$, where $\bar{u}$ is some exogenous variable that is given to us. What is the cheapest way to achieve $\bar{u}$? We can write down our expenditure minimization problem as

$$\min_{x_1, x_2, \ldots, x_n} \sum_{i=1}^{n} p_i x_i \text{ s.t. } u\left(x_1, x_2, \ldots, x_n\right) \geq \bar{u}.$$

If we use vector notation and denote $\mathbf{p}$ as our prices and $\mathbf{x}$ as our chosen quantities, we can write this more succinctly as

$$\min_{\mathbf{x} \geq 0} \mathbf{p} \cdot \mathbf{x} \text{ s.t. } u(\mathbf{x}) \geq \bar{u}$$

We can solve this problem similar to how we solved the cost minimization problem from our study of firm theory.

First Order Conditions

We can write our Lagrangian in vector form as

$$\mathcal{L}(\mathbf{x}, \lambda) = \mathbf{p} \cdot \mathbf{x} - \lambda[u(\mathbf{x}) - \bar{u}].$$

Our first order conditions can be written as

$$\begin{cases} \mathcal{L}_{x_i} = p_i - \lambda^* u_{x_i}(\mathbf{x}^*) = 0 \text{ for all } i \\ \mathcal{L}_\lambda = u(\mathbf{x}^*) = 0 - \bar{u}. \end{cases}$$

Notice how this looks similar to our conditions from utility maximization. The last condition is simply the requirement that we attain $\bar{u}$. Our first set of conditions regarding each $\mathcal{L}_{x_i}$ term is the same as what we had from utility maximization, except since our constraint is different, the values of λ^* and $\mathbf{x}^*$ are different that what they were in the utility maximization case. Note that we could also have written this first set of conditions in vector form as

$$\mathbf{p} = \lambda^* \nabla u(\mathbf{x}^*)$$

Recognize that our first order conditions give us $n+1$ linear equations with $n+1$ unknowns. We thus have a solution for λ^* as well as solutions for

$$x_i^*(\mathbf{p}, \bar{u})$$

for $i = 1, 2, \ldots, n$. Notice that in the case of utility maximization, our solutions were of the form $x_i^*(\mathbf{p}, y)$. The important difference now is that our solution to the expenditure minimization problem is in terms of the required utility level rather than budget. Whereas we commonly denote $x_i^*(\mathbf{p}, y)$ as Marshallian demand, the solution to the utility maximization problem, we will denote the solution we just derived for the expenditure minimization problem as

$$h_i(\mathbf{p}, \bar{u})$$

for $i = 1, 2, \ldots, n$. This expression is known as the **Hicksian demand**, also referred to as the **compensated demand**. Moving forward, we will often drop arguments and simply write x_i^* and h_i^* for Marshallian and Hicksian demand, respectively, but it is important to remember how the arguments of these functions differ.

7.2 Expenditure Function

We now turn to the value function for the expenditure minimization problem, which is called the **expenditure function**:

$$e(\mathbf{p}, \bar{u}) = \mathbf{p} \cdot \mathbf{h}(\mathbf{p}, \bar{u}) = \sum_{i=1}^{n} p_i h_i(\mathbf{p}, \bar{u}).$$

In the last chapter, we noted that there are some similarities between a consumer's indirect utility function from utility maximization and a firm's profit function from profit maximization. However, the analogy was not exact. First, while the profit function is convex in prices, the indirect utility function is quasi-convex in prices. Second, while the profit function is homogeneous of degree 1, the indirect utility function is not.

With the expenditure function, we will notice that it is an exact analogy to a firm's cost function from cost minimization. We encourage the reader to refer back to Section 5.2 to notice the parallels between properties of the cost function and the expenditure function.

Properties of the Expenditure Function

Shephard's Lemma $\frac{\partial e}{\partial p_i} = h_i(\mathbf{p}, \bar{u})$. In the case of cost minimization, we saw that the derivative of cost with respect to input prices was equal to the input demand. Analogously, in the case of expenditure minimization, the derivative of expenditure with respect to prices of goods is equal to the Hicksian demand of those goods. The lemma and the proof are the exact same as what we saw before.

Proof. Applying the constrained envelope theorem, we have

$$\frac{\partial e}{\partial p_i} = \frac{\partial}{\partial p_i} \left\{ \mathbf{p} \cdot \mathbf{h} + \lambda^*[\bar{u} - u(\mathbf{h})] \right\} = h_i(\mathbf{p}, \bar{u}).$$

□

Homogeneous of degree 1 in prices $e(\alpha \mathbf{p}, \bar{u}) = \alpha e(\mathbf{p}, \bar{u})$. The proof is the same as in Section 5.2 and is encouraged as an exercise for the reader. The intuition is also the same: if we changed the units of the prices of all goods, then the units of our expenditure function would scale accordingly.

Concave in prices For all $\alpha \in [0, 1]$ and any price vectors $\mathbf{p}_1$ and $\mathbf{p}_2$,

$$e(\alpha \mathbf{p}_1 + (1-\alpha)\mathbf{p}_2, \bar{u}) \geq \alpha e(\mathbf{p}_1, \bar{u}) + (1-\alpha)e(\mathbf{p}_2, \bar{u}).$$

Again, the proof is the same as in Section 5.2 and is encouraged as an exercise for the reader. The basic idea of the proof is that at both $(\mathbf{p}_1, \bar{u})$ and $(\mathbf{p}_2, \bar{u})$, purchasing $\mathbf{h}(\alpha \mathbf{p}_1 + (1-\alpha)\mathbf{p}_2, \bar{u})$ will achieve the utility constraint, but you can spend weakly less by purchasing $\mathbf{h}(\mathbf{p}_1, \bar{u})$ or $\mathbf{h}(\mathbf{p}_2, \bar{u})$ at these points, respectively.

A similar interpretation follows as well: as a consumer, you would rather have prices fluctuate than have them stay constant at their average value, since you can reoptimize your consumption bundle at each of the fluctuations to achieve a lower expenditure.

Non-decreasing in prices and utility If $\mathbf{p}$ of $\bar{u}$ increase, then $e(\mathbf{p}, \bar{u})$ weakly increases. The proof is straightforward: if prices or required utility fell from a previous value, then the previous bundle would still satisfy the constraint, so minimum expenditure cannot rise.

