Lecture Notes on Microeconomic Theory

Tilman Börgers
tborgers@umich.edu

August 9, 2022

I am very grateful to Ashira Chugh, Amy Orlov, Joshua Moss, and Nathan Shen for reading this manuscript and providing me with extremely helpful comments. I am also grateful to Nur Syafiqah Binti Adam Malik for correcting many errors in the first version of these notes.

The Master said, "A person is worthy of being a teacher if he is able to gain new insights from chewing over what he already knew."

Confucius, The Analects, 2.11.

I have used Annping Chin's translation of Confucius' *Analects*, Penguin Books, 2014.

Table of Contents

Topic 1: Modern Research in Economics

Before I start explaining the basics of microeconomic theory, I want to give you a bigger picture: where does microeconomic theory fit in contemporary economic research? To find out what contemporary research in economics does, we shall look into one of the most prominent academic journals in economics, *The American Economic Review.* The front page of the September 2017 issue of the *The American Economic Review*, which includes a table of contents, is shown on the next page.

There are 11 articles. Some of the titles of the articles are a little mysterious, and don't give away what the article is about, even for someone reasonably well trained in economics. For example, I can't guess without reading the paper what is meant by: "The Fundamental Surplus." On the other hand, there are articles the titles of which even a layperson understands easily: "Why are Indian Children So Short? The Role of Birth Order and Son Preference" or "Virtual Classrooms: How Online College Courses Affect Student Success." And some titles make sense to the expert, but to few other people. For example, I happen to be able to imagine what "Optimal Allocation with Ex Post Verification and Limited Penalties" might be about, but I guess most of my readers are not.

Perhaps you are surprised by the variety of topics. For some articles, you might even be surprised to see them in an economics journal at all, such as the article on "Why are Indian Children so Short?" Let's therefore look first at what economists do research about, and then ask how economists do their research. Finally, we'll discuss the role of microeconomic theory in the bigger picture.

What Do Economists Do Research About?

Economics is divided into many fields, depending on which aspect of economic life is studied. Let's see which fields of economics are represented in this issue of the *The American Economic Review*. Some fields are defined by the sort of market they are interested in.

The American Economic Review

ARTICLES

SEPTEMBER 2017

Figure 1: Titlepage of the September 2017 issue of *The American Economic Review*

Labor Economics: This field studies the functioning of labor markets. Two articles in this issue belong to this field. I identify articles by listing their authors.

- *Ljungqvist and Sargent*, the article with the mysterious title, study how to explain fluctuations in the unemployment rate.
- *Garicano and Rayo* write about the incentives created by contracts between an employer and a trainee, for example, the incentives for the employer to train the trainee properly.

International Trade: The name of this field is self-explanatory.

- *Antràs, Fort, and Tintelnot* ask how firms decide whether to import inputs, how firms that choose to import inputs are different from firms that do not import inputs, and whether firms that import inputs from abroad therefore import fewer inputs domestically.
- *Handley and Limão* investigate why China's becoming a member of the "World Trade Organization" in 2001 was followed by an increase in trade between the United States and China, even though the tariffs that applied to such trade did not really change.
- *Brandt, Van Biesebroeck, Wang and Zhang* study the effect of China's membership of the "World Trade Organization" on the productivity and on the price choices of Chinese manufacturing firms.

Development Economics: This is the field that is concerned with life in economies with a relatively low living standard.

- *Jayachandran and Pande* observe that children in India are smaller in size than children in other countries, say African countries, and ask whether this can be explained by parents' preferences among their children, in particular by favoritism for the eldest son.

Economic History:

- *Pascali* investigates the second half of the 1800s, a period in which steamships started to replace sailing ships in international trade. The article is about how the introduction of steamships changed the volume and patterns of trade, and how the participating countries were affected by these changes.

So far, I have classified only 7 out of 11 papers in this issue of *The American Economic Review*. 3 of the remaining articles represent the increasing tendency of economists to do research on areas that in the past have been regarded as other subjects' domain. An example is *Martin and Yurukoglu*, who try to estimate the impact of ideological slant in 24-hour news channels, such as CNN or Fox News, on the ideological polarization among American voters. There is also one paper, *Mylovanov and Zapechelnyuk*, that is directly about the subject of our course, microeconomic theory, and I postpone a discussion of this field of economics until the end of this section.

Going through these articles leaves us with a few simple insights: Economics studies a large variety of topics. Most articles ask practical questions about the real world. (Hopefully, you found the questions asked interesting. Perhaps you have found yourself tempted to take a peek at one or two of the articles I mentioned.) Economics is organized into fields. In fact, there are many more fields than are represented in this particular issue of *The American Economic Review.*

As we proceed in this Section we want to establish what the role of Microeconomic Theory is in this universe of research, and, in particular, how what you will learn in an introductory course in microeconomic theory relates to the world of contemporary economics research. The next section is a first step towards this goal.

How Do Economists Do Research?

To explain how economists do research, we shall focus on the article by Antràs and his co-authors (*Antràs et. al.*). Recall that it focuses on firms' choices of which inputs for their production to import from abroad. Is that a narrow question, not of general interest? In fact, this article is directly relevant to the political argument in the United States over the question whether trade with Canada, Mexico, or China, hurts the United States' workers. Indeed, *Antràs et. al.* provide a detailed analysis of the effects of trade with China on different sectors in the United States and find that in some sectors, increased trade with China encourages firms to also buy more domestically produced inputs.

Our discussion of *Antràs et. al.* will focus on methodology, however. There is a convention in the writing of economics articles that is useful for us. The convention is that each article has an Introduction, and that at the end of the Introduction there are typically a paragraph or two that explain how the paper is organized. The last paragraph of the Introduction to the article by Antràs and his co-authors (*Antràs et. al.*) is:

> The rest of the paper is structured as follows. We present the assumptions of our model in Section I, and solve for the equilibrium in Section II. In Section III we introduce the data and provide descriptive evidence supporting the assumptions underlying our theoretical framework. We estimate the model structurally in Section IV, and in Section V, we perform our counterfactual analysis and compare the predictions of the model to reduced-form evidence. Section VI concludes.

Obviously, if you are not a professional economist, you will not understand all of this. I want to draw your attention to two words: model and data. Let's first talk about data. Almost all of the articles in this issue of *The American Economic Review* that make real world claims have data. Moreover, for many of the articles, you can find all data on which the article relies on the website of *The American Economic Review*, together with instructions on how to reproduce the papers' results. Good economics research today is not primarily an expression of an opinion. It does not resemble in any way an opinion article in the *New York Times* or the *Wall Street Journal,* not even if that opinion article is written by an economics professor. Good economics research is evidence based, and reproducible.

What are the data in *Antràs et. al.*? Not surprisingly, data collected by the "US Customs and Border Protection" (before 2004: "US Customs Services") are used. But other data, collected by US government agencies, are also used. Some of these data are confidential. An online appendix describes how to apply for access to these data.

What do *Antràs et al.* do with their data? We now turn to the second word that I emphasized earlier, the word "model." Sections I and II of the article describe the model. These sections are full of equations. For example, the second sentence of Section I states:

> "Consider a world consisting of J countries in which individuals value the consumption of differentiated varieties of manufactured goods according to a

standard symmetric CES aggregator

$$U_{Mi} = \left(\int_{\omega\in\Omega_i} q_i(\omega_i)^{(\sigma-1)/\sigma} d\omega\right)^{\sigma/(\sigma-1)}, \tag{1}$$

... (the sentence continues).

What are we to make of this? There are more mathematical symbols here than, most likely, you will ever learn to understand in your career as an economics student. Nonetheless, let's try and unpack a little bit what is going on.

First, the obvious: a model is something mathematical. More specifically: it is a mathematical description of how that aspect of the economy works from which the data reported in the paper are taken. That is: how inputs are transformed into outputs, how decisions about inputs and outputs are made, how much of the output firms will be able to sell, how this depends on the prices firms choose, how firms will set their prices, etc.

As an example, consider equation (1), which I quoted above. It describes how much enjoyment people derive in *Antràs et. al.*'s model from the consumption of the manufactured goods that the paper deals with. The left hand side is a symbol for this enjoyment ("utility" in the economists' words; we shall later discuss this concept more carefully). The right hand is a formula that involves the quantities of the various goods that a person buys. Thus, the utility depends on these quantities. The utility will guide demand for the manufactured products in Antràs et. al.'s model, and the demand will be a contributing factor in firms' decisions which inputs to buy from abroad. For example, the size of demand might be important for the question whether it is worthwhile for a firm to make the investment into acquiring the skills needed to buy inputs abroad.

What the specific formula is, why there is an integral sign "$\int$," etc. need not concern us here.

But here is something puzzling: how can it be that the mathematical description of the world of production of manufactured goods, imports, prices of manufactured goods, etc., comes in the paper *before* we even see the data? Remember that the model is in Sections I and II, but the data are in Section III.

To understand this, we need to notice an important point about formula (1). It is not a complete description of peoples' utility. There is one symbol, σ (the Greek letter "sigma"), that has not been specified. It stands for some number greater than zero. But the formula for utility, and therefore the demand for the manufactured goods, will be different, depending on what the value of σ is.

When describing their model, *Antràs et. al.* don't tell us what the value of σ is. They describe their model for any arbitrary value of σ.

In fact, the value of σ is one of the things that *Antràs et. al.* want to figure out from their data. Indeed, in a later part of the paper, *after* invoking their data, they conclude that $\sigma = 3.85$. There is some economic interpretation of this number, but for this discussion, which is primarily about methodology, it does not matter what that interpretation is.

Thus, *before* even seeing the data, *Antràs et. al.* make assumptions about consumers' utility that are expressed in formula (1). They then look at their data *assuming* that this formula is correct, and that the only remaining open question is what is σ. It is only this last question that is resolved by the data. There are many more prior assumptions in Section I, and also in Section II, of *Antràs et. al.* article. These involve further symbols the value of which is not specified, but later derived from the data. Using a metaphor: the model constitutes the lenses through which *Antràs et. al.* look at their data, and those lenses, together with what is out there, determine what the authors see.

To stick with the metaphor, when taking pictures, you might use lenses that produce a black and white photograph. What you see, is determined by the lenses, and by what you photograph. The lenses will direct your attention: instead of focusing on colors, you will focus on light and dark when looking at the picture. If that is what you want to focus on, then the lenses are very useful.

But why do economists need lenses? Don't the lenses just distort what they are seeing? What is the benefit to *Antràs et. al.* of having a theoretical model? Like color filtering lenses, the lenses of economists, their models, allow them to focus on particular aspects of what they see. *Antràs et. al.*'s model provides them with a focus when they look at their data, and prevents them from getting lost in an ocean of data. Somehow, we have to come up with the aspect of the data that we want to investigate. By formulating a model, *Antràs et. al.* are very clear about what it is that they want to learn from their data.

There is another, separate benefit of having a model. It allows researchers to ask questions about economic policy: If *Antràs et. al.*'s model of how the world of trade works is correct, and if their estimates of the various variables are correct, then they can introduce government policies that don't yet exist in the real world and simulate their effects. This line of inquiry happens to not play a huge role in *Antràs et. al.*, but in other articles it is often an important part of the motivation for the use of a theoretical model. It is the main way in which economists can provide policy advice.

By the way: the model that *Antràs et. al.* use does not precisely fit the data. It would be a miracle if it did. Therefore, they consider the data assuming that the model is in principle true, but that there are also some random effects not captured by the model. To deal with these random effects, *Antràs et. al.* use methods from statistics. Some of the statistics methods that are used by economists have specifically been developed by economists for the study of economic data. In that case, the field that studies these methods is called econometrics. We can thus say more precisely that *Antràs et. al.* use econometric methods when turning to the data.

Summing up, we can distill from *Antràs et. al.*'s paper a stylized image of how economists do research:

- Start with an interesting question about the real world, or about economic policy.
- Develop a mathematical model that seems suitable for the purpose, but that is sufficiently flexible so that there are at least some value of the model's parameters such that the predictions of the model are not too far from the data.
- Use data to figure out the most likely values of the parameters.
- Use the understanding of the world that you have developed in this exercise to make predictions about how different economic policies will affect the world.

Does *all* research in economics have this structure? No. In fact, we can use the September 2017 issue of *The American Economic Review* to see some other styles of research. A number of papers by-pass the use of economic theory, and instead directly use tools from econometrics, in particular the most popular among them, the *regression analysis*. For example, a central hypothesis in *Jayachandran and Pande*'s study is that one important factor that determines the height of Indian children is the preferential treatment that some parents give to their first born son. To test this hypothesis they compare states of India in which the tradition is that the eldest son plays an important role

The details of what motivates this hypothesis are not important for the point that I am making.

in the family hierarchy, and states in which this tradition is not so prominent. They check whether in the former states children are on average shorter than in the latter states. This is a straightforward question that can be answered by looking at the data. *Jayachandran and Pande* don't need a model to think about this. They use some form of regression analysis instead, where the role of the regression analysis is to account for the fact that many other variables may also influence the children's' height.

Pascali, who is interested in the impact of steamships on trade volume, trade patterns, and economic well-being of countries, also uses regression analysis, without explicit economic theory, as a central tool in his study. For example, an important piece of his analysis is the question how different the trade volumes are for countries that are close to each other when steam boats are used, how different they are for countries that are close to each other when sailing boats are used to travel, and how these differences have changed over time in the 1800s. No sophisticated mathematical theory is needed to see that this is a relevant question, and it can be answered by straightforward regression analysis.

But the use of economic models is wide-spread. The two labor economics papers, and two of the three papers on international trade, base at least parts of their data analysis on theoretical models, as does the paper on the effect of bias in cable news. Why some papers are explicitly based on theoretical models, and other papers not is an interesting question that we cannot discuss here.

The Role of Microeconomic Theory in Economic Research

The "big picture" of research in economics that I have developed in the previous section suggests that, before you can do economic research, you have to first learn two sets of helpful skills. One is the set of skills needed to write mathematical models. The second is the set of econometric skills needed when taking microeconomic models to real world data.

There are actually two "fields" of economics in which research is primarily concerned with developing the ideas used in economic modeling and the ideas used in econometrics. These fields are sometimes referred to as the field of "theory," and the field of "econometric theory." Researchers in these fields sometimes, but not always, focus on specific economic application. Often they try to identify important general ideas and focus on these in the abstract.

These researchers' work often looks like pure mathematics rather than economics. In the September 2017 issue of *The American Economic Review*, the paper by *Mylovanov and Zapechelnyuk* belongs to the field of "theory." It will therefore appear to a lay reader particularly inaccessible.

A disclosure: I am myself a "theorist." But in writing these notes, I have had firmly in mind that my job is to bring you closer to being able to read, and perhaps some day conduct, applied economic research. My purpose here is not to steer you into the direction of specializing in economic theory.

These notes are meant to help you develop the first set of skills referred to above, that is, the skills needed to write mathematical models. These notes do not deal with the second set of skills. This is why in these notes I can not demonstrate to you all the steps that are involved when one wants to use the models discussed here to investigate real world economic policy questions. We would need to have data, and we would need have econometric skills, too, to accomplish this. In the language of economics, these notes focus on "theory."

Unfortunately, the mathematics used in the formulation of mathematical models these days is often very sophisticated. Most beginners don't have the knowledge needed to deal with such models. Here, we shall therefore only describe extremely simplified versions of the models actually used in real economics research. The models that we study here might seem to you sometimes too simple, too naive, to be at all connected to the real world. But this is a wrong impression. Yes, what we are going to do involves huge simplifications. But, even in our very simplified setting, you will, for example, encounter soon a formula that looks a lot like the one in equation (1), and you will soon understand more about the way in which such a formula is used in economics.

Moreover, even though we focus on "theory," we shall not completely neglect the real world, the data, the policy issues. We shall approach them, though very tentatively, and very cautiously. But, as you work through the material here, I shall try to not let you completely forget the bigger picture of economics research.

What we are going to study is called microeconomic theory. This suggests that there are other types of theory. And, indeed, there are books, classes, videos, etc., on another type of theory, macroeconomic theory. In all of this first topic, I have not differentiated between *microeconomic* and *macroeconomic* theory. The reason is simple: the distinction between *microeconomic* and *macroeconomic* theory is a little old-fashioned. There was a period in the development of economics when this was a very important distinction, say in the 1960s. Today, the situation is different: it makes most sense (to me) to think of *macroeconomics* as another field, alongside fields such as international

What follows is a personal opinion. Not everyone in the profession shares it.

trade, labor economics, etc. Specifically, macroeconomics is the field concerned with fluctuations in the total value of production in an economy, and with the growth or shrinkage of that production. Modern macroeconomics also formulates models using (microeconomic) theory, and takes these models to data, just as researchers in international trade or in labor economics do.

The paper by *Ljungqvist and Sargent*, that I earlier put into the "labor economics" category, could also go into the "macroeconomics" category.

So, let us now embark on our study of (microeconomic) theory!

Topic 2: Key Ideas of Microeconomic Theory

There are two basic ideas in economic theory: rationality, and equilibrium. Economic *theory* studies these ideas' meaning and their logical implications in general, and in a variety of applications. Both ideas are frequently misunderstood, by students, but also in textbooks, and, dare I say, even by economics professors. In this topic I shall offer a little preview of these key ideas.

Of course, I simplify a lot here. Many economists who read this will protest, and mention that they have had a couple more ideas than just these two. But to start with, it is useful to focus on these two ideas.

Rationality

"Rationality" is in economics a property that peoples' choices may, or may not have. That is, you may make rational choices, or perhaps you don't. But it is a black or white picture. We don't say that one choice is "more rational" than another.

Economics often hypothesizes that peoples' behavior *is* rational. This is often criticized as an "implausible" assumption. The rationality assumption has been the subject of much criticism, often from people who themselves are not economists.

Of course, economists test their hypotheses using data. Economists maintain the rationality hypothesis only if it is not clearly refuted by data. Introspection suggests that we all make irrational choices, and therefore you may be surprised that the theory of rational choices is not regarded in economics as unambiguously refuted, but only as refuted in certain areas of human behavior, not in others.

We need to be very clear about the meaning of the theory of "rational" choices in economics. The word "rationality" has a meaning in economics that is *not* the same as the meaning of the word "rationality" in everyday language. This is a source of endless confusions in the debate about the rationality assumption in economics. People think economics assumes "rationality" in the everyday sense. But we really assume it in a very specific, economic sense.

Of course, the meaning of "rationality" in everyday language cannot be defined with complete precision. We might mean by a "rational" choice one that pursues some “reasonable” objective. For example, we might say that buying a huge house with lots of glass windows in the coldest parts of Alaska is not “rational,” if you just think about the gas energy bills that the owner has to pay to keep the house warm. This makes sense in everyday language. But it is not something that we would say in economics. In economics, choices are “rational” if they are best according to the decision maker’s preferences. But what those preferences are is left to the decision maker. Everyone can have their own preferences. If you like huge houses with lots of glass windows so much that you are willing to pay the heating bill even in Alaska, that is up to you. Economics is not judgmental. Economics would not call this “irrational.”

In everyday language, another part of the definition of a “rational” choice is that the choice is the result of careful reasoning. Maybe a rational choice of college is a choice that is based on a table of pros and cons for each college that a student considers. But this is not what economics means by “rationality.” In economics, even a routine choice can be “rational.” What matters is only the action itself: is it optimal according to the decision maker’s preferences? It does not matter how the choice came about. Even if you picked the University of Michigan because all your family went here, it may still be a “rational” choice for you, perhaps because your share your family’s preferences.

Preferences

I have used the word preferences when explaining what rationality means in economics. But I should be more careful about what we mean by “preferences.” I shall explain the concept of “preferences” in this section. Then, once I have explained “preferences,” I go back and give a more precise definition of “rationality,” as the word is used in economics.

Let’s think about a simple case: suppose I went out with you to a restaurant, and looked with you over the menu. Before we order, I could attempt to find out your preferences over food. For example, I could ask you: among the “beef burger with fries” and the “stir fried vegetables with rice,” which one do you prefer? Suppose your answer were “stir fried vegetables with rice.” If I am very diligent, I could record your answer on a sheet of paper. To write down what you told me, I could use a shorthand notation that is quite common in economics. I could write down:

$$\text{stir fried vegetables with rice} \succ \text{beef burger with fries.}$$

The symbol "$\succ$," that looks a little bit like a mathematical "greater" sign ($>$), but is curved, is read as: "is strictly preferred." So what I have written down means: stir fried vegetables with rice are strictly preferred to a beef burger with fries.

To figure out your complete preferences, I could go through all pairs of items on the menu, and for each of them ask you which one you prefer, and then write down the answer. Sometimes, you might tell me that you don't really care. We call that in economics "indifference." We would write, for example:

$$\text{roasted cauliflower} \sim \text{baked potato.}$$

The symbol "$\sim$," that looks a little bit like the mathematical "equal" sign ($=$). The statement that I have written down is read as "the decision maker is indifferent between roasted cauliflower and baked potato.

Let X be the set of items on the dinner menu, where we refer to the individual item shorthand as: "$x, y, z, \ldots$" I can now say what we mean by a "preference." A preference is a *complete record* of the decision maker's pairwise comparisons, that is, it is a list in which, for every pair x, y of elements of X (where $x \neq y$) exactly one of the following three statements appears:

- $x \succ y$
- $y \succ x$
- $x \sim y$

For example, when $X = \{x, y, z\}$, then one preference would be the following list:

$$x \succ y, z \succ y, x \sim z.$$

Not every preference makes equal sense. For example, this preference seems contradictory:

$$x \succ y, y \succ z, z \succ x.$$

This preference has a built-in cycle. The decision maker prefers x to y, and also y to z, but she prefers z also to x. As we go around in a circle: $x \rightarrow y \rightarrow z \rightarrow x \rightarrow y \ldots$, things get worse and worse for the decision maker, with no end in sight. We want to rule out such preferences. We shall assume that preferences are transitive:

$$\text{Whenever } x \succ y \text{ and } y \succ z, \text{ then we also have: } x \succ z.$$

This is the only assumption that we shall make. Aside from this, peoples' preferences reflect their personal taste, and we don't judge personal tastes.

"De gustibus non est disputandum," as one says in Latin. An English explanation of this phrase can be found on Wikipedia.

Rational Choices

A rational agent chooses among any set of available choices the one that is best according to her preferences. So, suppose the set of available choices were some subset of the set X, perhaps because not all ingredients are available, and therefore the chef has had to remove some dishes from the menu. Let's call that subset Y. Suppose:

$$Y = \{y_1, y_2, \ldots, y_n.\}$$

The rational economic agent will choose among these options the one that maximizes her preferences, that is, she will choose a y_i such that:

$$y_i \succ y$$

for all other choices y in Y. If the agent is sometimes indifferent between different menu items, then this condition should more precisely say:

$$y_i \succ y \text{ or } y_i \sim y$$

for all other choices y in Y.

So far, it seems as if we were heading for this definition of rationality: the rational economic agent has transitive preferences, and chooses from the available choices one that is best according to her preferences. But that is still not exactly the definition that we shall use. Our actual definition of rationality is this:

The rational economic agent behaves as if she had transitive preferences, and as if she chose from the available alternatives one that is best according to these preferences.

Thus, it does not matter for rationality whether an agent is aware of her preferences or not, and whether she deliberately chooses optimally, or perhaps just by habit. When I interview you about your preference over the dishes on the menu, you might have no answers, or incoherent answers. But it might still be that I find, as I observe you making choices over many menus, that I can infer a preference of yours, even if you are not aware of it. You make your dinner choices as if you had a preference. For example, I could discover that you always choose the dish with the largest amount of carbohydrates. You might not have been aware of this. But the economist would call your choice "rational."

In this sense, even animals can make rational choices. Indeed, they might often do so, because, if their behavior is shaped by biological evolution, they will make choices that maximize their number of offsprings, so that it is as if they preferred more reproductive fitness over less, even though, of course, the animals might not know this, nor make deliberate choices. And if we humans are just another kind of animal, then maybe even we, without knowing, make choices that maximize what is called our "reproductive fitness," i.e. our expected number of offsprings? Certainly, without knowing it, we might act *as if* we were pursuing one or the other goal. In economics, if that goal is a transitive preference, we call any such behavior "rational." Thus, the "rationality" assumption is a much weaker assumption than critiques of the rationality assumption in economics usually think it is.

Utility

You may have heard, perhaps in an introductory economics class, that economics studies the "utility maximizing" agent. So far, we have not talked about utility. Again, this is a concept that is widely misunderstood, even by students who have taken an introductory economics class. Let me try to explain it clearly. A "preference," as I defined it earlier, is a complicated object. For example, if there are 100 possible items that you might be able to choose from, then I have to write for every pair of items which one you prefer, or that you are indifferent. There are 4950 such pairs. My list will become very long!

An easier way of writing down your preferences is to ask you to assign to each of the 100 items a number so that the number is higher the more you prefer an item. Thus, if you gave item a the number 10, and item b the number 5, then I can conclude that $a \succ b$. And if you gave item c the number 7, then I can conclude that $a \succ c$, but also that $c \succ b$. Thus, assigning numbers to items is a short way of explaining preferences. Formally, an assignment of numbers to items is a function that maps the set of potentially available alternatives X into real numbers. In mathematics, we write such a function as:

$$u : X \to \mathbb{R}.$$

I have denoted the function by u. The number assigned to some alternative, say a, is written as $u(a)$, and $u(a)$ is called the utility of a. The function u is called the utility function of the agent whom we are considering. But note that the fact that the utility of a is, say, 10, by itself does not tell you anything. All that utility does is that it encodes the comparison between alternatives. If I tell you that $u(a) = 10$, and $u(b) = 8$, I have told you that: $a \succ b$, and if I told you instead that $u(a) = 3$, and that $u(b) = 1$, I would have told you just the same, namely that $a \succ b$. The numbers themselves are meaningless. Only their comparison conveys information.

Utility functions are so easily, and so frequently, misunderstood that it is worth devoting some more paragraphs to their interpretation. Sometimes, the utility numbers are interpreted as "utils" by economics teachers. This is just a metaphor. The metaphor may mislead people. Economics nowhere assumes the existence of something called "utils" that we can measure.

Suppose someone's utility from their morning coffee is -1. Does this mean they don't enjoy their coffee? No! It doesn't. The number by itself says nothing. If their utility from a morning tea would be -2, then you would have some information, namely that they like coffee better than tea.

Suppose someone's utility from a plain bagel is 1, and their utility from an everything bagel is 2. Does this mean that they like the everything bagel twice as much as they like the plain bagel? No! it just means that they like the everything bagel more than the plain bagel. Only the comparison, but not the ratio, is meaningful.

Suppose someone assigns utility 1 to one bagel, utility 1.5 to two bagels, and utility 1.8 to three bagels. Does that mean that they have decreasing utility from every additional bagel? No! that statement actually makes no sense by itself. All that the utility function tells you is that the person prefers three bagels over two, and that the person prefers two bagels over one.

Suppose someone's utility from a beautiful apple is 3. Does that mean that they are willing to pay 3 Dollars for the apple? No! Utility has no units. In particular, it is not expressed in Dollars.

I could go on and on. But I shouldn't. I can imagine the impatience that is already filling you, as I have gone on about this topic for so long. So, just remember, utility is used to describe preferences, nothing more. We also say that utility is ordinal in economics.

On the other hand, I feel that I myself am getting all worked up ... That is not good, either.

Revealed Preferences

The hypothesis of economics is that people choose *as if* they had preferences, not that they actually do. So, when economists are acting as proper empirical scientists, they try to infer people's preferences from people's choices. They don't just assume what peoples' preferences are. For example, we don't assume that everyone is greedy in the sense of liking more money rather than less. Maybe our behavior shows that we are greedy, or maybe it doesn't. Thus, the critique that economists think that people are selfish and money hungry is based on a misunderstanding of economics.

Maybe I think, and say, that I am not greedy, but if you carefully observe my behavior, you conclude that I'm greedy without knowing it even myself.

So, suppose I try to infer your preferences from your choices. Suppose I saw that the table with free fruits had apples and oranges on it, and you chose an apple. Then I can infer, from your behavior, that your preference includes the statement that you prefer an apple over an orange. Making many more observations about which foods you pick when you face a buffet, I can infer more things about your food preferences. Maybe I conclude that you prefer foods that are rich in fat over other foods. You may say "no," but if I can bring you evidence, it may persuade you. In other words: your behavior may be a better guide to your preferences than your introspection. This is one more reason for economists to study the preferences that you reveal through your choices rather than the preferences that you think you have.

We have just now compared apples to oranges, by the way. Some people say that is impossible. But I may say, for myself, that I have no problem comparing them, and that I like the taste of oranges much better than that of apples.

It may be that I don't observe enough choices to infer your preferences completely, but that I can make just some incomplete inferences about your preferences. That is still useful. Economists use inferred preferences to try and predict you behavior in choice situations that, perhaps, you haven't yet confronted.

It is also possible that I observe behavior of an economic agent that is not compatible with *any* preferences. In other words, the assumption of rational choice behavior is not without empirical content. For example, if you choose a from the choice set $\{a, b\}$, but b from the choice set $\{a, b, c\}$, then, assuming that you have strict preferences, your first choice shows that your preference is: $a \succ b$, but your second choice reveals: $b \succ a$, and also $b \succ c$. But now we have a contradiction: which one is true: $a \succ b$, or $b \succ a$? This is the sort of observation that can refute the economists assumption of rational choice. And, indeed, it is not too rare that this is the sort of choices that people make.

Equilibrium

The second key idea of economic theory is that of "equilibrium." I shall say much less at this point about equilibrium than I have said about rationality. I postpone this discussion mostly to a later point. But let me give a little preview. You have perhaps encountered one notion of equilibrium already, where equilibrium is defined as an outcome where supply equals demand. The "law of supply and demand" is perhaps what economists are most famous for. I show the standard supply and demand diagram in Figure 2.

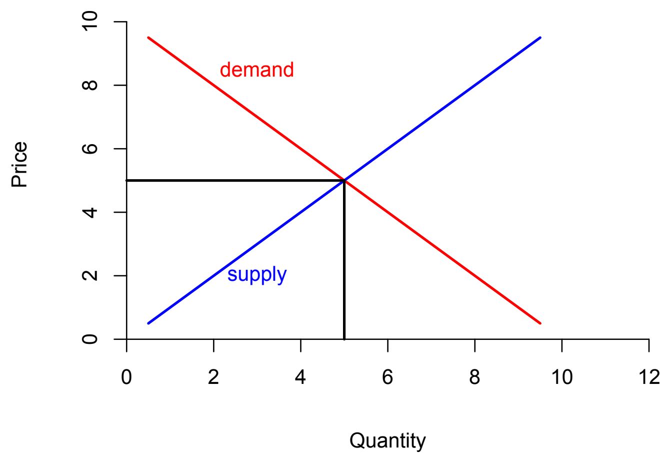
Or, maybe, economists are infamous for boring people with the "law of supply and demand."

This diagram corresponds to one market. On the vertical axis we have the price prevailing in that market. On the horizontal axis we have the quantity traded in the market. The blue upward-sloping line is the graph of the supply function. It indicates for every price the quantity that sellers want to sell. The red, downward sloping line is the graph of the demand function. It indicates for every price the quantity that buyers want to buy. The equilibrium is where these two intersect. There is thus an equilibrium price, marked by a horizontal black, line in Figure 2, and an equilibrium quantity, marked by a vertical line in Figure 2.

Note that the independent variable of the function, namely the price, is on the *vertical* axis, and the dependent variable, namely the quantity, is on the *horizontal* axis. This is the opposite of how one draws the graph of functions in mathematics. For some reason, the convention that I show in Figure 2 has persisted in economics over many decades.

I want to make two points about the law of supply and demand. First, it is a prediction. That is, we predict that the transactions in the market will occur at the equilibrium price, and the total quantity traded will be the equilibrium

Figure 2: The Supply and Demand Diagram



quantity. This may, or may not, be true. We'll discuss later the evidence.

Second, the motivation for this prediction is often described as follows: "if the price is above the equilibrium price, there will be more supply than demand, which will make the price fall. If the price is below the equilibrium price, then the price will rise. Thus, it is stable only at its equilibrium value." The point that I want you to notice is that this is a dynamic story, a story of adjustment over time. Thus, "equilibrium" is meant to describe the rest point of an adjustment process. This suggests that it will most likely come about if there is enough time for adjustment, and if the environment does not change too much during this time.

These two points are meant to help you interpret the notion of equilibrium in economics. They also apply to other notions of equilibrium than the equilibrium of supply and demand. For example, in game theory, the concept of Nash equilibrium has a similar interpretation. We shall learn about game theory and Nash equilibrium later in these notes.

The interpretation of the concept of "equilibrium" that I have suggested here makes immediately clear that not everything that we observe everywhere can be expected to be an equilibrium. Sometimes we observe the adjustment process, rather than its restpoint, and sometimes there is actually not enough time to reach the restpoint. Equilibrium is not a *universally* plausible concept. Depending on the facts, it may make sense in some environment, yet no sense in another.

Topic 3: Budget Sets, Utility Functions, and Indifference Curves

We shall now apply the model of rational choice to some economic choices. We begin with the choice of what to consume. In economics, we mean by "consumption" a larger variety of activities than we mean by "consumption" in every day life. In economics, when you eat an apple, you consume that apple, but also, when you wear a shirt, you consume that shirt, and when you have a haircut, you consume the service of the hairdresser. When we study rationality in consumption, we have in mind all these applications.

We could apply the rational choice model to many other decisions that people make, such as the choice whether to work, how much to work, in which job to work; or to the choice whether and how much to save; or to the choice whether to have children, and how many. We shall in fact study rationality in some of these choices later. But for the moment we focus on consumption.

Consumption might seem to you a very unlikely candidate for a successful application of the theory of rational choice. My own introspection suggests that I certainly don't give a lot of thought to my choices about every day consumption, such as dinner purchases, or clothes purchases. Indeed, some of them are driven by what you may call addiction. For example, if I didn't have coffee every day, I would get a bad headache. So, my coffee consumption is perhaps addiction driven. But remember from the previous topic that by rationality we don't mean that people make carefully considered choices, we only mean that they act as if they maximized some consistent preferences. Perhaps, we don't think that our consumption is as if it maximized some well-defined preference, but to an outside observer it appears so. So, let's keep an open mind about the applicability of the rational choice model to consumption. Ultimately, data have to decide whether it is a good model. We shall talk a little about consumption data later in this course.

In fact, some of the ideas that we shall develop when studying rational choice in consumption are central to most peoples' understanding of economics. Many economists believe that an important part of "thinking like an economist" is to think in terms of the concepts that I am about to teach you. So, it is worth our time to examine these concepts.

Budget Sets

In the previous section we studied choices among some small sets of alternatives, for example, choices among the items that are on the menu of a restaurant. In this section, we shall think of a very large set of alternatives. We shall try to cover all consumption choices, that is, we are trying to address how much someone consumes of every good. For each good, you choose how much you consume of it. Let's introduce some notation. For example, q_{milk} might denote the quantity of milk that you consume, measured in some units. For each consumption good you choose such a quantity. We want to study how a rational consumer would choose all these quantities at the same time, because a rational consumer, when choosing whether she wants to consume more of one good, will weight all alternatives on which the same money could be spent. As a consequence, the choice variable of our rational consumer will be a very long list of quantities, one quantity, that is, one number, for each good.

I promised that we would consider choice among all consumption goods, but, actually, I also want to draw diagrams, and I can only draw two-dimensional diagrams. Therefore, I shall pretend that the complete set of goods that the consumer has to consider consists actually only of two goods, which we shall imaginatively name good 1 and good 2. But what I will tell you is largely also true if there are a million goods rather than just two.

Maybe some readers will remember that Dr. Seuss wrote in a children's book about "Thing one and thing two."

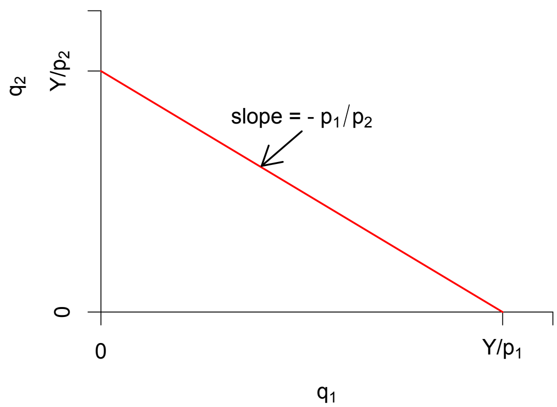
So we are going to write q_1 for the quantity of good 1 that the consumer consumes, and q_2 for the quantity of good 2 that the consumer consumes. We shall make the very idealized assumption that both quantities can take any non-negative value: $q_1 \geq 0$ and $q_2 \geq 0$.

Which choices are feasible will depend on the consumer's budget. The consumer has some given amount of money y, which we shall now refer to as "income." Let's assume that $y > 0$. We won't ask - yet - how income is determined. Each of the two goods has prices: $p_1 > 0$ and $p_2 > 0$. We shall assume that the consumer takes these prices as given, and does not negotiate over them. Sometimes in economics we call this price taking behavior. A price taking consumer will, for example, not bargain over prices. The consumer can only choose consumption quantities q_1 and q_2 that satisfy:

$$p_1 q_1 + p_2 q_2 \leq y.$$

Here, $p_1 q_1$ denotes the amount of money that the consumer spends on good 1 when buying q_1 units of good 1, and $p_2 q_2$ is the same for good 2. The inequality says that the total amount spent must not be more than the consumer has income. We call this inequality the budget constraint.

Figure 3: Budget Line



Now I want to draw in a diagram, with q_1 on the horizontal axis, and q_2 on the vertical axis, the set of all quantity pairs (q_1, q_2) that the consumer can choose. We call such pairs also consumption bundles. Let's first look at those bundles that satisfy the budget constraint as an equality:

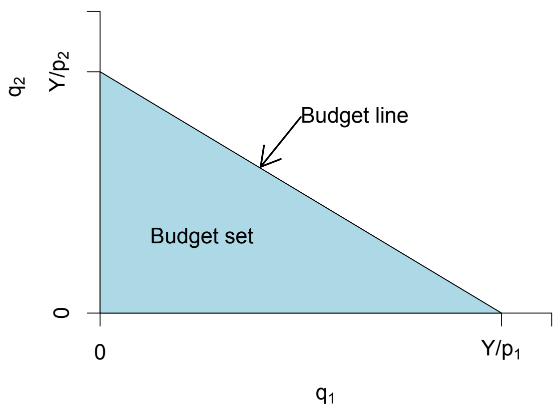
$$p_1 q_1 + p_2 q_2 = y.$$

To draw all these bundles in a diagram, it is best to solve the inequality for q_2 because, in our diagram, q_2 will be on the vertical axis. We have:

$$\begin{aligned} p_1 q_1 + p_2 q_2 &= y \Leftrightarrow \\ p_2 q_2 &= y - p_1 q_1 \Leftrightarrow \\ q_2 &= \frac{y}{p_2} - \frac{p_1}{p_2} q_1. \end{aligned}$$

Notice that if y, p_1 and p_2 are given constants, q_2 is just a linear function of q_1. The graph of all consumption bundles which satisfy the budget constraint as an equality is therefore a straight line. The intercept on the vertical axis is $\frac{y}{p_2}$, and the slope is $-\frac{p_1}{p_2}$. In Figure 1, I have drawn that line. We shall call these lines sometimes the budget line.

Figure 4: Budget Set



Now, of course, not only consumption bundles on that line, but also consumption bundles below that line are feasible. If the consumer picks such consumption bundles, she does not spend all her income. For the moment, we shall allow for that possibility. It will not play a big role. But why not, just for a second, consider the very frugal consumer? In Figure 2, I show all consumption bundles that are on, or below, the budget line. The set that I have shaded in blue in this figure is also called the budget set.

To summarize, we shall consider in this section consumers who make rational choices from sets of available choices that have the triangular form shown in the above figure.

Note that in both figures the intercepts of the budget line with the axes of the coordinate system are at $\frac{y}{p_1}$ and $\frac{y}{p_2}$. The former intercept corresponds to the case that the consumer spends all her income on good 1, and nothing on good 2, thus buying $\frac{y}{p_1}$ units of good 1, and the second intercept corresponds to the opposite case.

One final point on the subject of budgets: there is a simple intuition for why the slope of the edge of the budget set, that is, of the budget line, is equal to $-\frac{p_1}{p_2}$. Recall from calculus that the slope is the answer to the question how much of good 2 you have to give up if you want to buy one more unit of good 1, and satisfy the budget constraint. If you want to buy one more unit of good 1, you need to have p_1 Dollars. To find out how many units of good 2 this corresponds to, you just have to divide this amount by p_2, and thus you get: $\frac{p_1}{p_2}$. Because you are giving up units of consumption, a minus sign goes in front. And that is the slope of the budget line.

Utility Functions and Indifference Curves

After describing the type of sets that the consumer can choose from, next we have to describe the consumer's preferences. If we followed very strictly the logic of rational choice outlined in the previous topic, we should consider first the consumer's preference $\succ$ over consumption bundles, and only later introduce a utility function to represent that preference. But we shall right away use a utility function to describe the preferences. That is convenient, and conventional. But always keep in mind that the interpretation of utility functions is ordinal, i.e., only utility comparisons matter. The utility numbers themselves have no meaning.

We want to describe preferences over any consumption bundles (q_1, q_2) that the consumer might some time be able to choose from, and therefore we will not restrict attention to any particular budget set. Rather, we shall consider any such bundles where both quantities are positive. Mathematically speaking, the domain of our utility function will be $\mathbb{R}^2_+$. This is just mathematical notation for the set of all pairs of non-negative real numbers.