Continuous We will not prove this property of the expenditure function, but it is good to remember.

7.3 Hicksian and Marshallian Duality

The Hicksian and Marshallian have offered us two different demand functions. However, the Hicksiand and the Marshallian are related in two important ways.

Given prices, the indirect utility function and the expenditure function are inverses.
We observe that

$$v(\mathbf{p}, e(\mathbf{p}, \bar{u})) = \bar{u}.$$

Intuitively, we defined $e(\mathbf{p}, \bar{u})$ as the minimum cost to achieve $\bar{u}$, so if we have a starting budget of $e(\mathbf{p}, \bar{u})$, then the maximum utility we can achieve with this budget is $\bar{u}$. Similarly, we observe that

$$e(\mathbf{p}, v(\mathbf{p}, y)) = y.$$

This time, we defined $v(\mathbf{p}, y)$ to be the maximum utility we can achieve with budget y, so the minimum budget needed achieve utility level $v(\mathbf{p}, y)$ must be y. We conclude that if prices $\mathbf{p}$ are given, then the value functions e and v are inverses of each other.

At optimality, Marshallian and Hicksian demand coincide. We observe that

$$\mathbf{x}(\mathbf{p}, e(\mathbf{p}, \bar{u})) = \mathbf{h}(\mathbf{p}, \bar{u}).$$

Intuitively, $\mathbf{h}(\mathbf{p}, \bar{u})$ is the cheapest bundle to attain utility $\bar{u}$, and it costs $e(\mathbf{p}, \bar{u})$. Thus, if we had a budget of $e(\mathbf{p}, \bar{u})$, then most satisfying bundle to buy should also be $\mathbf{h}(\mathbf{p}, \bar{u})$. Similarly, we observe that

$$\mathbf{h}(\mathbf{p}, v(\mathbf{p}, y)) = \mathbf{x}(\mathbf{p}, y).$$

Now, $\mathbf{x}(\mathbf{p}, y)$ is the most satisfying bundle to buy with a budget y, and it gives us utility $v(\mathbf{p}, y)$. Thus, if we must achieve a utility of $v(\mathbf{p}, y)$, then the cheapest bundle that attains this utility should also be $\mathbf{x}(\mathbf{p}, y)$.

To summarize, since utility-maximization and expenditure-minimization both yield optimal consumption bundles, these bundles coincide when our budget from the utility-maximization problem is just enough to attain the required utility from the expenditure-minimization problem.

7.4 Slutsky Equation

One of the questions are likely interested in now that we know how people demand goods is figuring out how demand changes for a given change in price. However, the answer is not always straightforward. Intuitively, if the price of a single good increases, we can separate the effect on consumption into two channels. First, that good becomes relatively more expensive and other goods become relatively cheaper; this effect is called the **substitution effect**. Second, the consumer has effectively become poorer in terms of the options she can afford; this effect is called the **income effect**. This section will study the **Slutsky equation**, which uses our ideas of Marshallian and Hicksian demand to derive an equation that contains both of these pieces together.

Derivation

Recall from duality that at an optimum, the Marshallian and the Hicksian must be equal. Focusing on a single good i, this implies that,

$$x_i(\mathbf{p}, e(\mathbf{p}, \bar{u})) = h_i(\mathbf{p}, \bar{u}),$$

By the Implicit Function Theorem, we can differentiate both sides of this equation with respect to p_i, yielding

$$\frac{d}{dp_i} x_i(\mathbf{p}, e(\mathbf{p}, \bar{u})) = \frac{d}{dp_i} h_i(\mathbf{p}, \bar{u}).$$

Using the multivariate chain rule, this becomes

$$\frac{\partial x_i}{\partial p_i}(\mathbf{p}, e(\mathbf{p}, \bar{u})) + \frac{\partial x_i}{\partial y}(\mathbf{p}, e(\mathbf{p}, \bar{u})) \frac{\partial e}{\partial p_i}(\mathbf{p}, \bar{u}) = \frac{\partial h_i}{\partial p_i}(\mathbf{p}, \bar{u}).$$

By Shephard's Lemma, we know $\frac{\partial e}{\partial p_i} = h_i(\mathbf{p}, \bar{u})$, so we can simplify our earlier equation to

$$\frac{\partial x_i}{\partial p_i}(\mathbf{p}, e(\mathbf{p}, \bar{u})) + h_i(\mathbf{p}, \bar{u}) \frac{\partial x_i}{\partial y}(\mathbf{p}, e(\mathbf{p}, \bar{u})) = \frac{\partial h_i}{\partial p_i}(\mathbf{p}, \bar{u}).$$

We now use the first fact that the expenditure function and the indirect utility function

are inverses to rewrite this equation in terms of some budget y instead of $\bar{u}$. Define $y = e(\mathbf{p}, \bar{u})$. Then $\bar{u} = v(p, y)$ by the first fact, and $h_i(\mathbf{p}, \bar{u}) = x_i(\mathbf{p}, y)$ by the second fact. We can substitute these terms into our equation to get

$$\frac{\partial x_i}{\partial p_i}(\mathbf{p}, y) + x_i(\mathbf{p}, y)\frac{\partial x_i}{\partial y}(\mathbf{p}, y) = \frac{\partial h_i}{\partial p_i}(\mathbf{p}, v(\mathbf{p}, y)).$$

Rearranging these terms gives us the famous **Slutsky equation**

$$\boxed{\frac{\partial x_i}{\partial p_i}(\mathbf{p}, y) = \frac{\partial h_i}{\partial p_i}(\mathbf{p}, v(\mathbf{p}, y)) - x_i(\mathbf{p}, y)\frac{\partial x_i}{\partial y}(\mathbf{p}, y).}$$

Suppressing the arguments, we can write this more succinctly as

$$\frac{\partial x_i}{\partial p_i} = \frac{\partial h_i}{\partial p_i} - x_i\frac{\partial x_i}{\partial y}.$$

Interpretation

The effect of price changes on consumption is exactly the $\frac{\partial x_i}{\partial p_i}$ term on the left, while the right hand side of the equation separates the substitution and income effects we were interested in.