The range of the utility function will just be the set of real numbers, $\mathbb{R}$. Thus, the utility function is of the form $U(q_1, q_2)$, assigning to every consumption bundle some utility. Once again, all that matters are comparisons. Using the notation of the previous topic, $U(q_1, q_2) > U(\hat{q}_1, \hat{q}_2)$ just means $(q_1, q_2) \succ (\hat{q}_1, \hat{q}_2)$ where $(\hat{q}_1, \hat{q}_2)$ is just some consumption bundle other than (q_1, q_2).

If we wanted to draw the graph of a utility function, with two goods, then we would have to draw a 3-dimensional graph. We shall make our task a little easier. We shall only draw a 2-dimensional graph. Our graph will be analogous to a map. When you look at a map, you will see contour lines that indicate all locations on the map that are, for example, 800 feet high. Contour lines are very useful for a hiker. If you see very contour lines very close together on a map, then you know that in that area your hike is going to be very steep. A map is a 2-dimensional representation of a 3-dimensional reality.

For utility functions, in the case of two goods, we can depict the location of all consumption bundles between which give the consumer the same utility, that is, between which she is indifferent. Here, "same utility" is analogous to "same elevation" in maps.

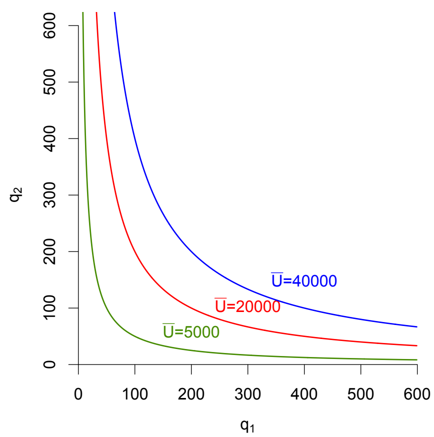


Figure 5: Indifference Curves

Consider a simple example. Suppose the utility function is:

$$U(q_1, q_1) = q_1 \cdot q_2$$

This is just an example, though an example that we shall use frequently. It is convenient. There is no reason why it would be plausible. Let's find the pairs (q_1, q_2) for which utility level is 10. These are the pairs which satisfy:

$$q_1 \cdot q_2 = 10.$$

We want to draw the graph of these pairs. We solve for the variable that is on the vertical axis, i.e. q_2:

$$q_2 = \frac{10}{q_1}.$$

Had we started with any other utility level $\bar{U}$, we would have obtained:

$$q_2 = \frac{\bar{U}}{q_1}.$$

For any given value of $\bar{U}$, we can draw the graph of this line in a 2-dimensional diagram. We call these graphs indifference curves, because they indicate all consumption bundles among whom the consumer is indifferent. In Figure 3, I have indicated for the above example a couple of indifference curves. They are all hyperbolas.

Really, the non-negative orthant is filled with indifference curves. If I were to draw all of them, however, then the non-negative orthant would just be one red area, and no individual curve could be distinguished from any other. So, I have just drawn a few examples.

Note that as consumption bundles that are on indifference curves further away from the origin of the orthant are strictly preferred to consumption bundles that are on indifference curves closer to the origin. By giving you this information, as well as the graph of the indifference curves, I have have, in fact, described for you the complete ordinal preferences of the consumer. The particular numbers that the utility function provides, convey, as we have often discussed, no further information. Thus, in principle, instead of writing down the utility function, I could just have drawn the indifference curves, and I could have told you the direction into which preferences increase. This would have indicated for you all that you need to know about the consumer's preferences. But it is, of course, much more convenient to just write down a utility function.

We could say that, unlike maps, systems of indifference curves, together with the direction of increase in preference, are a 2-dimensional representation of a 2-dimensional reality.

Let's do one more example. let's suppose utility is given by:

$$U(q_1, q_2) = q_1 + q_2.$$

What are indifference curves for this utility function? Well, let us look at some utility level $\bar{U}$, and find all consumption bundles that satisfy:

$$q_1 + q_2 = \bar{U}.$$

Solving for q_2, we get:

$$q_2 = \bar{U} - q_1,$$

and this shows that the indifference curves are straight lines with intercept $\bar{U}$

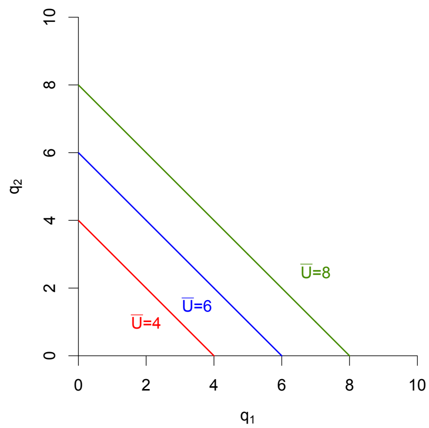


Figure 6: Another Example of Indifference Curves

on both axes, and with slope -1. Figure 4 shows some of these indifference curves. They look like budget lines, but they are not. They are indifference curves.

Monotone Preferences

Note that in the two figures showing indifference curves, we see that these curves are downward sloping. This will actually always be the case if the preferences represented by the utility function are monotone (actually: monotonically increasing, but we will never consider the case of monotonically decreasing utility functions), which means, that having more of any good is better. It is actually clearer to translate the property of utility functions being monotone into a corresponding property of preferences. Preferences are monotone if

$$(q_1, q_2) \succ (\hat{q}_1, \hat{q}_2)$$

whenever $q_1 \geq \hat{q}_1$ and $q_2 \geq \hat{q}_2$, and one of the inequalities is strict. Saying that the utility function is monotonically increasing is just the same as saying that preferences are monotone.

To see that for monotone preferences indifference curves are downward sloping, just note that along an indifference curves, as we move from the left to the right, we get more of good 1, i.e. higher q_1. A consumer with monotone preferences can remain indifferent then only if q_2 is decreasing. Otherwise, by monotonicity, he would be better off. Thus, monotonicity of preferences implies that all indifference curves are downward sloping, as is the case in both of our examples. In both examples you can also verify that the utility functions are monotone.

Marginal Rates of Substitution

In the previous section we explained that for monotone preferences, indifference curves are downward sloping, which means that the derivative of the equation that shows q_2 as a function of q_1 along any given indifference curve is negative. In other words:

$$\frac{dq_2}{dq_1} < 0.$$

In economics, we call the absolute value of the derivative $\frac{dq_2}{dq_1}$ the Marginal Rate of Substitution between goods 1 and 2. That is, the Marginal Rate of Substitution is the derivative $\frac{dq_2}{dq_1}$, but dropping the minus sign in front of the derivative (or, equivalently, multiplying the derivative by -1). It answers the following question: If we give the consumer one more unit of good 1, how many units of good 2 can he give up to be exactly indifferent between what he had before and what he has afterwards? Thus, it describes how the consumer substitutes good 1 for good 2. We abbreviate the term "Marginal Rate of Substitution" by writing MRS. The word "marginal" indicates that we consider the case that units are very, very small. More precisely, we are considering limits for unit size tending to zero. We do so automatically whenever we use the notion of a derivative. The word "marginal" simply indicates that derivatives are involved.

Note that we take the negative of $\frac{dq_2}{dq_1}$. This is because $\frac{dq_2}{dq_1}$ is itself a negative number: this indicates that we have to give up of good 2 if we gain more of good 1. But it is easier to calculate with positive numbers. We put a minus sign in front of $\frac{dq_2}{dq_1}$, and therefore get a positive number.

Let's calculate for the first of our previous example the MRS. The indifference curves in that example were given by:

$$q_2 = \frac{\bar{U}}{q_1}$$

The MRS is just the absolute value of the derivative of the right hand side.

Therefore, we get:

$$\text{MRS} = \frac{\bar{U}}{(q_1)^2}$$

(To see this, note that $q_2 = \frac{\bar{U}}{q_1}$ is the same as: $q_2 = \bar{U} \cdot (q_1)^{-1}$, apply the power rule of differentiation, and multiply by -1.) Note that the MRS that we calculated is positive. The slope of the indifference curve is the negative of this. The indifference curve is downward sloping.

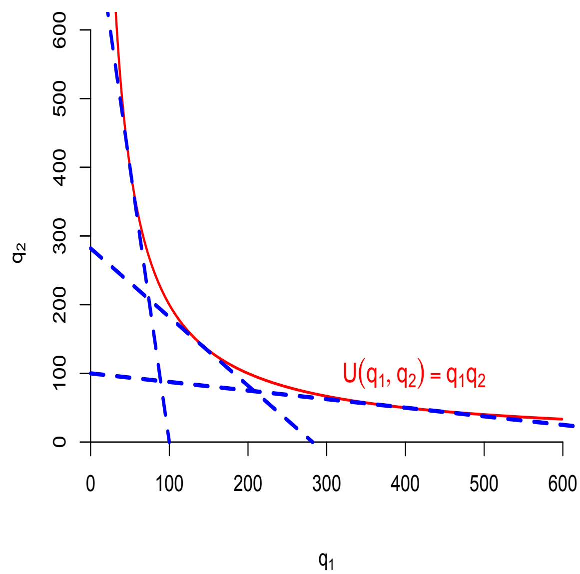


Figure 7: Marginal Rates of Substitution

There is a simple formula for calculating MRS that we will find useful, also because it gives us some intuition for the meaning of the MRS. I shall not derive the formula here, but I shall simple write it down, and then explain its meaning and some intuition. The formula is:

$$\text{MRS} = \frac{\frac{\partial U}{\partial q_1}}{\frac{\partial U}{\partial q_2}}.$$

We need to interpret the symbols on the right hand side. They are partial derivatives. This concept might frighten you, but there is nothing special about it. The symbol

$$\frac{\partial U}{\partial q_1}$$

for example, is read as the partial derivative of U with respect to q_1, which

There is really nothing special about partial derivatives, except that the notation is curious. Suppose I asked you to differentiate $f(x) = ax^2$. Presumably, you would quickly come up with the derivative: $f'(x) = 2ax$. (Aside: we use the two notations for the derivative: $\frac{df}{dx}$ and $f'(x)$ interchangeably.) But, strictly speaking, you have calculated here a partial derivative. This is because there are really two independent variables, a and x, but you have treated one of them, a, as a constant. Thus, you have really calculated the partial derivative of f with respect to x. There is nothing more to partial derivatives than this. No need to take a class in multi-dimensional calculus. What is difficult in such classes is not the concept of partial derivatives, but other things, which don't matter for our course.

means that you take the derivative of U with respect to q_1, pretending that the other variable, q_2, where a constant. So, for example, if

$$U(q_1, q_2) = q_1 \cdot q_2,$$

then, if we treat q_2 as a constant, this function is a constant times q_1, and therefore the partial derivative with respect to q_1 is just that constant:

$$\frac{\partial U}{\partial q_1} = q_2.$$

Similarly:

$$\frac{\partial U}{\partial q_2} = q_1.$$

Therefore, using our formula for the MRS we get in this example:

$$\text{MRS} = \frac{\frac{\partial U}{\partial q_1}}{\frac{\partial U}{\partial q_2}} = -\frac{q_2}{q_1}.$$

But, you might say, earlier we found:

$$\text{MRS} = \frac{\bar{U}}{(q_1)^2}.$$

Why are these two formulas different from each other? Well, suppose we look at any particular consumption bundle (q_1, q_2). The MRS may be different, of course, depending on which consumption bundle we start out with. But if we start out with (q_1, q_2), the utility level that we are looking at, i.e. $\bar{U}$, is:

$$\bar{U} = q_1 \cdot q_2$$

Plugging that in, our second formula for the MRS is:

$$\text{MRS} = \frac{\bar{U}}{(q_1)^2} = \frac{q_1 q_2}{(q_1)^2},$$

which, after we cancel out q_1, simplifies to:

$$\text{MRS} = \frac{q_2}{q_1}$$

which is the same as we got earlier.

Using the concept of MRS we can confirm what we observed in the previous subsection. If preferences are monotone, then we have to have:

$$\frac{\partial U}{\partial q_1} > 0 \quad \text{and} \quad \frac{\partial U}{\partial q_2} > 0,$$

and therefore the MRS, which is the ratio of these two is positive, and the slope of the indifference curves, which is the MRS with a minus sign in front, is negative, and hence the indifference curves are downward sloping.

Why is the MRS equal to the ratio of the partial derivatives? Well, if we increase q_1 by a small unit, we get $\frac{\partial U}{\partial q_1}$ more utility. If we increase q_2 by a small unit, we get $\frac{\partial U}{\partial q_2}$ more utility for each unit. Thus, to compensate for $\frac{\partial U}{\partial q_1}$ more utility, we have to reduce the consumption of good 2 by "$\frac{\partial U}{\partial q_1}$ divided by $\frac{\partial U}{\partial q_2}$" units of good 2, and that is the MRS. If one extra unit of q_1 increases utility by 3, and one extra unit of good 2 increases utility by 6, then, if we get one more unit of good 1, utility is kept constant by giving up half a unit of good 2, that is, 3/6.

Convex Preferences

The indifference curves in our first example above are not only decreasing, they are also convex. What does it mean that an indifference curve is convex? Informally, it means that, as you drive along the indifference curve with a car, you have to turn the steering wheel to the left. Mathematically speaking, the indifference curves get flatter and flatter. You can see this in Figure 5, where I have drawn some tangents to the indifference curve, that is, straight lines that have the same slope as the indifference curve has in the point of tangency. That the indifference curve is convex is reflected by the fact that, the further we move to the right, the flatter the tangents become.

I owe this tangible explanation of convexity to my late mathematics high school teacher, Herr Lehmann. (We never used his first name.)

What does convexity mean in economic terms? It means that the amount of good 2 that the consumer is willing to give up for one more unit of good 1 is getting smaller as the consumer has more and more of good 1. Very loosely speaking, the marginal utility of good 1, expressed in units of good 2, is diminishing: the more of good 1 the consumer has, the less is an additional unit of good 1 worth to her.

Another way of saying this is to say that the MRS is decreasing as we move along an indifference curve. For the first example above we have already calculated the MRS. It was given by:

$$\text{MRS} = \frac{q_2}{q_1},$$

As we move an indifference curve, two things happen: q_1 goes up, and q_2 goes down. What happens to the ratio $fracq_2q_1$? Well, the denominator increases, and the numerator decreases. Thus, the ratio decreases, and we see again that this is a case of convex preferences.

In our second example, they are straight lines. We might say that they are weakly convex. Strictly speaking, they are not convex, but they are also not the opposite of convex, i.e. they are not concave.

Of course, any particular consumers preferences need not be convex. There is nothing more or less rational about having convex preferences. It just happens to be mathematically nice if preferences are convex. So, it is an assumption that we shall frequently make, but we don't argue it is particularly realistic.

Another way of defining convex preferences is to say that the consumer, whenever she is indifferent between two consumption bundles, strictly prefers a mixture of these consumption bundles over the original bundles. Let's write that in utility terms. Suppose there are two consumption bundles, q_1, q_2 and $(\hat{q}_1, \hat{q}_2)$ such that $U(q_1, q_2) = U(\hat{q}_1, \hat{q}_2)$, i.e. the consumer is indifferent. What is a mixture of these bundles? Well, the exactly equally weighted mixture is:

$$\left(\frac{1}{2}q_1 + \frac{1}{2}\hat{q}_1, \frac{1}{2}q_2 + \frac{1}{2}\hat{q}_2\right).$$

What is this: it is a pair of two quantities: of good 1 the consumer gets the exact average of q_1 and $\hat{q}_1$, and of good 2 the consumer gets the exact average of q_2 and $\hat{q}_2$. If one bundle has 3 apples and 6 pairs, and the other one has 7 apples and 2 pears, then the mixture is (3+7)/2=5 apples, and (6+2)/2=4 pears.

Having defined mixtures, we can now say what convexity means: Preferences are convex if for any two consumption bundles q_1, q_2 and $(\hat{q}_1, \hat{q}_2)$ for which $U(q_1, q_2) = U(\hat{q}_1, \hat{q}_2)$ we have:

$$\left(\frac{1}{2}q_1 + \frac{1}{2}\hat{q}_1, \frac{1}{2}q_2 + \frac{1}{2}\hat{q}_2\right) > U(q_1, q_2) = U(\hat{q}_1, \hat{q}_2).$$

That is, if a consumer is indifferent between having 3 apples and 6 pairs, and having 7 apples and 2 pears, then she strictly prefers 5 apples and 4 pears over either of those bundles.

You may be surprised. Didn't we earlier define convexity as preferences with decreasing marginal rates of substitution. Why another definition? How does it relate to the earlier definition. The answer is: the two definitions are equivalent. It is not obvious, but we shall not prove it/explain it here. You should know this fact, though, even if you don't know exactly why it is true.

Does this make sense? Do you have convex preferences? If preferences are about what we like, then introspection suggests that this sometimes makes sense, and sometimes not. If you have to choose between spending all your money on food and nothing on clothes, or spending all your money on clothes, and nothing on food, you might say that you are indifferent. Both are equally bad. But spending half of your money on clothes, and half on food, might sound better. Then your preferences are convex. But do you prefer half a cup of tea plus half a cup of coffee over one cup of tea, one cup of coffee, if you are indifferent between coffee and tea? I doubt it.

This example demonstrates that convexity is perhaps more plausible when goods are defined broadly, such as "food" rather than narrowly, such as "tea."

Of course, really, as explained in Topic 2, we mean in economics by preferences not necessarily what our introspection tells us makes us happy, but we mean the preferences revealed by our choices. A very mysterious result, namely Afriat's Theorem, says that a finite data set can never refute the hypothesis that preferences are convex. That is a subtle, counter-intuitive fact, that we shall unfortunately not have time in this course to explain. In this sense, assuming convex preferences is actually without loss of generality.

Because it was discovered by Sydney Afriat.

Our earlier definition of convexity had a clear geometric meaning: indifference curves get flatter as one moves to the right. Can we give our second definition of convexity a geometric interpretation? Yes, and at first sight, it is different from the earlier one. In our new definition, we start with two bundles between which the consumer is indifferent. That is, we pick two bundles that are on the same indifference curves. Then we consider the mixture. Geometrically, this mixture is exactly the mid point of the straight line connecting the original two bundles. Convexity means that this midpoint must be above the indifference curve with which we started. Figure 6 illustrates this.

Maybe it is geometrically plausible that the two properties of indifference curves that we have discussed, decreasing marginal rate of substitution, and midpoints of two points on the same curve being above the curve, are the

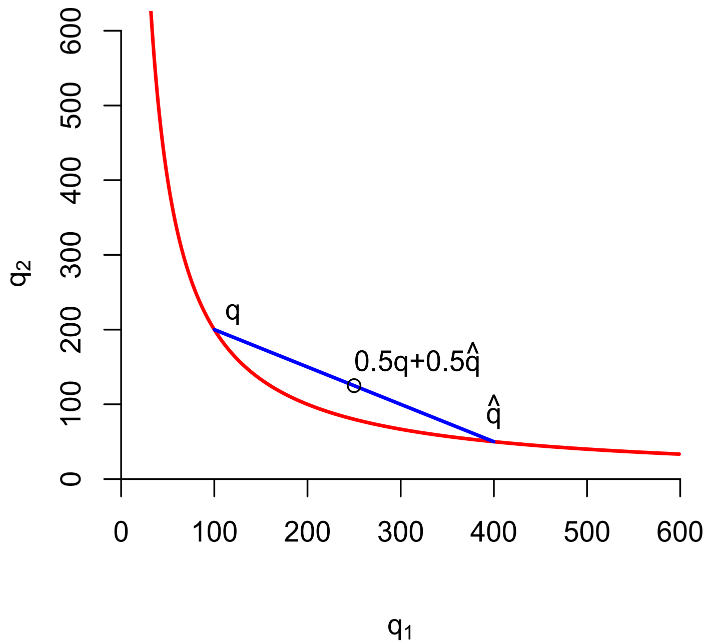


Figure 8: An Illustration of Convexity

same. But even if it is not plausible, it is true.

Is There A Law of Diminishing Marginal Utility?

You might have heard of a "law of diminishing marginal utility." It has a great tradition in economics. But note that in the framework that we have constructed in this section, at least its naive interpretation makes no sense. The law says: the more you consume of something, the less additional utility you derive from every extra unit. That sounds plausible: the more coffee I drink, the less helpful is one extra cup of coffee. But, in our setting this sense is meaningless: utility numbers are meaningless, and therefore, the phrase "additional utility" is meaningless. My advice to you: forget the naive law of diminishing marginal utility.

But, you might say, I myself spoke earlier in these notes about a law of diminishing marginal utility, namely in the first paragraph of the section on convexity. How can I extricate myself out of this contradiction? Well, this is what I wrote: "Very loosely speaking, the marginal utility of good 1, expressed in units of good 2, is diminishing: the more of good 1 the consumer has, the less is an additional unit of good 1 worth to her." What is "diminishing" is thus the quantity *expressed in units of good 2*, that an additional unit of good 1 is worth to the consumer. Expressed in units of good 2, not in units of imaginary utility.

Some people speak of "utils." Hopefully, you will never catch me use this word.

There is another point that you should notice about the above quote from the section on convexity. The quote occurred in the context of a discussion of convexity, and it referred to a movement *along an indifference curve*. That is, what we were asking was: as I give you more and more of good 1, at the same time compensating each additional unit of good 1 by a reduction in the quantity of good 2 that I give you, how does the quantity of good 2 that I can use to compensate you change. That is, I compensate you all along, so that at each point the starting point, that is, the quantity of good 2 that you currently hold, changes. This is different from keeping the quantity of the other good fixed. I illustrate the difference in Figure 7.

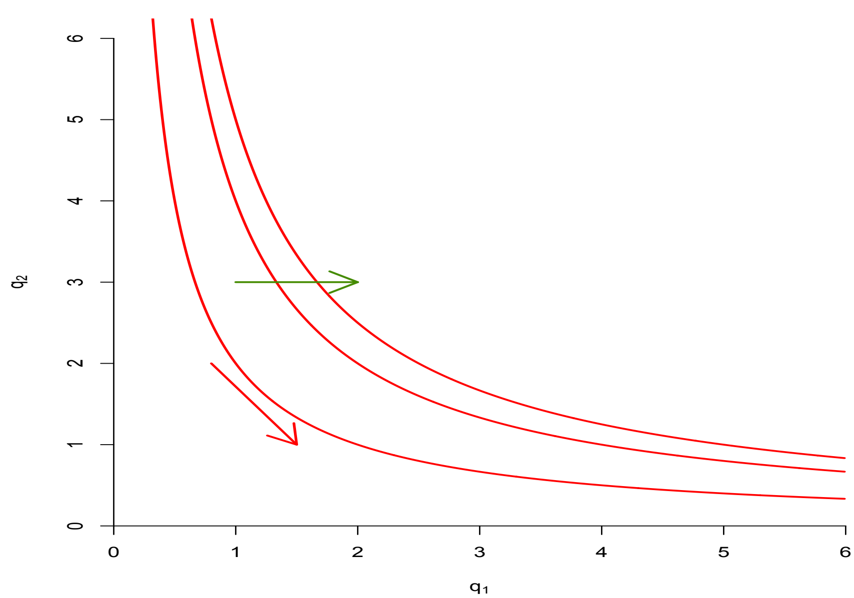


Figure 9: Directions Into Which the MRS Might Be Decreasing

The movement that we consider when we speak about diminishing marginal rates of substitution is along an indifference curve, indicated in Figure 7 by the red arrow. In other words, we ask ourselves how the MRS changes along an indifference curves. If we kept the quantity of good 2 fixed, then we would move in parallel to the horizontal axis, that is, along the green arrow. We would ask ourselves how the MRS changes as we move along the green arrow. That is a different question. It is meaningful in our setting, because it can be expressed in properties of indifference curves, and this green movement might be used to define a different law of diminishing marginal utility, so to speak. But here, in this section, we are only concerned with convexity, and so this alternative law of diminishing marginal utility is not of interest to us.

Topic 4: Utility Maximization

Now we shall put budget set and utility function together and begin to study the implications of the hypothesis of rational choice for consumption. In the framework of the last section, the rational choice now boils down to the solution to the following maximization problem: The consumer chooses $q_1 \geq 0$ and $q_2 \geq 0$ so as to maximize $U(q_1, q_2)$ where her choice is subject to the budget constraint: $p_1 q_1 + p_2 q_2 \leq y$. The consumer takes the prices as given, and we also take income as given, postponing until later the question how the income is acquired.

Before we proceed to studying this problem, let me remind you that the restriction to just two goods is mostly made only to make the presentation easier. Most of what I tell you also holds if there are million goods. Also, we shall use some calculus, but that does *not* mean that our consumer has to be good at calculus. She just acts as if she were able to solve the maximization problem. Whether real consumers do so, or not, is a question that we shall take up later.

Necessary Conditions for Utility Maximization

We shall begin by learning about two necessary conditions for maximizing utility. What is a necessary condition? A necessary condition is one that a consumption bundle *must* satisfy to be a candidate for a solution to the optimization problem. Satisfying the condition by itself does, however, not *guarantee* that a consumption bundle is a solution. Conditions which guarantee that we have a solution are called *sufficient* conditions, and we shall study them in the next section.

A *necessary* condition for being able to vote in US presidential elections is that one is a US citizen. But it is not *sufficient*. For example, nobody in prison is allowed to vote.

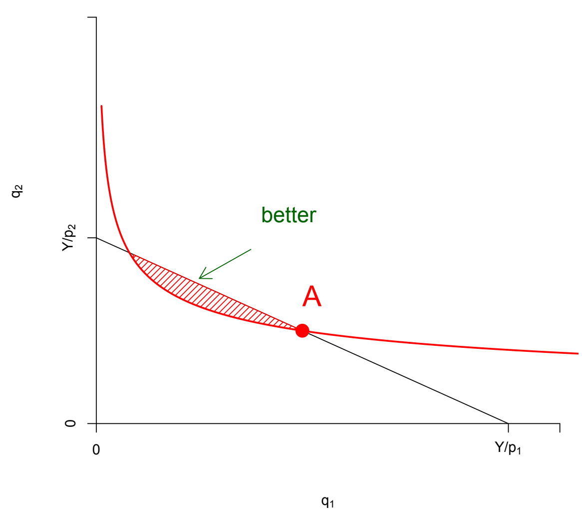
Let us assume for all of this section that preferences are monotone. Then the first necessary condition is that a utility maximizing consumption bundle must lie on the budget line, and not below it. This means in words that all money is spent, either on good 1 or on good 2. If it were not, then the consumer could buy more of one or both goods, and, if her preferences are monotone, she would have higher utility. This necessary condition is very obvious, except that you might wonder why it cannot be rational to save some of the income. This is because, at the moment, we are considering a very stylized model that is static, i.e., there is no future. If there is no future, then there is no point to saving. Later, we shall extend the model and also consider saving.

The second necessary condition is that the marginal rate of substitution must be equal to the price ratio:

$$\text{MRS} = \frac{p_1}{p_2}$$

Before we explain this condition, let us briefly pay attention to the details of this equation: the MRS on the left hand side is the the negative of the slope of the indifference curve, that is, it is: $-\frac{dq_1}{dq_2}$. Note that good 1 is in the numerator. The price ratio on the right hand side of our equation also has good 1 in the numerator. Thus, on both sides, good 1 is in the numerator. Of course, the choice of good 1 here is arbitrary. We could also have put good 2 into the numerator on both sides. We just have to put the same good on both sides into the numerator.

Figure 10: MRS < price ratio



Now, why should this have to hold in a utility maximum? I shall give two explanations, one in a graph, and one in formulas. First the geometric explanation. Recall that the MRS is the negative of the slope of the indifference curves, and that the price ratio is the negative of the slope of the budget line.

The condition therefore says that in an optimum these two slopes have to be the same. Let's consider in Figure 8 a case in which these slopes are *not* the same.

Suppose in Figure 8 the decision maker considered choosing consumption bundle A, which is marked by a red dot. It is on the budget line, thus satisfies our first necessary condition. But the second necessary condition is violated. I have drawn the indifference curve through A in the figure as a red line, and you can see that it is flatter than the budget line. Why is A not optimal? It is not optimal because there are choices with higher utility in the budget set: In fact, all consumption bundles that are in the red shaded set in Figure 8 are preferred over A because they are above the indifference curve through A. Moreover, they are also in the budget set, so the consumer can afford them. Note that I not saying that the points in the red shaded set *are* optimal. I refer to these points only to prove that A is *not* optimal.

Figure 8 shows the case in which the indifference curve is flatter than the budget line. The case that the indifference curve is steeper than the budget line is very similar. The only difference is that in the case of Figure 8 the "better" consumption bundles are to the left of the bundle A that we start with, whereas, when the indifference curve is steeper than the budget line, they will be to the right of the bundle A that we start with.

Moving to the left, or moving to the right, of A means that we are increasing, respectively decreasing, the consumption of good 1. Let's see in more detail why this change makes the consumer is better off through such a change. Let's stick with the case of Figure 8, that is, the indifference curve is flatter than the budget line. Mathematically speaking, this means:

$$\text{MRS} < \frac{p_1}{p_2}.$$

Now recall that we can calculate the MRS on the left hand side as the ratio of marginal utilities. Thus, we can write the inequality also as follows:

$$\frac{\frac{\partial U}{\partial q_1}}{\frac{\partial U}{\partial q_2}} < \frac{p_1}{p_2}.$$

Now let us multiply both sides by $\frac{\partial U}{\partial q_2}$, and also divide both sides by p_1. Both are positive, so that the direction of the inequality does not change, and we obtain:

$$\frac{\frac{\partial U}{\partial q_1}}{p_1} < \frac{\frac{\partial U}{\partial q_2}}{p_2}.$$

This inequality has a simple interpretation. The left hand side is the marginal utility per dollar spent on good 1, and the right hand side is the marginal utility per dollar spent on good 2. To see this for the left hand side suppose, for example, the marginal utility of good 1 were 4. That means, if the consumer were to buy an extra unit of good 1, then she would get 4 more units of utility. But maybe good 1 is quite expensive. For example, it might cost 10 Dollars per unit. Then, if the consumer spends one Dollar on good 1, she can only buy 0.1 units, and therefore, the extra marginal utility from just one Dollar spent on good 1 is only 0.1 times 4, that is, 0.4. Why do we normalize so that we calculate marginal utility per Dollar rather than marginal utility per unit of the good? We do it so that we can compare the utility of one Dollar spent on either of the two goods.

I adopt the ubiquitous, but annoying, habit of pretending that the Dollar is the only currency that matters.

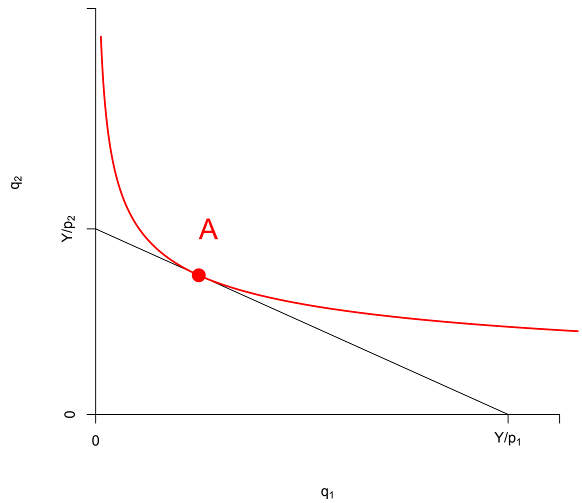
Now let's return to our inequality. Suppose we spend one more dollar on good 2. Then the gain in utility is the right hand side of the inequality. Where do we get the dollar from? We can spend one less dollar on good 1. Then we lose what is on the left hand side of the inequality. But the inequality says that the gain is larger than the loss. Thus, we can raise utility by spending more on good 2 and less on good 1. And that is why in Figure 8 we can raise utility by moving to the left: moving to the left means that we spend less on good 1, and more on good 2. if the inequality were reversed, and the indifference curve were steeper than the budget line, it would be the opposite.

The conclusion is that, for a consumption bundle to be a candidate for being optimal, it must be that the marginal rate of substitution in that bundle equals the price ratio, i.e. indifference curve and budget line have the same slope. Two lines have the same slope if they are tangential to each other. Thus, here, the budget line must be a tangent of the indifference curve. I show this in Figure 9. In point A, MRS and price ratio are just the same.

In Figure 9 it may seem as if the point A not only satisfied our two necessary conditions for a maximum, but as if it were indeed a maximum. But it would be a mistake that our two conditions were not just necessary, but also sufficient. We'll explain why in the next section.

Before we move on to sufficient conditions, there is one important caveat to what we have said so far which I still owe you. The logic explained above only

Figure 11: MRS = price ratio



makes sense if you *can* actually increase or decrease your consumption of good 1. How could it be that you can't? Well, if you already spend all your money on good 1, how can you increase your consumption of it? Conversely, if you consume none of good 1, how can you decrease your consumption of it? In either cases, the answer is you can't. The two cases correspond to the two corners of the budget set. in the bottom right corner, you already spend all your money on good 1. Therefore, even if the MRS is larger than the price ratio, you cannot increase your consumption of good 1. In the upper left corner, you consume nothing of good 1. Thus, even if the MRS is less than the price ratio, you cannot reduce your consumption of good 1.

In Figure 10 I show both possibilities. It is best to think of the two diagrams in Figure 10 as representing two different consumers. The one on the left has very flat indifference curves. Thus, in the top left corner this consumer may have an optimum even though the MRS is less than the price ratio. We are tempted to advise the consumer to consume less of good 1, but the consumer might rightly point out that she can't consume less than zero. The consumer on the right has very steep indifference curves. In the bottom right corner, she spends all of her income on good 1. Still, the MRS is larger than the price ratio. We would like to advise her to consume more of good 1, but once she has spent all her income on good 1, she can't consume more of that good.

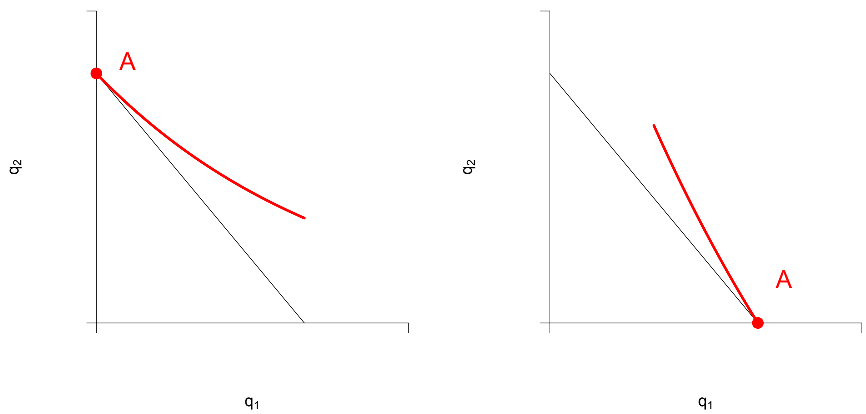


Figure 12: Corner Solutions

Thus, we have to modify our necessary condition a little bit. If the consumer consumes positive quantities of both goods, then MRS indeed has to equal the price ratio. If she only consumes good 1, then the MRS can be larger than the price ratio (which means that the marginal utility per dollar for good 1 is larger than it is for good 2), and if she consumes nothing of good 1, then the MRS can be smaller than the price ratio (which means that the marginal utility per dollar for good 1 is smaller than it is for good 2).

A Numerical Example

I'll now give a numerical example in which we use the insights of the previous section to calculate candidates for utility maximizing consumption bundles. The purpose of this example is really just to make sure that everything said up to this point is clear. There is not much (maybe: no?) additional economic insight that we gain from the example.

Let's pick some utility function:

$$U(q_1, q_2) = \sqrt{q_1} + \sqrt{q_2}.$$

Let us imagine the price of good 1 were $p_1 = 2$, the price of good 2 were $p_2 = 4$, and income were $y = 24$. We want to figure out the utility maximizing consumption bundle. Our first necessary condition is the budget constraint as an equality:

$$2q_1 + 4q_2 = 24.$$

To write down the second necessary condition, we first need to calculate the marginal rate of substitution. We use the formula that says that the MRS equals the ratio of the partial derivatives of the utility function.

$$\text{MRS} = \frac{\frac{\partial U}{\partial q_1}}{\frac{\partial U}{\partial q_2}} = \frac{\frac{1}{2\sqrt{q_1}}}{\frac{1}{2\sqrt{q_2}}}$$

Let's simplify the fraction on the right hand side a little, recalling that dividing by a fraction is the same as multiplying by the inverse of the fraction, and then dividing numerator and denominator by 2. We get:

$$\text{MRS} = \frac{\sqrt{q_2}}{\sqrt{q_1}}.$$

Now we are set to go: Let's first consider candidates for optimal consumption such that both q_1 and q_2 are strictly greater than zero. such a consumption bundle can be positive only if the MRS equals the price ratio. This means:

$$\frac{\sqrt{q_2}}{\sqrt{q_1}} = \frac{2}{4}$$

We square both sides, and reduce the fraction on the right hand side:

$$\frac{q_2}{q_1} = \frac{1}{4}$$

And we re-write this, after cross-multiplying, as:

$$q_1 = 4q_2$$

Thus, the MRS equals the price ratio if and only if the consumer consumers four times as much of good 1 as of good 2.

We now have two necessary conditions for consumption bundles for which both quantities are strictly positive: the budget constraint, and the condition that MRS=price ratio. Let's repeat them in one line:

$$2q_1 + 4q_2 = 24 \quad \text{and} \quad q_1 = 4q_2.$$

Both conditions must hold at the same time. We have a system of two very simple linear equations. Let's solve them. We plug the right hand side equation into the equation on the left hand side, i.e. we replace on the left hand side q_2 by $4q_2$, and we get:

$$2(4q_2) + 4q_2 = 24.$$

The left hand side equals $12q_2$, and thus q_2 equals 24 divided by 12, which gives us:

$$q_2 = 2.$$

We also need that q_1 is 4 times q_2, so that we also obtain:

$$q_1 = 8.$$

This is our first candidate for an optimal solution: the quantity consumed of good 1 is 8, and the quantity consumed of good 2 is 8.

We should also ask ourselves whether there are candidates for optimal solutions in which one of the two quantities is zero. These are the two corner points of the budget set. They can be calculated by setting one of the two quantities equal to zero in the budget constraint, and solving for the other one. In our case they are given by $q_1 = 0$ and $q_2 = 6$, and by $q_1 = 12$, and $q_2 = 0$.

Recall that a solution with $q_1 = 0$ is a candidate for an optimal choice only if it satisfies:

$$MRS < \frac{p_1}{p_2}.$$

But the MRS is given by: $\frac{\sqrt{q_2}}{\sqrt{q_1}}$, and therefore, when $q_1 = 0$, the MRS is infinity. Therefore, it is clearly not less than the price ratio, which is $\frac{2}{4}$. Thus $q_1 = 0$ and $q_2 = 6$ is not a candidate for an optimal solution.

How about $q_1 = 12$ and $q_2 = 0$? Recall that a solution with $q_2 = 0$ is a candidate for an optimal choice only if it satisfies:

$$MRS > \frac{p_1}{p_2}.$$

But, by our formula for the MRS, if $q_2 = 0$, then the MRS is zero, and therefore it is not larger than the price ratio. Therefore also $q_1 = 12$ and $q_2 = 0$ is not a candidate for an optimal solution.

We are left with a unique candidate for an optimal consumption bundle:

consume 8 units of good 1 and 2 units of good 2. We hope, of course, that the necessary conditions in this example are also sufficient. If they were not, i.e. if our only candidate were actually *not* optimal, we would be in true trouble. Let us now turn to sufficient conditions for maximization.

Sufficient Conditions for Utility Maximization

The basic insight in this section is quite simple. If preferences are convex, then the necessary conditions are at the same time also sufficient. To see this, look at Figure 9. If preferences are convex, then the indifference curves can never bend so as to cross the budget line again. Therefore, everything that is above the indifference curve through A, is not in the budget set. Put differently: all consumption bundles that are better than A cannot be afforded at the given prices and at the given income. The same is also true in the case of the corner solutions in Figure 10, provided that preferences are convex.

When preferences are convex then things are therefore very simple: the conditions derived earlier are *necessary and sufficient* for a consumption bundle to be optimal. Solving these conditions gives us the optimal consumption bundle, or the optimal bundles if there are several.

Let us check for our numerical example whether preferences are convex. The easiest way to do this is to calculate the MRS and check whether it is decreasing. We have already calculated the MRS for that example. It is:

$$\frac{\sqrt{q_2}}{\sqrt{q_1}}.$$

As we move on an indifference curve, q_1 becomes larger, and q_2 decreases. Because q_1 is in the denominator, the MRS gets smaller as q_1 increases. Because q_2 is in the numerator, the MRS also becomes smaller as q_2 creases. So, certainly it decreases as we move on an indifference curve. Therefore, we have indeed convex preferences. And thus we conclude that in this example the consumption bundle that we calculated in the previous section is indeed the unique optimal consumption bundle.

What about the case in which preferences are not convex? You can still use our necessary conditions to determine candidates for the optimal consumption. If there is just one such candidate, it must be optimal. Something must be optimal. But if there are several such candidates, the simplest method that I know of to determine which of these candidates is optimal is to calculate for each of the candidate solutions the utility, and to pick the one that has highest utility. Usually, that is a straightforward exercise. I give a numerical example in the next section.

A Numerical Example With Preferences That Are Not Convex

Let's consider this utility function:

$$U(q_1, q_2) = (q_1)^2 + (q_2)^2.$$

Let us suppose the price of good 1 were $p_1 = 4$, the price of good 2 were $p_2 = 3$, and income were $y = 60$. We shall use the necessary conditions to find candidates for utility maximizing consumption bundles. The first necessary condition is the budget constraint:

$$4q_1 + 3q_2 = 60.$$

Next, we calculate the marginal rate of substitution:

$$\text{MRS} = \frac{\frac{\partial U}{\partial q_1}}{\frac{\partial U}{\partial q_2}} = \frac{2q_1}{2q_2} = \frac{q_1}{q_2}.$$

Stop for a moment, and look at this expression. We see that the MRS *increases* rather than *decreases* as q_1 increases. Thus, we certainly don't have convex preferences. But our necessary conditions remain necessary.