Substitution effect. This component is given by the first term

$$\frac{\partial h_i}{\partial p_i}(\mathbf{p}, v(\mathbf{p}, y)).$$

Intuitively, this term represents the first-order change solely from good i becoming more expensive relative to other goods, assuming that we could still afford the original consumption bundle. Another explanation is that this term represents the effect of a price change if real income is kept constant, where real income is defined in terms of the ideal price index discussion from the previous section.

We claim that the substitution effect is always negative; intuitively, if an item becomes relatively more expensive, the substitution effect would cause a shift in spending to the now relatively cheaper items. This fact is called the **Law of Compensated Demand.**

Proof. Recall that the expenditure function is concave in $\mathbf{p}$, so

$$\frac{\partial^2 e}{\partial p_i^2}(\mathbf{p}, \bar{u}) < 0.$$

By Shephard's Lemma, this means

$$\frac{\partial}{\partial p_i}\frac{\partial e}{\partial p_i}(\mathbf{p}, \bar{u}) = \frac{\partial h_i}{\partial p_i}(\mathbf{p}, \bar{u}) < 0.$$

$\square$

Income effect. This component is given by the second term

$$-x_i \frac{\partial x_i}{\partial y}.$$

Intuitively, if the consumer becomes Δy poorer, the consumption of good i decreases by

$$\Delta x_i \approx \frac{\partial x_i}{\partial y} \Delta y.$$

If the consumer was originally buying x_i of good i, then a price increase of Δp_i effectively makes her poorer by

$$\Delta y \approx -x_i \Delta p_i.$$

Putting these pieces together, the income effect of a small price change Δp_i is

$$\Delta x_i \simeq -x_i \frac{\partial x_i}{\partial y} \Delta p_i,$$

which matches our result from Slutsky's equation.

What is the sign of the income effect? It depends! If the income effect is negative (i.e. if $\frac{\partial x_i}{\partial y} > 0$) then $\frac{\partial x_i}{\partial p_i} < 0$, since the substitution effect is always negative. This means that the demand function is downward sloping, as we might imagine intuitively. However, this might not always be the case; we might have a positive income effect when $\frac{\partial x_i}{\partial y} < 0$ for some items (e.g. cheap dorm food) that we consume less when we become richer.

General Slutsky Equation

In the original Slutsky equation, we were interested in the effect of a price change in good i on the consumption of good i, so we differentiated with respect to p_i. If we were instead interested in the effect of a price change in some other good j on the consumption of good i, we would differentiate with respect to p_j. The resulting derivation would look identical, and our final result would be the **general Slutsky equation**

$$\boxed{\frac{\partial x_i}{\partial p_j} = \frac{\partial h_i}{\partial p_j} - x_j \frac{\partial x_i}{\partial y}.}$$

The $-x_j \frac{\partial x_i}{\partial y}$ term represents the income effect, where the price of good j makes the consumer poorer in real terms, which has some effect on the consumption of good i depending on whether good i is an inferior or normal good. However, whereas the substitution effect used to be unambiguously negative, the new substitution effect $\frac{\partial h_i}{\partial p_j}$ in

the general equation can be positive or negative, depending on whether goods i and j are substitutes or complements, respectively.

Net and Gross Substitutes

Using this new substitution effect from the general Slutsky equation, we introduce two different ideas for how goods can be substitutes.

Definition (Net substitutes)**.** We say that goods i and j are **net substitutes** if

$$\frac{\partial h_i}{\partial p_j} = \frac{\partial h_j}{\partial p_i} > 0.$$

Intuitively, two goods are net substitutes if an increase in the price of one good causes the consumption of the other good to rise, assuming that utility must be kept constant. We note that this definition is symmetric, which we prove in the following proposition.

Proposition. *Net substitutability is a symmetric condition. That is,*

$$\frac{\partial h_i}{\partial p_j} = \frac{\partial h_j}{\partial p_i}.$$

Proof. By Shephard's Lemma,

$$h_i(\mathbf{p}, \bar{u}) = \frac{\partial e(\mathbf{p}, \bar{u})}{\partial p_i}.$$

Differentiating with respect to p_j gives

$$\frac{\partial h_i(\mathbf{p}, \bar{u})}{\partial p_j} = \frac{\partial^2 e(\mathbf{p}, \bar{u})}{\partial p_i \partial p_j}.$$

By Young's Theorem, we can swap the order of the partial derivatives, so

$$\frac{\partial^2 e(\mathbf{p}, \bar{u})}{\partial p_i \partial p_j} = \frac{\partial^2 e(\mathbf{p}, \bar{u})}{\partial p_j \partial p_i}.$$

Thus, we conclude

$$\frac{\partial h_i(\mathbf{p}, \bar{u})}{\partial p_j} = \frac{\partial h_j(\mathbf{p}, \bar{u})}{\partial p_i}$$

$\square$

We also claim that every good must have some net substitutes.

Proposition. *For all goods j, there exists good i such that $\frac{\partial h_i}{\partial p_j} > 0$.*

Proof. We know by the definition of the Hicksian that

$$u(\mathbf{h}(\mathbf{p}, \bar{u})) = \bar{u}.$$

Differentiating both sides by p_j and using the multivariate chain rule yields

$$\sum_{i=1}^{n} u_i \frac{\partial h_i}{\partial p_j} = 0.$$

By the assumption of non-satiation (i.e. marginal utility for every good is always positive), $u_i > 0$ for all i. By the law of compensated demand, which we showed previously in this chapter, we have $\frac{\partial h_j}{\partial p_j} < 0$. This means there must be some good i such that $\frac{\partial h_i}{\partial p_j} > 0$. $\quad\square$

If we want to examine possible substitution when budget is held fixed rather than when utility is held fixed, we have a similar definition using Marshallian demand instead of Hicksian demand.

Definition (Gross substitutes). We say that good i is a **gross substitute** for good j if

$$\frac{\partial x_i}{\partial p_j} > 0.$$

That is, good i is a gross substitute for good j if an increase in the price of j results in increased consumption of i under a fixed budget. Notice that this property is not symmetric. For a counterexample, consider two goods x and y with a quasilinear utility function $U(a, b) = a + \log b$. We encourage this counterexample as an exercise to the reader, but in this case, b is a gross substitute for a, but not vice versa.

Slutsky Equation with Two Goods

If we have only two goods, we can derive Slutsky's equation directly for more intuition. We start with the utility-maximization problem

$$\max_{a, b \geq 0} u(a, b) \text{ s.t. } p_a a + p_b b \leq y.$$

We substitute the budget constraint to obtain the unconstrained univariate optimization problem

$$\max_{a \in [0, y/p_a]} u\left(a, \frac{y - p_a a}{p_b}\right).$$

The first order condition is

$$u_a\left(a^*, b^*\right) - \frac{p_a}{p_b} u_b\left(a^*, b^*\right) = 0.$$

Together with the budget constraint

$$b^* = \frac{y - p_a a^*}{p_b},$$

these two conditions give implicit solutions for Marshallian demands $a^*(p_a, p_b, y)$ and $b^*(p_a, p_b, y)$. We differentiate the first order condition with respect to p_a with the Implicit Function Theorem, yielding

$$-\frac{u_b}{p_b} + \left(u_{aa} - \frac{p_a}{p_b}u_{ab}\right)\frac{\partial a^*}{\partial p_a} + \left(u_{ab} - \frac{p_a}{p_b}u_{bb}\right)\frac{\partial b^*}{\partial p_a} = 0.$$

Substituting our expression for b^* from the budget constraint yields

$$-\frac{u_b}{p_b} + \left(u_{aa} - \frac{p_a}{p_b}u_{ab}\right)\frac{\partial a^*}{\partial p_a} - \left(u_{ab} - \frac{p_a}{p_b}u_{bb}\right)\frac{1}{p_b}\left(a^* + p_a\frac{\partial a^*}{\partial p_a}\right) = 0.$$

We can rearrange these terms to get

$$-\left[u_{aa} - 2\frac{p_a}{p_b}u_{ab} + \left(\frac{p_a}{p_b}\right)^2 u_{bb}\right]\frac{\partial a^*}{\partial p_a} = -\frac{u_b}{p_b} - \frac{1}{p_b}\left(u_{ab} - \frac{p_a}{p_b}u_{bb}\right)a^*.$$

What does all of this mean? The left hand side is the comparative static $\frac{\partial a^*}{\partial p_a}$, multiplied by

$$S \equiv -\left[u_{aa} - 2\frac{p_a}{p_b}u_{ab} + \left(\frac{p_a}{p_b}\right)^2 u_{bb}\right],$$

which is the value that must be positive in the second-order condition for a maximum. We see the substitution effect

$$-\frac{u_b}{Sp_b} < 0,$$

which we can verify is the same as $\frac{\partial h_a}{\partial p_a}$ if we set up the expenditure-minimization problem accordingly. Lastly, we see the income effect

$$-\frac{1}{Sp_b}\left(u_{ab} - \frac{p_a}{p_b}u_{bb}\right)a^*$$

which is equal to $-\frac{\partial a^*}{\partial y}a^*$ if we compute $\frac{\partial a^*}{\partial y}$ with the implicit function theorem.

7.5 Income and Price Elasticity

One common way to measure how income and prices affect demand is to use elasticities. These are unit invariant ways of seeing how given change affects demand for a good. To motivate the use of elasticities, we can use them to examine some intuition behind the income and price effects of goods. To start, we can differentiate our original budget constraint,

$$y = \mathbf{p} \cdot \mathbf{x}$$

in order to identify a few important elements. Differentiating both sides with respect to y yields

$$1 = \sum_{i=1}^{n} p_i \frac{\partial x_i}{\partial y}$$

$$= \sum_{i=1}^{n} \frac{p_i x_i}{y} \frac{y \partial x_i}{x_i \partial y}.$$

We now define two new terms. We define

$$\alpha_i \equiv \frac{p_i x_i}{y}$$

to be the **budget share** of good i, which represents the fraction of our budget that we spend on good i. We define

$$\varepsilon_y^i \equiv \frac{y \partial x_i}{x_i \partial y}$$

to be the **income elasticity of demand** for good i. Intuitively, this tells us the percentage change in demand for a small percentage change in income. You may have seen the elasticity of income in previous economic classes defined as,

$$\varepsilon_y^i = \frac{\% \text{ change in demand}}{\% \text{ change in income}}$$

How can we square these two definitions? We can start by rewriting our definition of the elasticity to be

$$\varepsilon_y^i \equiv \frac{\partial x_i}{x_i} \frac{y}{\partial y} = \frac{\partial x_i / x_i}{\partial y / y}$$

We abuse the definition of the derivative slightly and treat it as a fraction even though this is not the case formally. Notice however that $\frac{\partial x_i}{x_i}$ is a small change in x_i divided by current consumption of x_i, which is a small percentage change in x_i. Following the same logic, $\frac{\partial y}{y}$ is a small percentage change in income. So, our definition of income elasticity is in fact equivalent to the interpretatation as small percentage changes.

Another sometimes useful fact is that the income elasticity is definable in terms of log derivatives:

$$\varepsilon_y^i = \frac{\partial x_i y}{x_i \partial y} = \frac{\partial \log x_i}{\partial \log y_i}$$

In general, you can think of a change in the value of the logarithm of a value as a percentage change of said value.

Notice the relationship between income elasticity and the income effect for a good: the income effect for a good i is negative if and only if $\varepsilon_y^i > 0$, which makes sense given our intuitive understanding of elasticity.

Substituting these new terms into our previous equation, we have the relationship

$$\sum_{i=1}^{n} \alpha_i \varepsilon_y^i = 1.$$

The term on the left is a weighted average of the different ε_y^i terms, since the some of all our α_i terms must be 1. Since the term on the left is in units of elasticity, we can almost think of this relationship as the idea that with a 1% increase in budget, we can buy 1% more "stuff" (not necessarily 1% more quantity of consumption), where the composition of this "stuff" is determined by each good's budget share and income elasticity. Notice that this relationship also tells us that it is impossible for every good in our option set to have a negative income elasticity; if our budget increases, we should not be buying less of everything. Later, we will see that this means that it is impossible for every good in our option set to be an inferior good.