Now we start looking for candidates for optimal consumption bundles. If we don't have a corner solution, then we have to have MRS = price ratio:

$$\frac{q_1}{q_2} = \frac{4}{3}.$$

Cross-multiplying we find:

$$q_1 = \frac{4}{3}q_2.$$

That sounds a little counterintuitive: even though goods 1 and 2 enter the utility function in the same way, and good 1 is more expensive than good 2, the condition says that consumption of good 1 is more than consumption of good

2. But the maths is right, and so, we continue.

If we don't have a corner solution, the two necessary conditions that we found are:

$$4q_1 + 3q_2 = 60 \quad \text{and} \quad q_1 = \frac{4}{3}q_2.$$

Let us plug the right hand side equation into the left hand side equation. We get:

$$\frac{16}{3}q_2 + 3q_2 = 60 \Leftrightarrow \frac{25}{3}q_2 = 60 \Leftrightarrow q_2 = \frac{180}{25} = 7.2.$$

Substituting this back into our equation for q_1 we get:

$$q_1 = \frac{4}{3}q_2 = \frac{4}{3} \cdot 7.2 = 9.6.$$

Thus, in our first candidate for an optimal consumption bundle, the consumer consumes $q_1 = 9.6$ units of good 1, and $q_2 = 7.2$ units of good 2.

Let's now think about the two possible corner solutions. One is that the consumer spends all money on good 1, and none on good 2. What is then the MRS? Because q_2 is in the denominator, we have that the MRS is then infinity. Our necessary condition says that the MRS has to be larger than the price ratio, and indeed it is. Thus, the corner satisfies the necessary condition, and our next candidate for an optimal solution is:

$$q_1 = \frac{60}{4} = 15 \text{ and } q_2 = 0.$$

Let's look into the other corner: no money is spent on good 1, and all money is spent on good 2. The MRS is zero. The necessary condition is that it is smaller than the price ratio, and indeed it is. We have a third candidate for an optimal solution:

$$q_1 = 0 \text{ and } q_2 = \frac{60}{3} = 20.$$

We have found three candidates for a utility maximizer. Now we follow the naive procedure and check utility in each of them. The first candidate was $(q_1, q_2) = (9.6, 7.2)$. This yields utility $U(q_1, q_2) = 9.6^2 + 7.2^2 = 144$. The second candidate, $(q_1, q_2) = (15, 0)$, yields utility $U(q_1, q_2) = 15^2 + 0^2 = 225$. The third candidate, $(q_1, q_2) = (0, 20)$ gives the consumer utility $U(q_1, q_2) = 0^2 + 20^2 = 400$. The third candidate gives the highest utility. The optimal consumption is zero units of good 1 and twenty units of good 2. This confirms the intuition that in the optimum the consumer consumes more of the good with the lower price. In fact, she only consumes that good in this example.

It helps to draw a diagram with indifference curves and budget line for this example. The indifference curves are interesting. Utility $(q_1)^2 + (q_2)^2$ is actually the square of the distance of the point (q_1, q_2) from the origin of the coordinate system, that is, the point with coordinates $(0, 0)$. From geometry, you perhaps remember that the distance of any point (x, y) from the origin is: $\sqrt{x^2 + y^2}$. This expression is the length of the straight line connecting $(0, 0)$ and (x, y). If we square it, the square root goes away. Therefore, in our example, the utility of a consumption bundle is the square of the distance of the consumption bundle from the origin.

An indifference curve shows where utility is constant. This means in our example that it shows where the square of the distance from the origin is constant. But this square can only be constant if the distance itself is constant. Thus, an indifference curve shows the location of points all of which have the same distance from the origin. But this means that the indifference curves are quarter circles around the origin. They are just *quarter* circles, rather than *full* circles because we only consider points were both quantities are positive.

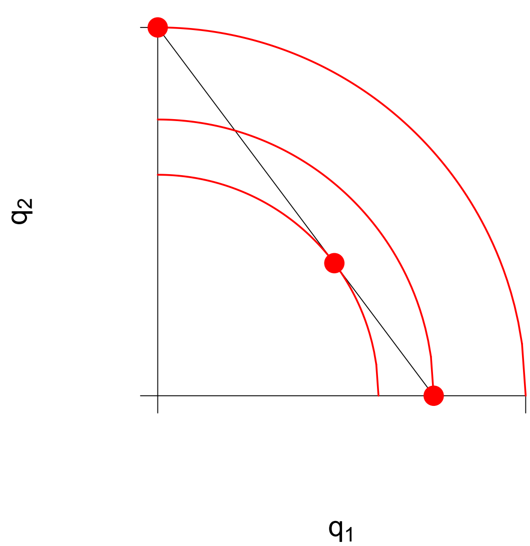


Figure 13: Non-Convex Preferences and Optimality

In Figure 11 I have sketched indifference curves and budget line for our example. I have marked the three consumption bundles that satisfy the necessary conditions for optimality as red dots. I have drawn through each of them the corresponding indifference curve. The point where the MRS equals the price ratio is the one in the middle. You can see that in the two corner points of the budget set the slope of the indifference curve is exactly as required by the necessary condition, and as shown in Figure 10. The top left corner is the optimum. It is on the highest indifference curve. You can see that the point where the indifference curve is tangential to the budget line is the *worst* point on the budget line, not the best.

Topic 5: Demand Functions and Compensated Demand Functions

The Demand Function

Which consumption bundle is optimal for the consumer depends obviously on the prices, and also on the consumer's income. We can construct a function that tells us for every income $y > 0$, and every pair of prices $p_1 > 0$ and $p_2 > 0$, how much a consumer with a given utility function buys of each of the two goods. We shall call this function the consumer's "demand function." There is a demand function for good 1, and one for good 2. The demand function for good 1 we shall denote by

$$D_1(p_1, p_2, y).$$

and the demand function for good 2 we shall denote by

$$D_2(p_1, p_2, y).$$

Thus, $D_1(p_1, p_2, y)$ is the quantity of good 1 (we previously denoted it by q_1) that the consumer is going to buy when income is y and prices are p_1 and p_2, and similarly $D_2(p_1, p_2, y)$ is the quantity that the consumer buys of good 2.

Which demand function a consumer has obviously depends on the consumer's utility function. In our notation, we don't show that, that is, we don't make the utility function U one of the arguments of the demand function. That makes our notation simpler. But it is important to keep in mind that everyone may have a different utility function, and that therefore everyone may have a different demand function.

When we pick some given and fixed numbers for p_2 and y, and plot $D_1(p_1, p_2, y)$ as a function of the remaining variable p_1 only, then we obtain what is most frequently called the "demand function" (thus using a slightly different terminology than we use here). It is often drawn as a linear function, as in Figure 12.

Let's calculate the demand function for a particular example of a utility function. This is just to make sure that we don't misunderstand each other,

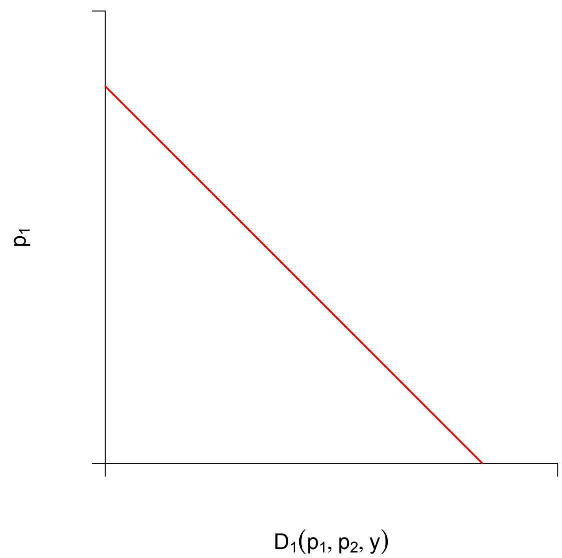


Figure 14: A Demand Function

and to make sure that we have an example at hand that we can use later to illustrate more results. Let's suppose that the utility function is:

$$U(q_1, q_2) = \sqrt{q_1} + \sqrt{q_2}.$$

We have encountered this utility function before, in Topic 4. We remember that it represents convex preferences. So, our necessary conditions for optimal consumption are also sufficient.

Recall the conditions. First there is the budget constraint:

$$p_1 q_1 + p_2 q_2 = y.$$

Secondly, we have to set the marginal rate of substitution equal to the price ratio. One can show that corner points cannot be optimal for this utility function, applying the same argument that we used in Topic 4 in our numerical example. We calculated the marginal rate of substitution in Topic 4. Plugging the formula that we got in Topic 4 into the equation MRS=price ratio, we get:

$$\frac{\sqrt{q_2}}{\sqrt{q_1}} = \frac{p_1}{p_2}.$$

We have to solve these two conditions for the two quantities, q_1 and q_2 . We start by solving for q_2. As we did in the previous Topic we shall square both sides of the condition MRS=price ratio:

$$\begin{aligned}
\frac{\sqrt{q_2}}{\sqrt{q_1}} &= \frac{p_1}{p_2} \Leftrightarrow \\
\frac{q_2}{q_1} &= \frac{(p_1)^2}{(p_2)^2} \Leftrightarrow \\
q_2 &= \frac{(p_1)^2}{(p_2)^2} q_1
\end{aligned}$$

To read this sequence of equations, you first need to know that I use the sign "$\Leftrightarrow$" at the end of an equation to indicate that the equation in the line below is equivalent to the equation to the left of the "$\Leftrightarrow$" sign. The second equation is equivalent to the first because we have squared both sides. The third equation is equivalent to the second because we have multiplied both sides by q_1.

Now let us plug in the re-written MRS=price ratio constraint into the budget constraint. This has the effect that we get an equation with only one unknown, q_1. Then we solve that equation for q_1. I go through many steps in the calculation below. It looks like a lot of algebra, but that is because I try to show every detail. You may read through it and try to understand each step by yourself. But at the end of the following equations, I shall also explain each step. So, bear with me.

$$p_1 q_1 + p_2 \frac{(p_1)^2}{(p_2)^2} q_1 = y$$

Cancelling out q_2 in the second term in the sum on the left hand side:

$$p_1 q_1 + \frac{(p_1)^2}{p_2} q_1 = y$$

Dividing the previous equation by p_1 on both sides:

$$q_1 + \frac{p_1}{p_2} q_1 = \frac{y}{p_1}$$

Factoring out q_1 on the left hand side:

$$\left(1 + \frac{p_1}{p_2}\right) q_1 = \frac{y}{p_1}$$

Replacing "1" on the left hand side by $\frac{p_2}{p_2}$ (hat looks a little silly, but trust me, it has its purpose):

$$\left(\frac{p_2}{p_2}+\frac{p_1}{p_2}\right)q_1=\frac{y}{p_1}$$

Writing the brackets on the left hand side as just one fraction:

$$\frac{p_1+p_2}{p_2}q_1=\frac{y}{p_1}$$

Dividing both sides by the fraction on the left hand side in front of q_1, we have solved for q_1, as we wanted to:

$$q_1=\frac{\frac{p_2}{p_1+p_2}y}{p_1}$$

Let's use our demand function notation. We have found:

$$D_1(p_1,p_2,y)=\frac{\frac{p_2}{p_1+p_2}y}{p_1}.$$

What is the demand function for good 2? We won't go through all the calculations again. We can just replace index "1" by index "2" and index "2" by index "1" in the demand function for good 1, and we get the demand function for good 2 (but note that we may do this only because the utility function is perfectly symmetric in q_1 and q_2):

$$D_2(p_1,p_2,y)=\frac{\frac{p_1}{p_1+p_2}y}{p_2}.$$

In Figure 13 I show the graph of this demand function, plugging in some fixed numbers for p_2 and y.

Finally, why have I written these demand functions in the form that I have chosen? The formulas look a little complicated. But it has a simple interpretation. Consider the demand function for good 1. The demand function for good 2 is analogous. Consider the numerator of the fraction that defines the demand function for good 1. It is: $\frac{p_2}{p_1+p_2}y$. Thus, it is a fraction times income. In fact, it is the fraction of income that the consumer spends on good 1. The demand function for good 1 is actually that fraction divided by the price of good 1. Thus, the numerator is the expenditure on good 1. The fraction of income spent on good 1 is $\frac{p_2}{p_1+p_2}$. If the price of good 1 is almost zero, then this fraction becomes 1. If the price of good 1 is almost infinity, then this fraction

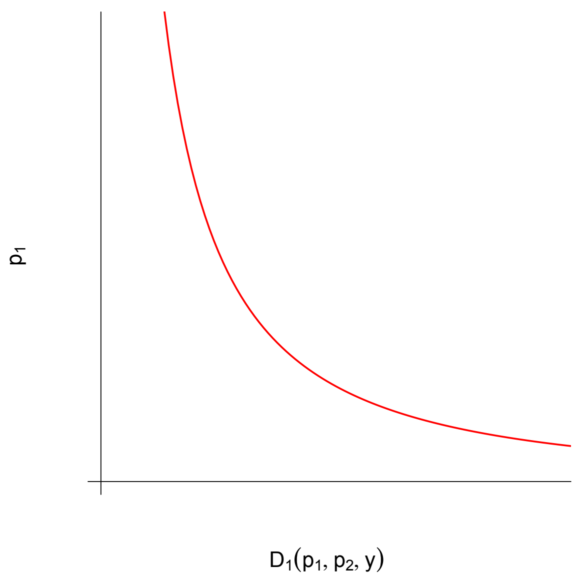


Figure 15: Demand Function in the Numerical Example

is almost zero. Thus, the fraction of income spent on good 1 goes down as the price of good 1 increases. This is one reason why demand decreases as p_1 increases. The second reason is, of course, that the quantity of good 1 that the consumer can buy with a given fraction of income decreases as p_1 goes up.

I hope this long discussion has made sense to you. The bottom line is just that we now have an example of a utility function, and the corresponding demand function, and, moreover, the demand function has a somewhat intuitive interpretation. Not much else matters about this long calculation.

The Compensated Demand Function

Our next step will certainly appear puzzling at first sight, but, as you read on, you'll see why it is useful for the study of utility maximizing demand. We shall study the expenditure minimization problem. Well, if you want to minimize expenditure, there is a simple solution: buy nothing. Unfortunately, that is not quite the solution to the problem that we have in mind. We shall study how to minimize expenditure if you want to reach a certain target level of utility. That is, unlike we did so far, we do not take "income" as given, but instead, we take a utility level as given. That is, suppose you have a certain utility function, and want to reach the utility level of, say, 10. Among all consumption bundles that give you utility 10, that is, all consumption bundles that are on the indifference curve corresponding to utility level 10, which one is cheapest?

Who in their right mind would decide to set themselves a target utility level, and then determine the cheapest way of finding this utility level? Think of it this way: you set yourself a target for the "living standard" that you want to achieve. Solving the expenditure minimization problem allows you to determine how much income you need to afford that utility level. The expenditure minimization problem is also, perhaps surprisingly, relevant for the question how to measure inflation. I shall demonstrate that later in this Topic. And, finally, even if your are only interested in the "normal" demand functions, certain properties of these are best understood if, at the same time, we study in parallel the compensated demand functions. This will hopefully become clear in Topics 6 and 7.

The solution of the expenditure minimization problem, of course, depends on the prices. For given utility level, and given prices, you obtain an optimal quantity of good 1, and an optimal quantity of good 2. We shall write these two optimal quantities as:

$$H_1(p_1, p_2, \bar{U})$$

and

$$H_2(p_1, p_2, \bar{U}).$$

Here, $\bar{U}$ is the utility level that the consumer wants to achieve. We use the letter "H" to denote these functions because they were invented by John Hicks, a Nobel prize winner in economics, and the two functions described above are also called the Hicksian demand functions. Another name for them is compensated demand functions. The motivation for this name is that these demand functions reflect what you would buy if in response to any price change someone would compensate you, that is, raise or lower your income, so that you have just enough to maintain your utility level.

Hicks is also the inventor of the "IS-LM model" that you learn about in Intermediate Macroeconomics. He invented it to clarify some of Keynes' macroeconomic theories.

As in the case of utility maximization, we are now going to study necessary, and then sufficient conditions for expenditure minimization. Don't worry. It's going to be quick. Suppose the consumer has some utility function, has a target utility level $\bar{U}$, and looks for the cheapest way of achieving that utility level. Well, to advise her, we shall first sit down with her and draw for her the indifference curve that corresponds to the target utility level. If she has nicely convex preferences, this indifference curves looks like the red line in Figure 14.

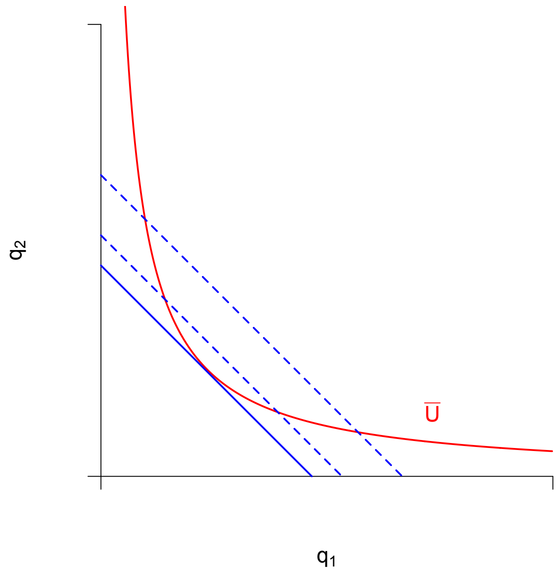


Figure 16: Expenditure Minimization

To find the cheapest point on this red indifference curve, let us just take a guess. Let's guess it is y Dollars. It is of course, a little trick that I denote this expenditure level by y, the letter that so far we have used as our symbol for income. I hope you forgive this trick. Let us find all consumption bundles for which the expenditure level is y. They are given by this equation:

$$p_1 q_1 + p_2 q_2 = y.$$

Well, of course that is our familiar budget equation. Thus, it is a straight line with slope $-\frac{p_1}{p_2}$. Suppose I had guessed y too high: then our straight line might look like one of the dashed lines in Figure 14. With the expenditure levels y that correspond to these lines you can certainly achieve the given utility level: this is because these dashed lines intersect with the indifference curve, actually they do so each in two points. Pick one of these points, and you get the target utility level with the guessed expenditure level y. But you can achieve $\bar{U}$ in a cheaper way. Consider the unbroken blue line. Because it is closer to the origin it corresponds to a lower expenditure level than the dashed lines. Still, it has one point in common with the indifference curve. By picking that point, you can lower your expenditure. Can you go further? No. If you shift the blue line parallel closer to the origin, then there will no longer be an intersection point with the red indifference curve, and therefore you can no longer achieve the target utility level $\bar{U}$. Thus, the point where the blue unbroken line touches the red indifference curve is the expenditure minimizing consumption bundle. It represents for this utility level, for the given prices, the compensated demand for goods 1 and 2.

What does Figure 14 tell us about necessary conditions for expenditure minimization? Actually, of course, it is a very familiar figure, and it shows that the following condition is necessary for expenditure minimization:

$$\text{MRS} = \frac{p_1}{p_2}.$$

Marginal rate of substitution must equal the price ratio. And, of course, this is true unless one of the quantities is zero. If the marginal rate of substitution is larger than the price ratio, then the quantity of good 2 must be zero. If it is smaller than the price ratio, then the quantity of good 1 must be zero. The logic is exactly the same as it is in the corner solutions that we discussed in Topic 4.

We have one necessary condition, but there is another one: it is the condition that the utility level is the one we are aiming for:

$$U(q_1, q_2) = \bar{U}.$$

Finally, you guessed it, if preferences are convex, then the necessary conditions are also sufficient, that is, each solution to the necessary condition is not just a *candidate* for an expenditure minimizing consumption bundle, but we can be sure that it *is* an expenditure minimizing consumption bundle. Thus, the structure of the expenditure minimization problem is very similar to the structure of the utility maximization problem.