We can derive a similar relationship to help us understand the substitution effect if we instead differentiate our original budget constraint with respect to the price of some good p_j. From the product rule, we get

$$0 = x_j + \sum_{i=1}^{n} p_i \frac{\partial x_i}{\partial p_j}.$$

Multiplying by $\frac{p_j}{y}$ and rearranging yields

$$-\frac{p_j x_j}{y} = \sum_{i=1}^{n} \frac{p_i x_i}{y} \frac{p_j \partial x_i}{x_i \partial p_j}.$$

We define

$$\varepsilon_{p_j}^i \equiv \frac{p_j \partial x_i}{x_i \partial p_j}$$

as the **cross-price elasticity of demand**, which represents how sensitive our demand for good i is in response to changes in the price of some other good j. When $i = j$, then this term is simply the **price elasticity of demand**. If $\varepsilon_{p_j}^i = 1$ always, then every 1% increase in the price of good j results in a 1% increase in the consumption of good i.

Substituting our definitions for cross-price elasticity and budget share, we get the relationship

$$-\alpha_j = \sum_{i=1}^{n} \alpha_i \varepsilon_{p_j}^i.$$

This equation tells us that the weighted average of cross-price elasticities of all goods with respect to price changes for good j is equal the negative budget share of good j. We will more closely examine the $\frac{\partial x_i}{\partial p_j}$ piece of our definition of cross-price elasticity later in the section on the general Slutsky equation.

Types of Goods

We might already have some intuitive idea of what goods have a negative income effect (i.e. we consume more when we get richer) and what goods have a positive income effect (i.e.e we consume less when we get richer. We can further categorize these types of goods using our definition of income elasticity:

Inferior goods have a negative income elasticity $\varepsilon_y^i < 0$, so their income effect is positive. This is equivalent to saying that $\frac{\partial x_i}{\partial y} < 0$. That is, we buy less of such goods as we become richer (e.g. cheap food). Recall from the previous section that it is possible for some or no goods to be inferior, but it is impossible for all goods to be inferior.

Normally, the demand-curve for inferior goods is still downward-sloping; that is, we still have $\frac{\partial x_i}{\partial p_i} < 0$. This is because even though we have positive income effect $-x_i \frac{\partial x_i}{\partial y} > 0$, typically the negative substitution effect $\frac{\partial h_i}{\partial p_i} < 0$ still dominates. Intuitively, an increase in the price of good i has two effects for an inferior good: (1) it makes the consumer poorer, which influences consumption positively from the positive income effect of inferior goods, and (2) it makes the consumer substitute away to relatively cheaper options, which influences consumption negatively. If the second effect dominates, as is often the case, then the demand curve for an inferior good is still downward sloping.

However, it is theoretically possible for the former effect to dominate, causing demand to be *upward-sloping*, or $\frac{\partial x_i}{\partial p_i} > 0$. That is, when prices for good i rise, the consumption of good i rises. These goods are called **Giffen goods** and are a rare subset of inferior goods. This phenomenon occurs if

$$\left| x_i \frac{\partial x_i}{\partial y} \right| > \frac{\partial h_i}{\partial p_i},$$

which generally happens when

$$x_i \frac{\partial x_i}{\partial y} \ll 0.$$

Intuitively, for the income effect to be so large, the good must be very inferior (i.e. $\varepsilon_y^i \ll 0$) and the budget share of x_i must be very large, which is generally the case when the consumer is poor. A famous example of this phenomenon was presented by Jensen and Miller (2008). They found that subsidizing the price of rice in rural Hunan, China, resulted in decreased consumption of rice among poor households. Why? Poor households who used to spend a large fraction of their income on rice became effectively richer through this subsidy, so they no longer needed to buy as much rice since they could afford other foods in their diets.

Normal goods have positive income elasticity $\varepsilon_y^i > 0$, so their income effect is negative. This is equivalent to saying that $\frac{\partial x_i}{\partial y} > 0$. That is, we consume more of them when we become richer; most goods we typically think of fall under this category. The

demand curve is also unambiguously downward-sloping, since both the income effect and the substitution effect are negative. By the same reasoning as before, some or all of the goods in our option set must be normal.

It makes sense to buy more of a good when we become richer, but the relative size of this increase might be bigger for some goods than others. A normal good is a **luxury** if $\varepsilon_y^i > 1$, so a 1% increase in income results in more than a 1% increase in consumption of that good. The budget share of these goods increases as the consumer becomes richer. For example, a consumer might spend no money on luxury cars initially, but they might spend a sizable amount of money on luxury cars when they become rich.

Alternatively a normal good is a **necessity** if $\varepsilon_y^i < 1$, so a 1% increase in income results in an increase in consumption of that good by between 0% and 1%. The budget share of these goods decreases as the consumer becomes richer. For example, Engel's law is the observation that the budget share for food decreases when income rises, so food is a necessity.

In the special case where $\varepsilon_y^i = 1$ for all goods in the option set, we say that preferences are **homothetic**. That is, budget shares of all goods are fixed, so a 1% increase in income causes a 1% increase in the consumption of all goods.

7.6 Measuring Welfare Changes

One of the challenges in dealing with utility in economics is that it is not directly observable. However, it would be nice to be able to quantify the change in utility from a change in prices. Fortunately, the expenditure function offers us a nice way to do so.

Money Metric Utility

The money metric utility function offers us a way to measure utility changes in dollars, which is observable in the real world. In particular, it allows us a way to map changes in prices to a change in income. Suppose we have some utility function u that we want to be able to measure in dollars. Take two utility levels, $\bar{u}_1, \bar{u}_2$ such that $\bar{u}_1 > \bar{u}_2$. Recall that for any set of prices, $\mathbf{r}$, we have

$$e(\mathbf{r}, \bar{u}_1) > e(\mathbf{r}, \bar{u}_2)$$

This tells us that $e(\mathbf{r}, \cdot)$ is a monotonic transformation in its second argument, so it can be applied to any valid utility function to create a new utility function. So, for any allocation $\mathbf{x}$, we can define the **money metric utility function** with reference prices $\mathbf{r}$

as,

$$m(\mathbf{x}) = e(\mathbf{r}, u(\mathbf{x}))$$

The idea here is that we can measure how much utility we have for a given allocation by examining what is the minimum amount that we would need to spend to achieve the utilty for that allocation.

This gives us a natural way to compare utilities for two different sets of prices. Consider prices $\mathbf{p}$ and $\mathbf{q}$. For a fixed level of income y, we can compare their utilities through their respective value functions, $v(\mathbf{p}, y)$ and $v(\mathbf{q}, y)$. We can then convert these utility values into a money metric values to measure how much better (or worse) in dollars $\mathbf{p}$ is compared to $\mathbf{q}$

$$m(v(\mathbf{p}, y)) - m(v(\mathbf{q}, y)) = e(\mathbf{r}, v(\mathbf{p}, y) - e(\mathbf{r}, v(\mathbf{q}, y))$$

If the above expression is positive, then $\mathbf{p}$ is preferred to $\mathbf{q}$, and if it is negative, then $\mathbf{q}$ is preferred to $\mathbf{p}$. Importantly, these quantities are ones that we could in principle measure, by asking or observing how much money an individual requires at prices $\mathbf{r}$ to be indifferent with having prices $\mathbf{p}$ and income y.

It is important to note, however, that while the money metric utility can be measured, it still does *not* allow us to compare utilities across individuals.

Equivalent and compensating variation

So far, we have used the money metric utility function to measure the differences between two sets of prices, $\mathbf{p}$ and $\mathbf{q}$, with respect to some set of reference prices $\mathbf{r}$. However, there is a question of how we select this reference price $\mathbf{r}$. Two reasonable prices we might pick are either $\mathbf{p}$ or $\mathbf{q}$. For the sake of this analysis, we will refer to $\mathbf{p}$ as the *old* prices and $\mathbf{q}$ as the *new* prices.

Equivalent Variation

The **equivalent variation (EV)** tells us the welfare effect of a price change from $\mathbf{p}$ to $\mathbf{q}$ through an equivalent change in income at *old* prices, $\mathbf{p}$. Mathematically, we are setting the reference price $\mathbf{r} = \mathbf{p}$,

$$EV = e(\mathbf{p}, v(\mathbf{q}, y)) - e(\mathbf{p}, v(\mathbf{p}, y)) = e(\mathbf{p}, v(\mathbf{q}, y)) - y$$

Intuitively, we can think of EV as how much you would be willing to pay to keep prices at $\mathbf{p}$ instead of $\mathbf{q}$. An easy way to remember this is that the equivalent variation tells you what would have been the *equivalent* change in income for a given price change.

Compensating Variation

The **compensating variation (CV)** tells us the welfare effect of a price change from $\mathbf{p}$ to $\mathbf{q}$ through an equivalent change in income at *new* prices, $\mathbf{q}$. Mathematically, we are setting the reference price $\mathbf{r} = \mathbf{q}$,

$$CV = e(\mathbf{q}, v(\mathbf{q}, y)) - e(\mathbf{q}, v(\mathbf{p}, y)) = y - e(\mathbf{q}, v(\mathbf{p}, y))$$

Intuitively, we can think of CV as how much you would be willing to pay to change prices from $\mathbf{p}$ to $\mathbf{q}$. Or equivalently, it is the negative of how much you would need to be paid after a price change to make you just as happy as you were before. This second interpretation is useful for remembering the definition of compensating variation, as it tells you how much you need to be *compensated* to be just as happy for a given price change.

Hicksian Integral

The equivalent and compensating variations have an interesting connection to the Hicksian demand function when we consider the change in only one price. Suppose that we have a single price change, from p to q.

First, we will examine the equivalent variation, which is

$$EV = e(p, v(q, y)) - e(p, v(p, y)) = e(p, v(q, y)) - y$$

However, we also have that $y = e(q, v(q, y))$. So we can rewrite the equivalent variation as,

$$EV = e(p, v(q, y)) - e(q, v(q, y))$$

Next, notice that for an arbitrary price s, we have by Shephard's Lemma,

$$\frac{\partial e(s, v(q, y))}{\partial s} = h(s, v(q, y))$$

Where h is the Hicksian demand for the single good. Then, by the fundamental theorem of calculus, we have

$$\begin{aligned}
EV &= e(p, v(q, y)) - e(q, v(q, y)) \\
&= \int_q^p \frac{\partial}{\partial s} e(s, v(q, y)) \, ds \\
&= \int_q^p h(s, v(q, y)) \, ds
\end{aligned}$$

We can perform a similar process with the compensating variation to obtain,

$$CV = e(p, v(p, y)) - e(q, v(p, y))$$
$$= \int_q^p \frac{\partial}{\partial s} e(s, v(p, y)) \, ds$$
$$= \int_q^p h(s, v(p, y)) \, ds$$

Note that while we handled these in the single good case, it generalizes to multiple goods. A simple way to see this is to consider multiple goods, but only changing the price of one good. Then we obtain the same results as above except the integral is with respect to the price of that specific good, and the the Hicksian is the Hicksian demand of that specific good.
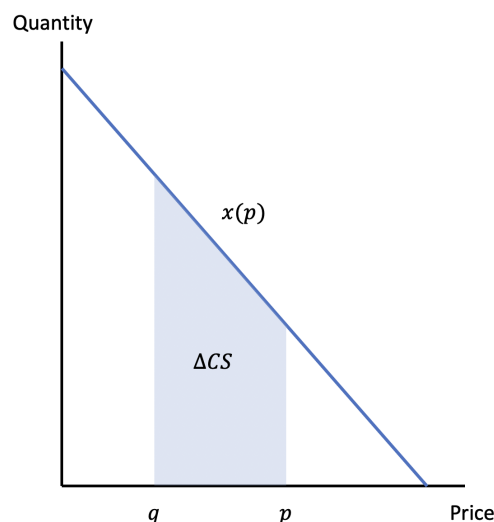
Consumer Surplus

You may be more familiar with another way of measuring changes in welfare: consumer surplus. In an introductory economics class, consumer surplus is presented as the the difference between how much consumers are willing to pay and the price they are charged, aggregated over all consumers.