Let us solve again a numerical example. We postulate the same utility function as before:

$$U(q_1, q_2) = \sqrt{q_1} + \sqrt{q_2}.$$

We wrote down and simplified the MRS=price ratio condition earlier already:

$$q_2 = \frac{(p_1)^2}{(p_2)^2} q_1.$$

(We don't worry about corner solutions for this calculation. There are no such solutions in this example.) The second necessary condition is:

$$\sqrt{q_1} + \sqrt{q_2} = \bar{U}.$$

We want to solve for q_1 and q_2. Let us plug the first necessary condition into the second:

$$\sqrt{q_1} + \sqrt{\frac{(p_1)^2}{(p_2)^2} q_1} = \bar{U}$$

We shall solve this for q_1. Using the fact that the square root of a product equals the product of the square roots of the factors

$$\sqrt{q_1} + \sqrt{\frac{(p_1)^2}{(p_2)^2}} \sqrt{q_1} = \bar{U}$$

Taking the square root of the first factor in the second term on the left hand side:

$$\sqrt{q_1} + \frac{p_1}{p_2} \sqrt{q_1} = \bar{U}$$

Bracketing out on the left hand side:

$$\left(1 + \frac{p_1}{p_2}\right) \sqrt{q_1} = \bar{U}$$

Re-writing the bracket on the left hand side:

$$\frac{p_1 + p_2}{p_2} \sqrt{q_1} = \bar{U}$$

Dividing both sides by $\frac{p_1+p_2}{p_1}$:

$$\sqrt{q_1} = \frac{p_2}{p_1 + p_2} \bar{U}$$

Squaring both sides, we have solved for q_1, as we wanted to:

$$q_1 = \left(\frac{p_2}{p_1 + p_2}\right)^2 (\bar{U})^2$$

The corresponding calculation for good 2 is analogous. Thus, we obtain these compensated demand functions:

$$H_1(p_1, p_2, \bar{U}) = \left(\frac{p_2}{p_1 + p_2}\right)^2 (\bar{U})^2$$

and

$$H_2(p_1, p_2, \bar{U}) = \left(\frac{p_1}{p_1 + p_2}\right)^2 (\bar{U})^2.$$

There is nothing too exciting about these formulas. Perhaps, by working through these calculations, you could check that you really understood what compensated demand is. This understanding does matter in what follows. It will also be handy to have the above formulas as an example at hand as we proceed.

The Indirect Utility Function and The Expenditure Function

While we are at it, we might as well introduce two more functions, one related to the utility maximization problem, and one related to the expenditure minimization problem. My excuse for introducing these functions is, as always, that they will matter later.

Consider the utility maximization problem. Suppose we had done all our work, and had determined the quantities demanded $D_1(p_1, p_2, y)$ and $D_2(p_1, p_2, y)$. Then we could plug these quantities back into the utility function, and we could figure out which utility level the consumer actually achieves when she maximizers her utility. Obviously, this utility level will depend on prices and income. We shall denote it by:

$$V(p_1, p_2, y).$$

The function V is also called the "indirect utility function." For any given prices and incomes it tells the consumer what is the highest utility level that she can achieve if she spends her income in a utility maximizing way.

Similarly, consider the expenditure minimization problem. Suppose we had determined the expenditure minimizing quantities $H_1(p_1, p_2, \bar{U})$ and $H_2(p_1, p_2, \bar{U})$. Then we could multiply these quantities by the prices, and add up the expenditures on both goods, to find the minimum expenditure needed to achieve utility level $\bar{U}$ when prices are p_1 and p_2. We denote this expenditure level by:

$$E(p_1, p_2, \bar{U}).$$

The function E is also called the "expenditure function." For any given prices and target utility level it tell the consumer what is the lowest expenditure level that allows her to achieve the target utility level.

To make things as clear as possible (to me), let me write down for our example the indirect utility function and the expenditure function. The indirect utility function can be obtained by plugging in the terms that we have found earlier as the quantities demanded into the utility function:

$$V(p_1, p_2, y) = \sqrt{\frac{\frac{p_2}{p_1+p_2}y}{p_1}} + \sqrt{\frac{\frac{p_1}{p_1+p_2}y}{p_2}}.$$

Of course, I used here our earlier calculation of the demand functions. Note that I never promised you that the expression that we get would be interesting! But you can check your understanding of the definition of the indirect utility function by checking whether you know exactly why the above formula is correct. The situation for the expenditure function is similar. The expression for it is not exactly inspiring, but it is worth your while making sure that you understand where it comes from.

$$E(p_1, p_2, \bar{U}) = p_1 \left(\frac{p_2}{p_1 + p_2}\right)^2 (\bar{U})^2 + p_2 \left(\frac{p_1}{p_1 + p_2}\right)^2 (\bar{U})^2.$$

Let us simplify this last expression further. We'll use it below.

$$\begin{aligned}
E(p_1, p_2, \bar{U}) &= p_1 \left(\frac{p_2}{p_1 + p_2}\right)^2 (\bar{U})^2 + p_2 \left(\frac{p_1}{p_1 + p_2}\right)^2 (\bar{U})^2 \\
&= \frac{p_1 (p_2)^2}{(p_1 + p_2)^2} (\bar{U})^2 + \frac{p_2 (p_1)^2}{(p_1 + p_2)^2} (\bar{U})^2 \\
&= \frac{p_1 (p_2)^2 + p_2 (p_1)^2}{(p_1 + p_2)^2} (\bar{U})^2 \\
&= \frac{p_1 p_2 (p_1 + p_2)}{(p_1 + p_2)^2} (\bar{U})^2
\end{aligned}$$

$$= \frac{p_1 p_2}{p_1 + p_2}(\bar{U})^2.$$

If you want to do more calculations: you might think that the expenditure function should be increasing in the prices. Check for this example that this is indeed the case. And, getting more ambitious, can you prove that this is true in general, not just in this example?

Measuring Inflation

I mentioned earlier that the expenditure minimization problem might help us to understand better the problem of inflation measurement. Inflation measures are supposed to capture by how much prices change *on average*. Some prices might go up a lot from one year to the next, some others might even fall, so when we say that inflation from one year to the next was 5%, we must have some way of averaging the percentages by which different prices change. There are many ways of doing this, and they are often taught in courses on descriptive statistics. For example, you might learn in such a course about the "Laspeyres" and the "Paasche" indices. You can look these up on Wikipedia, too. Here, the detailed definitions of these price indices will not matter. The main point is that the number that you get depends on the formula that you use for calculating inflation, and that we can use the expenditure minimization problem to understand *why* there are different ways of calculating the inflation rate.

The main point is this: we need to decide, when averaging over the price changes of different goods, which weight we should give to each good. Roughly speaking, goods that we spend much money on should have more weight than goods that we spend little money on. But this leaves open an important question: Should we use the quantities that we purchase before the price change, or after the price change? For example, we might buy a lot of cereal of one brand at some point. Then the price of that cereal goes up by a lot. Maybe this price increase induces us to switch to another, much cheaper cereal brand that we like just as much. Then the price increase should have little weight, and we might capture this by making the weight proportional to the quantity that we buy after the price increase. But if we do that, we might also give little weight to other cases in which the reason that we buy little of a good after the price increase is not that we have found a close substitute, but simply that we cannot afford the good anymore. Shouldn't this good receive a lot of weight in our price index?

The expenditure function gives us a nice way of constructing a price index. Suppose we fix a utility level $\bar{U}$. Then, we can calculate the percentage change in expenditure needed to achieve that utility level, and that will be a natural measure of the inflation rate. It reflects how the price changes affect the consumer who maximizes utility, and who adjusts their consumption as prices change. IF a consumer can switch in response to a price increase to some other close substitute, then the required expenditure to achieve a given utility level will not change much. But, if a drastic increase in car prices forces the consumer to give up cars at her given income, then in the expenditure function this will be reflected by a large change, because the expenditure function keeps the utility level, not the income, fixed.

In our notation, if prices change from p_1, p_2 to $\hat{p}_1, \hat{p}_2$, then the natural measure of inflation is the percentage by which $E(\hat{p}_1, \hat{p}_2, \bar{U})$ differs from $E(p_1, p_2, \bar{U})$. But a new problem emerges: Which utility level $\bar{U}$ should we choose? And for which consumer's expenditure function should we use? Different consumers have different expenditure functions because they have different utility functions, and the expenditure function depends on the utility function. There is no perfect answer to either of these two questions. And it seems to me that precisely because these two questions don't have a perfect answer, we don't have a perfect measure of inflation. The existing measures of inflation can be interpreted as imperfect guesses to how we might answer the two questions.

It is interesting, though, to consider the inflation measure for the consumer whom we described in the previous section. To find the percentage change in expenditure, we would calculate this ratio:

$$\frac{E(\hat{p}_1, \hat{p}_2, \bar{U})}{E(p_1, p_2, \bar{U})}.$$

Let's plug in the formula for the expenditure function that we found in the previous section:

$$\frac{E(\hat{p}_1, \hat{p}_2, \bar{U})}{E(p_1, p_2, \bar{U})} = \frac{\frac{\hat{p}_1\hat{p}_2}{\hat{p}_1+\hat{p}_2}(\bar{U})^2}{\frac{p_1 p_2}{p_1+p_2}(\bar{U})^2} = \frac{\hat{p}_1\hat{p}_2}{p_1 p_2} \cdot \frac{p_1+p_2}{\hat{p}_1+\hat{p}_2}.$$

That is undoubtedly a strange looking measure of inflation. But, for this consumer, it would be precisely right. Note, by the way, that the utility level $\bar{U}$ dropped out in our calculation. So, we don't have to solve the problem of choosing the "right" utility level. What does our calculation mean? Suppose the price of good 1 increases from 5 to 8, and the price of good 2 decreases from 6 to 5, what is our inflation rate? We first calculate the above ratio:

$$\frac{\hat{p}_1 \hat{p}_2}{p_1 p_2} \cdot \frac{p_1 + p_2}{\hat{p}_1 + \hat{p}_2} = \frac{8 \cdot 5}{5 \cdot 6} \cdot \frac{5+6}{8+5} \approx 1.128,$$

so that the inflation rate would be 12.8%. The price of good 1 increased by 60%, and the price of good 2 dropped by 17%. Our sophisticated way of average price changes lead to the conclusion that the average price change is 12.8%. Certainly not an obvious conclusion.

Duality

Let us return to the observation that the optimal solution in Figure 14 that represents the expenditure minimizing consumption bundle looks at the same time like the graph that we have shown to find utility maximizing consumption bundles. There is a somewhat more precise and formal way of stating this fact.

Suppose we had maximized utility for a given budget y, and at given prices p_1 and p_2. Let $D_1(p_1, p_2, y)$ and $D_2(p_1, p_2, y)$, as usual, denote the utility maximizing consumption bundle. The utility level that we have achieved then is $V(p_1, p_2, y)$. Let us denote this utility level now by $\bar{U}$. Now suppose we asked a different question: What is the expenditure minimizing consumption bundle that achieves the utility level $V(p_1, p_2, y)$? In this question, we leave aside for the moment that there was a given budget y, and we just ask how we can reach this utility level in the cheapest possible way? Figure 14 shows that the answer to this will again be: $D_1(p_1, p_2, y)$ and $D_1(p_1, p_2, y)$. That is:

$$H_1(p_1, p_2, \bar{U}) = D_1(p_1, p_2, y) \quad \text{and} \quad H_2(p_1, p_2, \bar{U}) = D_2(p_1, p_2, y).$$

That is, quantities demanded are equal to the expenditure minimizing quantities for utility level $\bar{U}$. Figure 14 also shows that the level of expenditure that is needed to reach utility level $\bar{U}$ is exactly y, the income with which we started this discussion:

$$E(p_1, p_2, \bar{U}) = y.$$

These equalities are clear from Figure 14, but we can also explain to ourselves in words why they must be true. Suppose I came up with a consumption bundle to minimize expenditure that is different from $D_1(p_1, p_2, y)$ and $D_1(p_1, p_2, y)$. Then I must have figured out a way to reach utility level $\bar{U}$ with expenditure that is less than the expenditure that $D_1(p_1, p_2, y)$ and $D_1(p_1, p_2, y)$ required, and this expenditure is y. That is, it is possible to reach $\bar{U}$ with expenditure of less than y. But then, when we maximize utility with income y, we must come up with a utility level larger than $\bar{U}$. But that contradicts our assumption that $\bar{U}$ is the largest utility level that the consumer can reach with income y.

This is a very general argument. It applies just as well when there are more than two goods. Note that we have used almost no assumption, except one: preferences are monotone. Do you see where we have used this assumption? The answer is that we used it when asserting that, if it is possible to reach $\bar{U}$ with expenditure of less than y, then, when we maximize utility with income y, we must come up with a utility level larger than $\bar{U}$. The argument that justifies is that we can choose the expenditure minimizing consumption bundle, and have a little bit income left. So, we can buy some more of one or both goods. Monotonicity then implies that this makes the consumer better off.

The equalities that we derived above are called the duality equalities. I'll skip the story that explains this name. We shall use these equalities in Topic 7.

Topic 6: The Effect of Income and Price Changes on Demand

A consumer's demand function for any particular good, say good 1, depends on good 1's price, and on income. We shall discuss how it depends on the price of good 2 in Topic 8. We shall study whether the theory of utility maximization that we have developed so far makes predictions about how the quantity demanded of a good depends on the good's own price, and on income. For example, we shall ask: does the theory of utility maximization imply that the higher the price of good 1 the less you demand of it? Interestingly, the answer is that this does not follow automatically. There are cases, in which a rational consumer demands more of a good when the good's price rises. In this section you will learn why.

The Effect of Income Changes on Demand

Let's start with income changes. If income goes up, will a rational consumer buy more of each good? The answer is "not necessarily." In Figures 15 and 16 I illustrate this in two diagrams. In both figures, I keep the prices fixed, and consider what happens if we increase income. Here, Y denotes the old income, and $\hat{Y}$ denotes the new, higher income. When prices don't change, an increase in income leads to a parallel outward shift of the budget line. Thus, in both figures you see two budget lines. They correspond to the two income levels that I am considering. The line further away from the origin corresponds to the higher income level, $\hat{Y}$.

For each income level I have indicated the utility maximizing demand. It is where an indifference curve is tangential to the budget line. I have focused on the simple case of convex preferences and no corner solution. In each figure, there are two points of tangency, corresponding to the two income levels. By dropping a vertical line from the utility maximizing point on the horizontal axis, we can read off how much of good 1 the consumer buys at the given income level. In Figure 15 this amount goes up as income increases, but in Figure 16 demand for good 1 goes down as income increases.

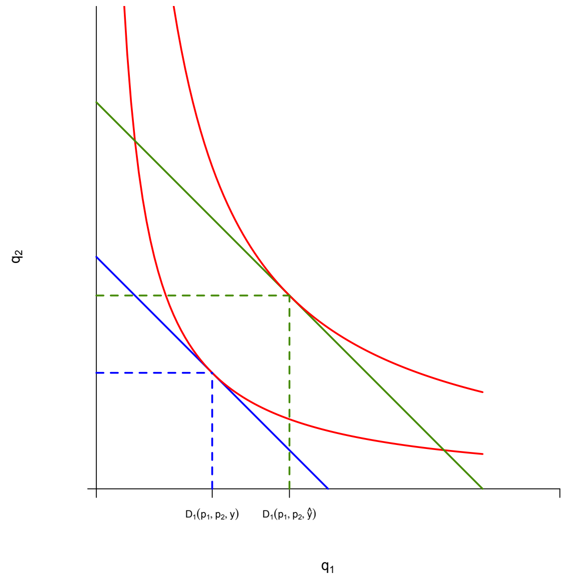


Figure 17: Normal good: demand for good 1 increases as income goes up

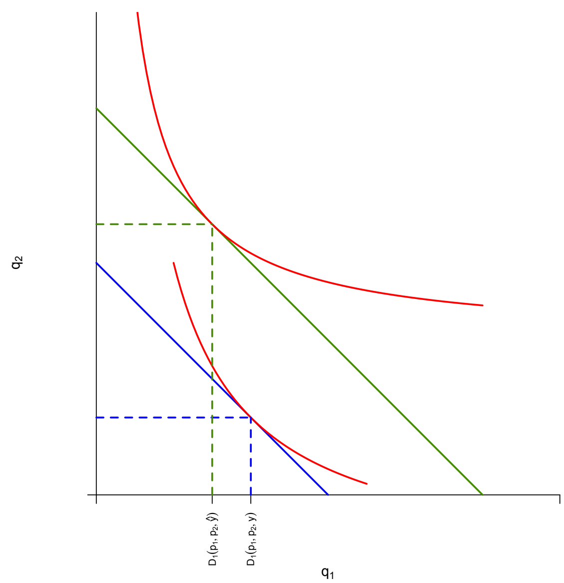


Figure 18: Inferior good: demand for good 1 decreases as income goes up

Note that I did not have to do anything strange to indifference curves to obtain the example in Figure 16. Indeed, there is nothing too strange about demand for a good falling as income goes up. We call such goods inferior goods. For some people, furniture that you have to assemble yourself, and that is made of plywood is an inferior good. As they get rich, they buy higher quality furniture. A good for which demand goes up as income goes up is called a normal good. Vacations might be an example of a normal good. As your income rises, you buy more vacations. But of course, what are inferior goods, and what are normal goods, is an empirical question. Here, we just speculate.

One final point: A good need not be "universally" inferior, or normal. It can be inferior at some income and price levels, but normal at other income and price levels. For example, fixing prices, demand may go up and down, and up and down again, as income increases. Nothing in the theory of utility maximization rules that out. One thing is ruled out, though. No good can be inferior for all income levels. That is so because, at income level zero, demand must be zero. If demand never went up as income increases, it couldn't go down either, because then it would have to be negative. So, if a good is inferior at some income levels, it must be normal at some lower income levels.

The Effect of Price Changes on Demand

The demand for any good, say good 1, potentially depends on the prices of all goods, not just the good's own price p_1. But in this section we shall investigate the effect of a change in the good's own price. For concreteness, we shall ask how the demand for good 1 changes as the price of good 1 goes up. But the case that the price of good 1 goes down is exactly analogous.

I promised you earlier that I would explain to you why the theory of utility maximization does *not* necessarily imply that an increase in the price of a good causes demand for that good to go down. You might have heard that the law of demand in economics says that a utility maximizing consumer always reduces demand for a good if the price of that good goes up. What I am telling you is that this is simply false. There is however a true law of demand, and our objective in this section is to find it!

First, let me just demonstrate to you that the effect of a price increase may be to reduce demand, but can also be to increase demand. I do this in Figures

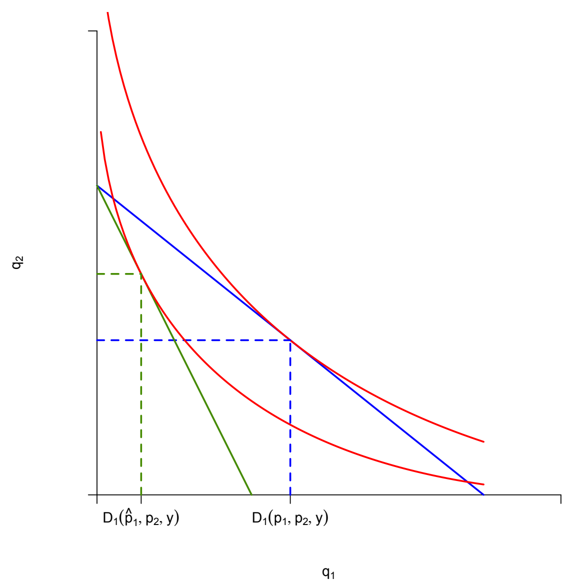


Figure 19: Regular good: demand for good 1 decreases as the price of good 1 goes up

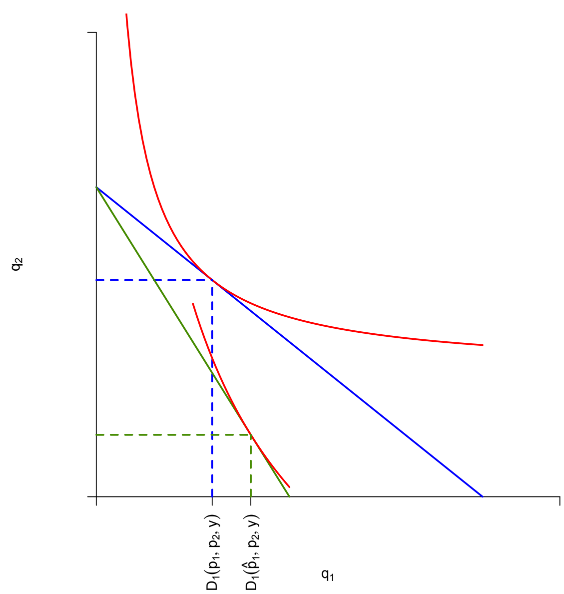


Figure 20: Giffen good: demand for good 1 increases as the price of good 1 goes up

17 and 18. In those figures, I denote the original price of good 1 by p_1, and the new, higher price of good 1 by $\hat{p}_1$. Keeping income, and all other prices, fixed, what does an increase in the price of good 1 do to the budget line? The slope of the budget line is $-\frac{p_1}{p_2}$. Thus, as p_1 goes up, the budget line becomes steeper. In fact, it rotates. The intercept with the vertical axis is not changed. The budget line rotates around this point. Why is this? The intercept with the vertical axis indicates the amount of good 2 that the consumer can buy if she buys none of good 1, and spends all her income on good 2. This amount does not change if the price of good 1 increases, obviously. It only depends on income and the price of good 2. But the intercept with the horizontal axis changes. This intercept indicates the amount of good 1 that the consumer can buy if she spends all her income on good 1. As the price of good 1 increases, this quantity decreases.