Suppose we faced a change in prices from p to q where $q < p$. Under the introductory economics interpretation, the change in consumer surplus (ΔCS) would look something like this:
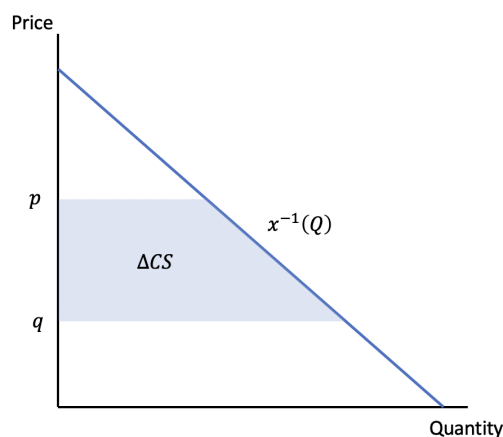
However, the problem with this approach relies is that it relies on there being many consumers each willing to buy one unit of the good for this line of logic to work. It might also make sense for an individual to have consumer surplus. Ideally, we would want a way to find the change in consumer surplus given an individual's marshallian demand curve. To do so, we can express the above chart with flipped axes so that we can write our demand function as the more familiar Marshallian demand, $x(p)$,

It now becomes clear that the change in **consumer surplus** is the area under the Marshallian demand curve from q to p. We can express this as an integral to obtain,

$$\Delta CS = \int_{q}^{p} x(s)\, ds$$

This might look familiar because it is the same as the integral for EV and CV, except we integrate the Marshallian demand function rather than the Hicksian demand functions. This measure also works both for an individual, or an aggregate demand curve derived by summing many individual Marshallian demands.

Comparing the welfare measures

Now that we have three different measures of welfare changes, it is worth asking how they compare to each other. For simplicity, we will examine the case where the price of a good decreases from p to q.

Since we have expressed equivalent variation, compensating variation, and consumer surplus as integrals of Hicksians and Marshallians, we can determine their relative values by comparing the Hicksian and Marshallian demand functions. We will try to examine these functions graphically.

Normal good

First, we will consider the case where the good is normal. This means that the income effect is positive for a price decrease, so the slope of the Marshallian demand curve is steeper than that of the Hicksian demand curve. We also know that $v(q, y) \geq v(p, y)$ because $q < p$, so the agent prefers price q. Since the good is normal, this implies that $h(s, v(q, y)) \geq h(s, v(p, y))$ for all prices s because if we need to achieve a greater utility level, we will demand more of the normal good. Finally, recall the optimality

conditions that $h(p, v(p, y)) = x(p)$ and $h(q, v(q, y)) = x(q)$. This gives us sufficient sufficient information to plot the functions as follows:
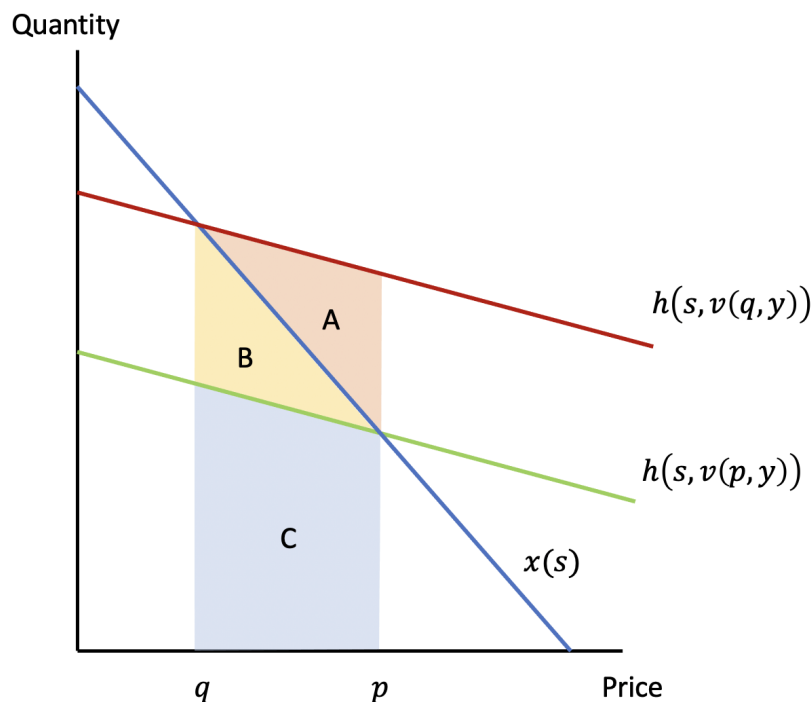
Based on this graphic, we have

$$
\begin{cases}
EV & = A + B + C \\
\Delta CS & = B + C \\
CV & = C
\end{cases}
$$

So, in the case of a normal good with a price decrease, we have

$$EV \geq \Delta CS \geq CV$$

The inequalities are reversed if we have a normal good with a price increase. It is worth noting the condition under which we have equality. In order for the thre quantities to be equal, we would need the slope of the Marshallian demand curve to be the same as the slope of the Hicksian demand curves. The only way this happens is if there is no income effect.

Inferior good

Next, we will consider what would happen if the good is inferior. In this case, the slope of the Marshallian demand curve would be less steep than the the slope of the Hicksians since the income and substitution effects would move in opposite directions. We would