Each of Figures 17 and 18 thus contains two budget lines, one, the flatter one, corresponds to the budget *before* the increase in the price of good 1, and the other one, the steeper one, corresponds to the budget *after* the increase in the price of good 1. For each budget line I have indicated the utility maximizing consumption bundle. It is where an indifference curve is tangential to the budget line. If we drop a vertical line from the utility maximizing consumption bundle on the horizontal axis, we can see the amount of good 1 that is bought when the consumer maximizes utility. In Figure 17 the amount by which the quantity demanded of good 1 goes down as price increases. In Figure 18, the quantity demanded of good 1, however, goes up. Thus Figure 18 is my proof that there is no naive law of demand.

For one of the two indifference curves in Figure 18 I only show a short piece of that indifference curve. This is due to the limitations of my ability to produce nice graphs. It has no further importance.

A good for which demand increases as the price goes up is called a Giffen good, named after a 19th century Scottish economist and statistician Robert Giffen. Robert Giffen is reported to have observed that poor people's purchases of low quality of bread went up rather than down as the price of bread went up. We shall return to Giffen's supposed discovery at the end of this Topic. There is no standard name in the literature for goods that are *not* Giffen goods. In these notes I shall call such goods regular goods. I would have liked to use the phrase *normal* goods. But we already introduced that terminology earlier in this topic for a different idea. So, I needed a different word. "Regular" sounds good enough.

Substitution and Income Effects

With some work, we can obtain a deeper understanding of how it can be that there are Giffen goods. Suppose the price of good 1 goes up from p_1 to $\hat{p}_1$. Suppose also, as before, that the price of good 2, and income, remain the same. Our starting point is the simple observation that the consumer will be disappointed in the sense that after the price increase her utility level will be lower than it was before. Using the notation introduced in the previous topic:

$$V(p_1, p_2, y) > V(\hat{p}_1, p_2, y).$$

Geometrically: since the budget line pivots inwards, we are inevitably moving to a lower indifference curve. Let's call the utility level on the left hand side of the above inequality $\bar{U}$, and let's call the utility level on the right hand side $\hat{U}$.

We shall now make up a little piece of fiction. Suppose, after the price of good 1 went up, the consumer came to us and complained that her life had gotten worse. Because we are nice people, we decide to help: we give the consumer exactly the extra amount of money that she needs to maintain her old utility level $\bar{U}$. How much should we give her? She could say to us: well, to buy exactly the same quantities as I bought before, you need to give me the amount of the price increase times the quantity of good 1 that I am used to buying. We would respond: No, not necessarily. You previously had utility level $\bar{U}$. All that we are going to give you is what is needed to maintain utility level $\bar{U}$. Because prices have changed, you might choose a different consumption bundle than before to achieve that utility level, and then we would have to pay you less than you think. Now, in the notation that we introduced in the previous topic, the cheapest way in which the consumer can maintain her utility level is actually by purchasing the compensated demand, i.e. $H_1(\hat{p}_1, p_2, \bar{U})$ and $H_2(\hat{p}_1, p_2, \bar{U})$, because that is how compensated demand was defined. To facilitate that consumption, we would just pay the consumer $E(\hat{p}_1, p_2, \bar{U})$, which is less than she asked for. She will be OK. It is enough for her to maintain her utility level. As our focus is on the demand for good 1, note that, after we have compensated her, the consumer's demand for good 1 would adjust from $D_1(p_1, p_2, y)$ to $H_1(\hat{p}_1, p_2, \bar{U})$.

Next, suppose that some change in the circumstances of our lives made it impossible for us to continue compensating this consumer, or suppose that our character changed unexpectedly: whereas before we were nice, now we are decidedly not nice. We withdraw our support for the consumer. Now she has to face reality. Her utility drops from $\bar{U}$ to $\hat{U}$, and her demand for good 1 adjusts from $H_1(\hat{p}_1, p_2, \bar{U})$ to $D_1(\hat{p}_1, p_2, y)$.

This was a little piece of fiction. But it motivates dividing the total effect of the increase in the price of good 1 from p_1 to $\hat{p}_1$ into two components, the initial change when the consumer is compensated:

$$H_1(\hat{p}_1, p_2, \bar{U}) - D_1(p_1, p_2, y),$$

and the subsequent change, when the compensation is withdrawn:

$$D_1(\hat{p}_1, p_2, y) - H_1(\hat{p}_1, p_2, \bar{U}).$$

Adding up these two differences gives us the total change in demand:

$$H_1(\hat{p}_1, p_2, \bar{U}) - D_1(p_1, p_2, y) + D_1(\hat{p}_1, p_2, y) - H_1(\hat{p}_1, p_2, \bar{U}) = D_1(\hat{p}_1, p_2, y) - D_1(p_1, p_2, y).$$

We shall study the two components separately. We shall call the first part of the difference the substitution effect, and the second part of the change the income effect. You will soon see the motivation for this terminology. Just to make things clear, let's write down the previous equation one more time, indicating where the two effects are.

$$\underbrace{H_1(\hat{p}_1, p_2, \bar{U}) - D_1(p_1, p_2, y)}_{\text{substitution effect}} + \underbrace{D_1(\hat{p}_1, p_2, y) - H_1(\hat{p}_1, p_2, \bar{U})}_{\text{income effect}} = \underbrace{D_1(\hat{p}_1, p_2, y) - D_1(p_1, p_2, y)}_{\text{total effect}}.$$

We'll now study the two effects, substitution and income effect, separately in detail. Let's begin with the substitution effect. This is the adjustment to the quantity of good 1 that the consumer makes when we compensate her for so that she can maintain her utility. We want to argue that for a rational consumer this effect has to be negative, i.e. she has to reduce her demand for good 1 (or keep it constant), but it cannot be that she will increase her demand for good 1. This seems eminently plausible. One can also give a mathematical proof. The following paragraph is a translation of this mathematical proof into words.

> The argument is indirect. We start by supposing that, to maintain the consumer's utility level with the lowest possible expenditure, once the price of good 1 rises, it is optimal for the consumer to purchase *more* of good 1 and *less* of

> good 2 than before. We shall argue that this leads to a contradiction. First, we note that under our hypothesis it must be that the extra expenditure on good 1 is less than the reduction in expenditure on good 2. If this is true *after* the price has increased, it must also be true *before* the price has increased, because before the price increase the extra expenditure on good 1 would have been even lower, and the reduction in expenditure on good 2 would have been the same. Thus, even before the price increase this adjustment in the consumer's utility level would have maintained her utility level, and would have reduced her expenditure. But this implies that before the price increase the consumer was not maximizing utility, and that is our contradiction. Why was she not maximizing utility? She could have increased her expenditure on good 1, reduced her expenditure on good 2, saved some money, and then use that saved money to obtain a higher utility than she did.

This was a somewhat complex argument. But, the bottom line is: the substitution effect is negative.

Now consider what we have called the "income effect." It is the difference in the consumer's demand for good 1 that is caused by a change in the consumer's income: we withdraw the income support for the consumer. It is not caused by any change in prices. As we discussed earlier, however, the theory of utility maximization does not predict unambiguously positive or negative income effects. Here, we are dealing with a reduction in income. If we have an inferior good, then this may well cause demand to go up, i.e. the income effect of a price increase would be positive. For a normal good, it would be negative. For the total effect to be positive, i.e. demand increases as the price goes up, good 1 thus has to be an inferior good, and, moreover, the income effect has to be large enough to overcome the substitution effect.

Let me illustrate substitution and income effects for the cases of a regular good, and a Giffen good. Figure 19 shows the case of a regular good. The blue budget line rotates around its intersection with the vertical axis, indicating an increase in the price of good 1. Had we compensated the consumer just enough to maintain her old utility level after the price increase, we would have paid her the expenditure that allows her to purchase the bundle marked as $H(\hat{p}_1, p_2, \bar{U})$. As is always the case, this consumption bundle involves a smaller quantity of good 1. The reduction in the quantity of good 1 is the substitution effect. The remaining change in the demand for good 1 is the income effect. In the figure, the income effect is also negative. So, certainly, the total effect is negative, and we have a regular good.

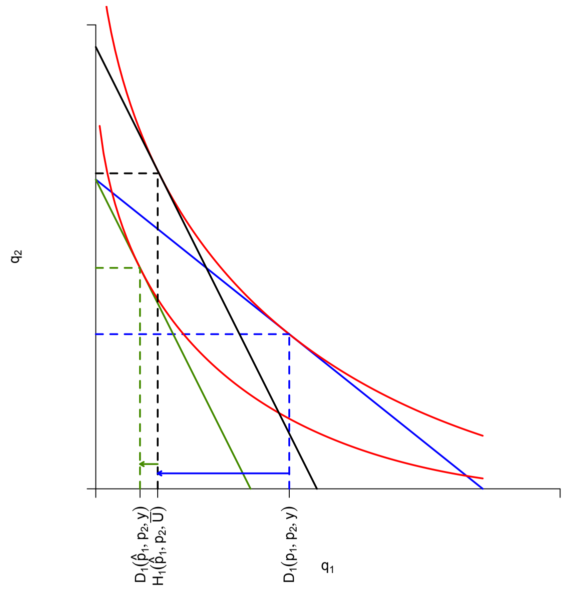


Figure 21: Substitution and Income Effects for a Regular Good. (The length of the blue arrow is the substitution effect and the length of the green arrow is the income effect.)

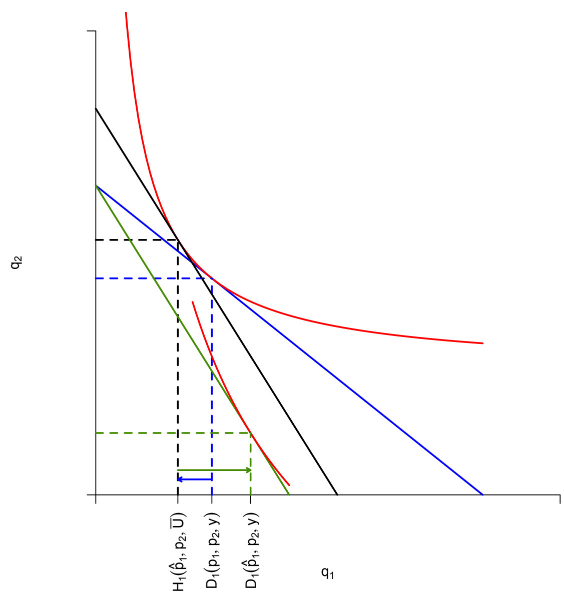


Figure 22: Substitution and Income Effects for a Giffen Good. (The length of the blue arrow is the substitution effect and the length of the green arrow is the income effect.)

Now consider Figure 20. In this figure, you can see that the substitution effect is again negative, but it is small. The income effect, although it corresponds to a small shift in income only, is very large, and it is positive. As the consumer's income drops, she buys much more of good 1. This second effect is larger than the first effect, and we have a Giffen good.

In the figures it is clear that the substitution effect is negative because it corresponds to a rotation to the left along an indifference curve. This geometric intuition seems to be related to the convexity of the indifference curve. But note that the verbal argument that we provided earlier to show that the substitution effect is negative is actually not related to convex preferences. The substitution effect is always negative, regardless of the shape of indifference curves.

A Numerical Example

In the lecture notes for Topic 5 we studied the utility function

$$U(q_1, q_2) = \sqrt{q_1} + \sqrt{q_2}$$

and found that in this example the demand function for good 1 is:

$$D_1(p_1, p_2, y) = \frac{p_2}{p_1 + p_2} \frac{y}{p_1}$$

Suppose $y = 12$, $p_2 = 2$, and the price of good 1 rises from $p_1 = 1$ to $\hat{p}_1 = 2$. Plugging into the formula for demand, we can see that demand for good 1 drops from

$$\frac{2}{1+2} \cdot \frac{12}{1} = 8 \qquad \text{to} \qquad \frac{2}{2+2} \cdot \frac{12}{2} = 3.$$

Let us calculate how much of this drop in demand for good 1 by 5 units is due to the substitution effect, and how much of it is due to the income effect. Thus, we have to calculate $H_1(\hat{p}_1, p_2, \bar{U})$, that is, what would be demand for good 1 if the increase in the price of good 2 were compensated by a corresponding increase in income. We calculated the compensated demand function for good 1 in the lecture notes for Topic 5. It is:

$$H_1(p_1, p_2, \bar{U}) = \left(\frac{p_2}{p_1 + p_2}\right)^2 (\bar{U})^2.$$

We know which prices to plug in, but we don't know yet which utility level $\bar{U}$ to plug in. Let's calculate it. Recall that $\bar{U}$ is the utility level *before* the price increase. Therefore, we just have to plug in the quantities demanded before the price increase into the utility function. We calculated above the quantity demanded of good 1 before the price increase, which was 8 units. But we also have to know the quantity of good 2 demanded before the price increase. We know from Topic 5:

$$D_2(p_1, p_2, y) = \frac{p_1}{p_1 + p_2} \frac{y}{p_2}.$$

Plugging in, we get as the quantity demanded of good 2 before the price increase:

$$\frac{1}{1+2} \cdot \frac{12}{2} = 2.$$

Now we can calculate $\bar{U}$:

$$\bar{U} = \sqrt{8} + \sqrt{2} = \sqrt{4 \cdot 2} + \sqrt{2} = 2\sqrt{2} + \sqrt{2} = 3\sqrt{2}.$$

Now we can calculate the compensated demand that we need to determine the substitution effect:

$$H_1(\hat{p}_1, p_2, \bar{U}) = \left(\frac{p_2}{\hat{p}_1 + p_2}\right)^2 (\bar{U})^2 = \left(\frac{2}{2+2}\right)^2 (3\sqrt{2})^2 = \frac{1}{4} \cdot 18 = 4.5.$$

Thus, if there had been an income compensation after the price increase, demand for good 1 would have dropped from 8 to 4.5. This drop by 3.5 units is the substitution effect. The remaining drop in demand from 4.5 to 3 units, i.e. 1.5 units, is the income effect. With this utility function, for these price and income levels, the substitution effect of the price increase was thus much larger than the income effect.

The Law of Demand

So, what about the law of demand that we mentioned earlier, according to which, supposedly, consumers always reduce demand when the price goes up? I told you before that this is not true. It is not true because of the possibility of Giffen goods. But there is a modified law of demand that is true: if a good is normal, then, if the price of that good goes up, demand for that good goes down. Bear in mind, "normal "here does not mean what it normally means, but a good is normal if, when income goes up, demand for the good goes up. For normal goods, of course, per definition also the reverse holds: if income goes down, demand for that good goes down. For normal goods, if the price of such goods increases, the substitution effect is negative, as always, and the income effect is also negative, because a price increase corresponds to a reduction in real income, and for normal goods the response to a fall in income is a reduction in demand. Thus, substitution effect and income effect are negative, and therefore also their sum is negative.

But, remember, there is the theoretical possibility of a Giffen good. For such goods demand functions look utterly unfamiliar, as shown in Figure 21. In Figure 21, by the way, the demand function curves backwards as prices become very high because even for Giffen goods, demand cannot increase even for arbitrarily high prices, because at some stage even if the consumer spends all her income on good 1, she cannot buy more of it, because she only has limited income.

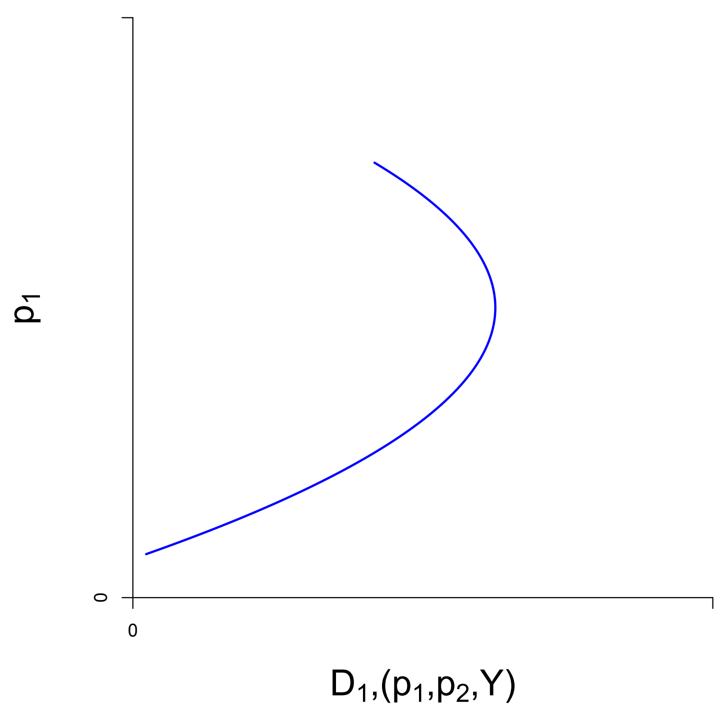


Figure 23: Demand Function for a Giffen Good

Do Giffen Goods Exist in the Real World?

This part of the lecture notes is based on Robert Jensen and Nolan Miller's article "Giffen Behavior and Subsistence Consumption," *The American Economic Review* 2008 (Volume 98), pages 1553-1577.

The fact that the theory of rational consumption does not imply that demand for a good, as a function of its own price, must be downward sloping, but that theoretically it may be upward sloping has been a staple of undergraduate microeconomics teaching for a very long time. The discovery that this theoretical possibility actually does show up in the real world is by tradition attributed to Robert Giffen, a 19th century statistician. There is no book or article by Giffen in which he would make this assertion however. I do not know what the origin of the rumor of Giffen's discovery is. Perhaps it is the following quote from the 1895 edition of British economist Alfred Marshall's book "Principles of Economics."

Although 1895 is a long time ago, and economics has evolved a lot since then, one can still find many influences of Alfred Marshall's book in the teaching of undergraduate economics. I am not convinced that this demonstrates the timeless validity of Alfred Marshall's insights. I believe that more likely it is simply an example of the slowness with which undergraduate teaching in economics has adjusted to the progress that economics research has made.

> As Mr. Giffen has pointed out, a rise in the price of bread makes so large a drain on the resources of the poorer labouring families and raises so much the marginal utility of money to them, that they are forced to curtail their consumption of meat and the more expensive farinaceous foods: and, bread being still the cheapest food which they can get and will take, they consume more, and not less of it.

Note that it is no accident that the good that (supposedly) was a Giffen good is a good that poor people consume more than rich people. As I explained in these notes, this is a necessary feature of Giffen goods. The income effect has to be negative.

The time period in which the upward sloping demand curve for bread was supposed to have been observed by Giffen were the late 1700s. Interestingly, in 1977, the economist Roger Koenker published a paper based on data from this time period that refuted the belief that bread demand increased as its price increased. Other textbook examples have also not stood up to modern empirical investigation. The most frequently cited such example is the demand for potatoes during a famine in Ireland in the years 1845 to 1849, in which about one million people died. But there seems to be no evidence of an upward sloping demand curve for potatoes in this period.

Until recently, the situation was therefore quite curious: Economists' model of consumer behavior could not rule out upward sloping demand curves, yet, in practice, nobody could find any. That suggested that there were regularities in the world not adequately captured by the model. This changed when in 2008 two economists, Jensen and Miller, published an article in the *American*

Economic Review called "Giffen Behavior and Subsistence Consumption," in which they used data that they had collected in a field experiment in China to demonstrate the existence of Giffen goods. They selected randomly households from the relatively poor part of the population, and lowered artificially for these households the price of their primary dietary staple by giving them vouchers for the purchase of these staples. Jensen and Miller focused on two provinces of China: Hunan in the south, where rice is the staple good, and Gansu in the north, where wheat is the staple. They found that poor households in Hunan exhibited Giffen behavior with respect to rice. That is, lowering the price of rice via the experimental subsidy caused households to reduce their demand for rice, and removing the subsidy had the opposite effect. In Gansu, the evidence for Giffen behavior is somewhat weaker. They attribute the finding in Gansu, among other things, to the fact that the households in their sample are not so poor that they consume only staple foods.

Topic 7: Elasticities and the Slutsky Equation

Substitution and Income Effects Expressed in Derivatives

A famous equation of demand theory, *the Slutsky equation*, is a re-writing of the equation: "total effect = substitution effect + income effect" in terms of derivatives of the demand function and the compensated demand function. An interesting implication of this equation will be that, if we observe, approximately, a consumer's demand function, then we can infer from this properties, specifically the derivative, of the compensated demand function, a function that we don't observe in practice, and that is just a hypothetical construction. In particular, we can then check whether it has the properties that are predicted by utility maximization, such as whether it has a negative derivative. That is, indirectly, testing the rationality of the consumer.

Suppose the price of good 1 increases from p_1 to $\hat{p}_1$. In the notes for Topic 6, we explained that the definition of income effects and substitution effects can be summarized by the following equation:

Version 1:

$$\underbrace{D_1(\hat{p}_1, p_2, y) - D_1(p_1, p_2, y)}_{\text{total effect}} = \underbrace{H_1(\hat{p}_1, p_2, \bar{U}) - D_1(p_1, p_2, y)}_{\text{substitution effect}} + \underbrace{\left(D_1(\hat{p}_1, p_2, y) - H_1(\hat{p}_1, p_2, \bar{U}\right)}_{\text{income effect}}.$$

We call this form of the equation "Version 1." We shall proceed in several steps, and at the end of all these steps, we arrive at "Version 5." That version of the equation will only involve derivatives. We shall put the derivation of the various versions of the equation into small, cursive letters. It is worth your while to read these derivations, because they will deepen your understanding of demand theory, but on first reading you can also skip these derivations.