still however need the points of intersection for the Hicksians and the Marshallian to remain the same. This would result in the following:
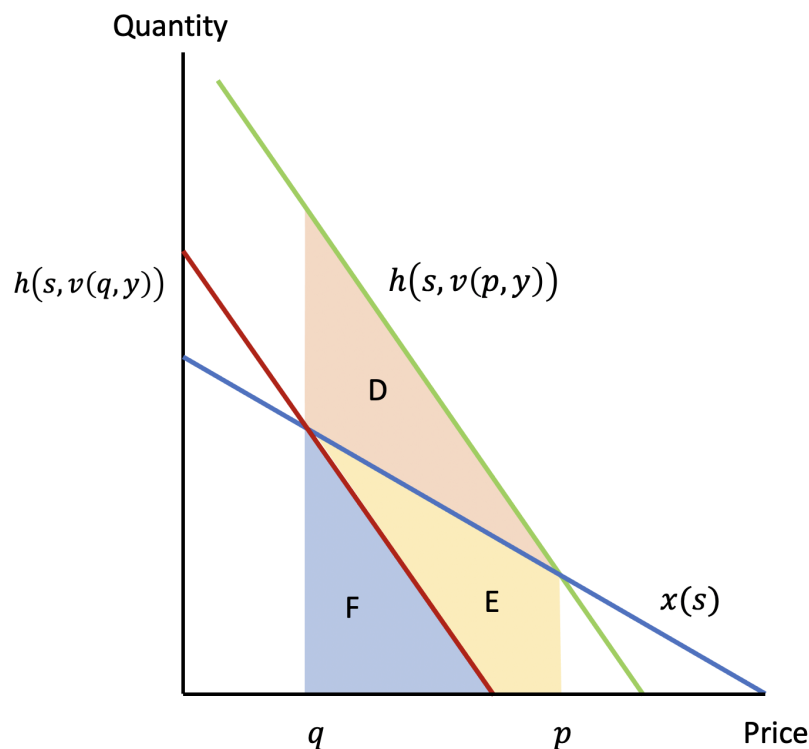
This results in the following values:

$$
\begin{cases}
EV & = F \\
\Delta CS & = E + F \\
CV & = D + E + F
\end{cases}
$$

So, for an *inferior good*,

$$EV \leq \Delta CS \leq CV$$

Once again, if you have a price increase instead of a price decrease, the inequalities are reversed.

Note however that regardless of whether we have a normal or inferior good, the change in consumer surplus is always between the equivalent and compensating variations.

7.7 Measuring Price Changes

One important application of our study of the expenditure function is a metric for measuring price changes, or a **price index**. Prices for goods often change, but what would be a good way to aggregate different price changes across multiple goods into a single metric?

One appealing approach is to use the expenditure function. Suppose we start with utility u_0 when prices are at $\mathbf{p}^0$, and then prices change to $\mathbf{p}^1$. A natural metric might measure how much expenditure would need to change in order to maintain utility u_0 at this new price level $\mathbf{p}^1$. Mathematically, this definition would look like

$$\frac{e(\mathbf{p}^1, u_0)}{e(\mathbf{p}^0, u_0)}.$$

As another way to interpret this definition, if we had income y_0 when prices were $\mathbf{p}^0$, and then prices changed to $\mathbf{p}^1$ and our income changed to $\frac{e(\mathbf{p}^1, u_0)}{e(\mathbf{p}^0, u_0)} y_0$, then our utility would be the constant.

The definition above is wonderful in theory but cannot be measured in practice: since we cannot observe consumers' utilities, we also cannot observe their expenditure functions. However, one workaround is to recognize that

$$e(\mathbf{p}, u) = \mathbf{p} \cdot \mathbf{h}(\mathbf{p}, u)$$

by definition, and that

$$\mathbf{x}(\mathbf{p}, y) = \mathbf{h}(\mathbf{p}, v(\mathbf{p}, y))$$

by duality. It is important to understand why the equation above holds: the utility-maximizing consumption given a budget (Marshallian demand) is the same as the expenditure-minimizing consumption to achieve this realized utility level (Hicksian demand). This observation is useful because although we cannot measure the expenditure function, we can measure Marshallian demand, since $\mathbf{x}, \mathbf{p}$, and y are all observed.

Using these formulations, we can approximate

$$\frac{e\left(\mathbf{p}^1, u_0\right)}{e\left(\mathbf{p}^0, u_0\right)} \approx \frac{\sum_{i=1}^n p_i^1 x_i^0}{\sum_{i=1}^n p_i^0 x_i^0},$$

where the approximation on the right-hand side is known as the **Laspeyres price index**. Notice where this approximation is not exact. The denominators are equal, since

$$e\left(\mathbf{p}^0, u_0\right) = \sum_{i=1}^n p_i^0 x_i^0$$

by definition of the expenditure function, but the numerators differ:

$$e\left(\mathbf{p}^1, u_0\right) \leq \sum_{i=1}^n p_i^1 x_i^0.$$

Intuitively, this is because we know $\mathbf{x}^0$ achieves our minimum utility u_0, but purchasing this same bundle might no longer be the minimum expenditure to achieve this utility

once prices change. Thus,

$$\frac{e\left(\mathbf{p}^1, u_0\right)}{e\left(\mathbf{p}^0, u_0\right)} \leq \frac{\sum_{i=1}^n p_i^1 x_i^0}{\sum_{i=1}^n p_i^0 x_i^0},$$

so the Laspeyres price index is an overestimate of our theoretical ideal. If we had an income y_0 initially and multiplied it by the Laspeyres price index after the price change, we would be able to weakly increase our utility, since we can still afford our old consumption bundle, but we may be able to reoptimize and do even better.

A similar approach to the same problem is the **Paasche price index**

$$\frac{e\left(\mathbf{p}^1, u_1\right)}{e\left(\mathbf{p}^0, u_1\right)} \approx \frac{\sum_{i=1}^n p_i^1 x_i^1}{\sum_{i=1}^n p_i^0 x_i^1},$$

where we now weight prices by $\mathbf{x}^1$ instead of $\mathbf{x}^0$. The theoretical ideal on the left is a similar idea but uses the new utility level as the baseline rather than the old utility level. The Paasche price index on the right is now an underapproximation of the theoretical ideal on the left, since while the numerators are equal, the denominators differ as

$$e\left(\mathbf{p}^0, u_1\right) \leq \sum_{i=1}^n p_i^0 x_i^1$$

by a similar reasoning to before. Intuitively, the Laspeyres price index measures how much a consumption bundle that was optimal in the base year costs now, whereas the Paasche price index measures how much a consumption bundle that is optimal now would have cost in the base year.

Recap

In this chapter, we discussed how we could derive the Hicksian demand function and the resulting expenditure function. This allowed us to derive the famous Slutsky equation, which decomposed the effects of a price change into a substitution effect and an income effect. We defined normal goods as those with a negative income effect in response to a price increase, while inferior goods have a positive income effect in response to a price increase. We also examined how we could use the expenditure function to derive a measurable utility function, as well as how to use that utility function to measure changes in welfare from price increases. Finally, we examined how the expenditure function could be used to construct an index of prices to see how prices have changed in the economy as a whole.