Before we can introduce derivatives, we need to re-write the right hand side of Version 1 a little bit. This is because derivatives will be substituted for differences, and this can only be done if the differences involve one and the *same function. But in Version 1 income and substitution effect both involve differences of different functions, namely demand and compensated demand functions.*

We shall use an equality that we introduced in Topic 5, and that indicates a connection between the utility maximization problem and the expenditure minimization problem. The equality was:

$$D_1(p_1, p_2, y) = H_1(p_1, p_2, \bar{U}),$$

where:

$$y = E(p_1, p_2, \bar{U}).$$

Recall from Topic 5 that we could see this geometrically. The diagram showing a solution to the utility maximization problem and the diagram showing a solution to the expenditure minimization problem are identical if the fixed utility in the expenditure minimization problem is the maximized utility in the utility maximization problem.

The connection between the two problems now allows us to replace in Version 1 for income and substitution effect "$D_1(p_1, p_2, y)$" by "$H_1(p_1, p_2, \bar{U})$" where $y = E(p_1, p_2, \bar{U})$. Conversely, we can also replace "$H_1(\hat{p}_1, p_2, \bar{U})$" by "$D_1(\hat{p}_1, p_2, \hat{y})$," where $\hat{y} = E(\hat{p}_1, p_2, \bar{U})$. Note that we have to refer to two different incomes: y is the actual income of the consumer. The symbol $\hat{y}$ stands for the actual income of the consumer plus the compensation that allows her to return to the original utility level after the price increase.

We get:

Version 2:

$$\underbrace{D_1(\hat{p}_1, p_2, y) - D_1(p_1, p_2, y)}_{\text{total effect}} = \underbrace{H_1(\hat{p}_1, p_2, \bar{U}) - H_1(p_1, p_2, \bar{U})}_{\text{substitution effect}} + \underbrace{(D_1(\hat{p}_1, p_2, y) - D_1(\hat{p}_1, p_2, \hat{y}))}_{\text{income effect}}.$$

Now let us think of very small changes in prices, so that the terms all become derivatives. Starting on the left hand side, this term will become the partial derivative of D_1 with respect to p_1:

$$\underbrace{\frac{\partial D_1(p_1, p_2, y)}{\partial p_1}}_{\textit{total effect}}$$

The first term on the right hand side will be the partial derivative of H_1 with respect to p_1:

$$\underbrace{\frac{\partial H_1(p_1, p_2, \bar{U})}{\partial p_1}}_{\textit{substitution effect}}$$

The second term will become the negative of a partial derivative. This is because we are subtracting demand at one income from demand at a higher *income. But the negative of which partial derivative? We are tempted to answer (where we already include the minus sign):*

$$\underbrace{-\frac{\partial D_1(\hat{p}_1, p_2, y)}{\partial y}}_{\textit{not really the income effect}}$$

But there is a little problem here. We are trying to trace the effect of an increase of p_1, not of an increase in y. The change in y that we are considering in this term is the change that is triggered by the compensation offered to the consumer if the price p_1 increases. Thus, we have to multiply the above partial derivative by the increase in y triggered by an increase in p_1:

Formally, we are applying here the chain rule.

$$\underbrace{-\frac{\partial D_1(\hat{p}_1, p_2, y)}{\partial y} \cdot \frac{\partial y}{\partial p_1}}_{\textit{income effect}}$$

We now get:

Version 3:

$$\underbrace{\frac{\partial D_1(p_1, p_2, y)}{\partial p_1}}_{\text{total effect}} = \underbrace{\frac{\partial H_1(p_1, p_2, \bar{U})}{\partial p_1}}_{\text{substitution effect}} + \left(\underbrace{-\frac{\partial D_1(\hat{p}_1, p_2, y)}{\partial y} \cdot \frac{\partial y}{\partial p_1}}_{\text{income effect}} \right).$$

But what is $\frac{\partial y}{\partial p_1}$? That is, by how much do we have to raise income if the price of good p_1 goes up by one (infinitesimally small) dollar, but the consumer wants to maintain her utility level? If the consumer did not adjust her demand at all to the price increase, then we would have to pay her for every unit of good 1 that she bought before the price increase of one Dollar, that is, she would need to get $D_1(p_1, p_2, y)$ Dollars. But, of course, she should rearrange her demand *after the price increase. Here comes a miraculous fact: for calculations with infinitesimally small Dollars, that is, for the calculus expression of income and substitution effect, this rearrangement of demand does not matter, i.e. we have:*

$$\frac{\partial y}{\partial p_1} = D_1(\hat{p}_1, p_2, y).$$

We shall not prove this here. The equation is a special case of a result that is known among economists as the "envelope theorem," presumably because one can prove it on the back of an envelope. But we won't try that trick here.

Also, nobody any longer uses envelopes. We send emails, not letters.

Plugging our last equation into the equation before, we obtain for the income effect:

$$\underbrace{-\frac{\partial D_1(\hat{p}_1, p_2, y)}{\partial y} \cdot D_1(\hat{p}_1, p_2, y)}_{\textit{income effect}}$$

Now we can put together the equation: "total effect = substitution effect + income effect," using all the derivative expressions that we have found.

Version 4:

$$\underbrace{\frac{\partial D_1(p_1, p_2, y)}{\partial p_1}}_{\text{total effect}} = \underbrace{\frac{\partial H_1(p_1, p_2, \bar{U})}{\partial p_1}}_{\text{substitution effect}} + \underbrace{\left(-\frac{\partial D_1(\hat{p}_1, p_2, y)}{\partial y} \cdot D_1(\hat{p}_1, p_2, y)\right)}_{\text{income effect}}.$$

Note that on the right hand side, in the first factor of the expression for the income effect, we have :"$\hat{p}_1$," and not: "p_1". But for infinitesimally small changes this difference is not important. We might as well write our equation as:

Version 5:

$$\underbrace{\frac{\partial D_1(p_1, p_2, y)}{\partial p_1}}_{\text{total effect}} = \underbrace{\frac{\partial H_1(p_1, p_2, \bar{U})}{\partial p_1}}_{\text{substitution effect}} + \underbrace{\left(-\frac{\partial D_1(p_1, p_2, y)}{\partial y} \cdot D_1(p_1, p_2, y)\right)}_{\text{income effect}},$$

This is the bottom line of this section, and our first version of the Slutsky equation.

Note that the equation with which we started, and which we took from the previous topic, is really just a definition. It is tautologically true. The Slutsky equation is not quite a tautology. We used the connection between the utility maximization problem and the expenditure minimization problem, and we used the envelope theorem. So, the equation with which we ended up is a little achievement, and not just a definition. It seems fair that the equation is named after someone, namely after Eugen Slutsky. The equation is an assertion about demand function and the compensated demand functions. In examples, we can check whether we have derived these functions correctly by checking whether the Slutsky equation holds. We could do this for our example in Topic 5, but we shall save us the algebra here.

Our earlier claim that the substitution effect of a price increase is always negative, means in the Slutsky equation that the first expression is negative:

$$\frac{\partial H_1(p_1, p_2, U)}{\partial p_1} < 0.$$

This is indeed true, although we omit the proof here. The second assertion, namely that the income effect may be positive or negative, translates now into the assertion:

$$-\frac{\partial D_1(p_1, p_2, y)}{\partial y} \quad \text{may be positive or negative,}$$

because the factor by which this gets multiplied in the Slutsky equation, $D_1(p_1, p_2, y)$, is always positive. Of course, our discussion in Topic 6 of normal and inferior goods shows that $\frac{\partial D_1(p_1,p_2,y)}{\partial y}$ may be positive or negative.

Elasticities

Economists like to replace the concept of a derivative, as we have used it in the previous section, by the concept of elasticities. The motivation is simple: Suppose we study the derivative of the regular demand function with respect to a good's own price. The derivative gives us the answer to the following question: By how many units of the good does demand change if we raise the price of the good by one monetary unit? Obviously, the answer will depend on the monetary unit with which we calculate, and also on the unit of measurement for the good. In other words, the number that answers the question: "How many more grams of sugar do I get if I spend one more Cent?" is different from the number that answers the question: "How many more pounds of sugar do I get if I spend one more cent?"

It would be nice to characterize the reaction of demand to a price increase by a number that is independent of units of measurement. We can get such a number if we ask ourselves: "By how many percents does the quantity demanded change if we raise the price by one percent?" The size of the percentage increase in quantity is always the same, regardless of which units we express the quantity in. For example: An increase from 2 feet to 2.5 feet of lumber is an increase by 25%. The same increase can also be expressed as an increase from 60.96 centimeters to 76.2 centimeters, but that is still an increase by 25%. The same applies to prices.

How can we calculate elasticities? We simplify our notation a little bit by considering briefly some mathematics. Suppose y is a function of x: $y = f(x)$. and we want to know by how much does y change as we increase x by some amount $\varepsilon > 0$. The answer is obviously:

$$f(x+\varepsilon) - f(x).$$

Now suppose we wanted to do the same calculation in percentages. If we increase the argument of f from x to $x+\varepsilon$, by how many percentages have we raised x? The answer is:

$$\left(100 \cdot \frac{\varepsilon}{x}\right)\%.$$

By how many percentages has f changed? The answer is:

$$\left(100 \cdot \frac{f(x+\varepsilon) - f(x)}{f(x)}\right)\%.$$

Let us divide the percentage change in f by the percentage change in x, to obtain the percentage change in f for each percentage by which we raise x:

$$\frac{\left(100 \cdot \frac{f(x+\varepsilon)-f(x)}{f(x)}\right)\%}{\left(100 \cdot \frac{\varepsilon}{x}\right)\%}.$$

It seems fair that we are allowed to remove the percentage sign in numerator and denominator, and to cancel out the factor 100.

$$\frac{\frac{f(x+\varepsilon)-f(x)}{f(x)}}{\frac{\varepsilon}{x}}.$$

Now we want to pass from this expression to derivatives. For this, we use a little bit of algebra to re-write the last expression as:

$$\frac{\frac{f(x+\varepsilon)-f(x)}{f(x)}}{\frac{\varepsilon}{x}} = \frac{f(x+\varepsilon)-f(x)}{f(x)} \cdot \frac{x}{\varepsilon} = \frac{f(x+\varepsilon)-f(x)}{\varepsilon} \cdot \frac{x}{f(x)}.$$

As always with derivatives we are interested in the case that ε tends to zero. The first factor in the product on the right hand side of the above equation tends to the derivative $f'(x)$ as ε tends to zero. This is, in fact, the definition of a derivative. The second factor does not depend on ε. Thus, we obtain:

$$f'(x) \cdot \frac{x}{f(x)}.$$

This is what economists call the elasticity of y with respect to x.

In the setting of demand theory, we might, for example, be interested in the elasticity of demand for a good with respect to its own price. To find this, we use the above formula, replacing the function f by the demand function D_1, and we replace the derivative f' by the partial derivative $\partial D_1/\partial p_1$, and we get the price elasticity of demand:

$$\frac{\partial D_1(p_1, p_2, y)}{\partial p_1} \cdot \frac{p_1}{D_1(p_1, p_2, y)}.$$

Analogously, we can calculate the income elasticity of demand:

$$\frac{\partial D_1(p_1, p_2, y)}{\partial y} \cdot \frac{y}{D_1(p_1, p_2, y)}.$$

We can also define the elasticity of compensated demand with respect to prices, or with respect to utility. Indeed, we can, and sometimes will, define the elasticity of almost everything with respect to almost everything.

Income and Substitution Effect Expressed in Elasticities

Let us use the concept of elasticities to re-write the Slutsky equation. The motivation for this is simply that economists like to think in terms of elasticities, for the reasons explained in the previous section. Let us pick up where we left off:

Version 5:

$$\underbrace{\frac{\partial D_1(p_1,p_2,y)}{\partial p_1}}_{\text{total effect}} = \underbrace{\frac{\partial H_1(p_1,p_2,U)}{\partial p_1}}_{\text{substitution effect}} + \underbrace{\left(-\frac{\partial D_1(p_1,p_2,y)}{\partial y}\cdot D_1(p_1,p_2,y)\right)}_{\text{income effect}}.$$

In the following paragraphs we transform this equation into an equation about elasticities. I shall again set the details of this derivation into small, cursive letters, so that you can skip them on first reading. This derivation is not too insightful, but if you are the sort of person who does not automatically trust what is in the books, but wants to derive it herself or himself, then you should go through this derivation.

We would like to have an elasticity on the left hand side rather than a derivative. To achieve this, we have to multiply the left hand side by $\frac{p_1}{D_1(p_1,p_2,y)}$. *But if we multiply the left hand side by this expression, we also have to multiply the right hand side by this expression. We get on the left hand side as the "total effect" the own price elasticity of demand:*

$$\underbrace{\frac{\partial D_1(p_1,p_2,y)}{\partial p_1}\cdot\frac{p_1}{D_1(p_1,p_2,y)}}_{\text{total effect}}.$$

On the right hand side, we get first the substitution effect, which, after multiplication, now becomes:

$$\underbrace{\frac{\partial H_1(p_1,p_2,U)}{\partial p_1}\cdot\frac{p_1}{D_1(p_1,p_2,y)}}_{\text{substitution effect}}$$

Now recall the equation: $D_1(p_1,p_2,y) = H_1(p_1,p_2,U)$*, which describes the connection between utility maximization and expenditure minimization. Substituting, we can therefore re-write the substitution effect as:*

$$\underbrace{\frac{\partial H_1(p_1,p_2,U)}{\partial p_1}\cdot\frac{p_1}{H_1(p_1,p_2,U)}}_{\text{substitution effect}},$$

which is the own price elasticity of compensated demand.

Finally, we multiply the income effect by the same factor as we have multiplied *all other terms, and we obtain:*

$$\underbrace{-\frac{\partial D_1(p_1,p_2,y)}{\partial y}\cdot D_1(p_1,p_2,y)\cdot\frac{p_1}{D_1(p_1,p_2,y)}}_{\text{income effect}}.$$

Cancelling out $D_1(p_1,p_2,y)$, this becomes:

$$\underbrace{-\frac{\partial D_1(p_1,p_2,y)}{\partial y}\cdot p_1}_{\text{income effect}}.$$

Unlike the other two expressions that we obtained, this does not look like any elasticity! That is disappointing. Most naturally, this should involve the income elasticity of demand. Let's simply make that happen. We are going to insert after the derivative $\frac{\partial D_1(p_1,p_2,y)}{\partial y}$ the factor $\frac{y}{D_1(p_1,p_2,y)}$. That is exactly the factor that we need to obtain the income elasticity. Unfortunately, after multiplying by this factor, we have to divide again by it, so as to not change the value of the expression. Dividing by a fraction is the same as multiplying by its inverse. So, we shall put at the end of the expression the factor: $\frac{D_1(p_1,p_2,y)}{y}$. We get:

$$\underbrace{-\frac{\partial D_1(p_1,p_2,y)}{\partial y}\cdot\frac{y}{D_1(p_1,p_2,y)}\cdot p_1\cdot\frac{D_1(p_1,p_2,y)}{y}}_{\text{income effect}}.$$

This looks ugly, but will soon become simple. Let us pull the factor p_1 into the numerator of the last fraction:

$$\underbrace{-\frac{\partial D_1(p_1,p_2,y)}{\partial y}\cdot\frac{y}{D_1(p_1,p_2,y)}\cdot\frac{p_1 D_1(p_1,p_2,y)}{y}}_{\text{income effect}}.$$

The first two factors together are the income elasticity of demand. The third factor is the expenditure on good 1 divided by income, that is, it is the proportion of income that the consumer spends on good 1.

Before we can write down the Slutsky equation in elasticities, we introduce some notation. Let us denote by:

- ε (read: "epsilon") the price elasticity of demand,
- ε^* (read: "epsilon star") the price elasticity of compensated demand,
- ξ (read: "ksi") the income elasticity of demand,
- θ (read: "theta") the proportion of income that the consumer spends on good 1.

Then we can write our equation in short as:

Version 6:

$$\underbrace{\varepsilon}_{\text{total effect}} = \underbrace{\varepsilon^*}_{\text{substitution effect}} + \underbrace{-\xi \cdot \theta}_{\text{income effect}} .$$

This is the Slutsky equation written in elasticities. Let us say this in words: the price elasticity of regular demand equals the price elasticity of compensated demand minus income elasticity of regular demand, where the last term is multiplied by a weight, namely the proportion of income spent on the good that we are considering.

There is some economic insight that we can derive from this. The last term, the income effect, will be small, i.e. it won't matter much, when we consider goods on which we spend only a small share of our income. For such goods, mostly the substitution effect matters, not the income effect, and therefore the total effect of a price increase is likely to be negative. For goods on which we spend a very large share of our income, the income effect matters, and because this effect can be positive or negative, the total effect of a price increase for such goods may sometimes be positive rather than negative.

Topic 8: Complements and Substitutes

The Effect of Other Goods' Price Changes on Demand

In Topic 6, we have considered the effect of income changes on the demand for some particular good, say good 1, and we have considered the effect of changes of the good's own price, p_1 on the demand for good 1. It remains to think about changes in the price of some other good. For concreteness, let us think about the effect of an increase in the price of good 2 on the demand for good 1. Such an increase will cause the budget line to rotate inwards, towards the origin, around the intersection point of the budget line with the horizontal axis, that is, the axis for good 1. This intersection point remains the same because the maximum amount of good 1 that the consumer can purchase doesn't change as the price of good 2 changes. The rotation occurs because an increase in the price of good 2 makes the budget line, whose slope is $-p_1 / p_2$, flatter. In Figure 22 on the next page, the blue budget line indicates the budget line before the increase in the price of good 2, and the green budget line indicates the budget line after the increase in the price of good 2.

As in Topic 6, when we considered the effect of an increase in p_1 on demand for good 1, we can also decompose the effect of an increase in the price of good 2 on the demand for good 1 into two parts, the income effect and the substitution effect. The sign of the effect of the substitution effect is clear: as p_2 increases, good 1 becomes relatively cheaper, and therefore the demand for good 1 increases. Thus, the substitution effect of an increase in the price of good 2 on demand for good 1 has the opposite sign, namely a "+" sign, of the substitution effect of an increase in the price of good 1 on the demand for good 1, which has a "-" sign.

The income effect of an increase of the price of good 2 is the effect of a reduction in real income. For a normal good, this effect will be negative. For an inferior good, it will be positive. Thus, if good 1 is an inferior good, then an increase in the price of good 2 will unambiguously increase demand for good 1. But if good 1 is a normal good, then the income and substitution effects of an increase in the price of good 2 work into opposite directions, and demand for good 1 may increase or decrease.

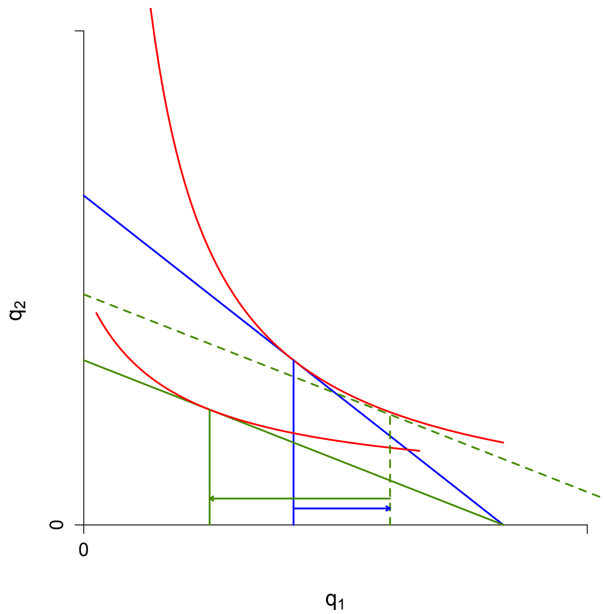


Figure 24: Income and Substitution Effect When the Price of Good 2 Increases: Complements

I illustrate two possibilities in Figures 22 and 23. In Figure 22 demand for good 1 goes down as the price of good 2 goes up because the income effect is stronger than the substitution effect. In Figure 23 the demand for good 1 goes up as the price of good 2 goes up because the income effect is smaller than the substitution effect. If, as in Figure 22, the demand for some good goes down as the price of the other good increases, then we say that the first good is a complement to the other good, and if it is true that both goods are complements to each other, then we simply call them complements. The idea is that these are two goods that the consumer typically wants to consume together: such as, for us Germans, coffee and cake. Thus, if coffee's price goes up, you reduce not only the demand for coffee, but also for cake, because the cake is less enjoyable if there is no coffee to go with it. Cake is a complement to coffee.

One more small observation on complements: Note that we define separately what it means that coffee is a complement to cake, and that cake is a complement to coffee, and only if both are complements for each other do we call them simply "complements." This is a subtle point. It may in theory be that one good is a complement for another, but not vice versa. But let's not worry about this point here.

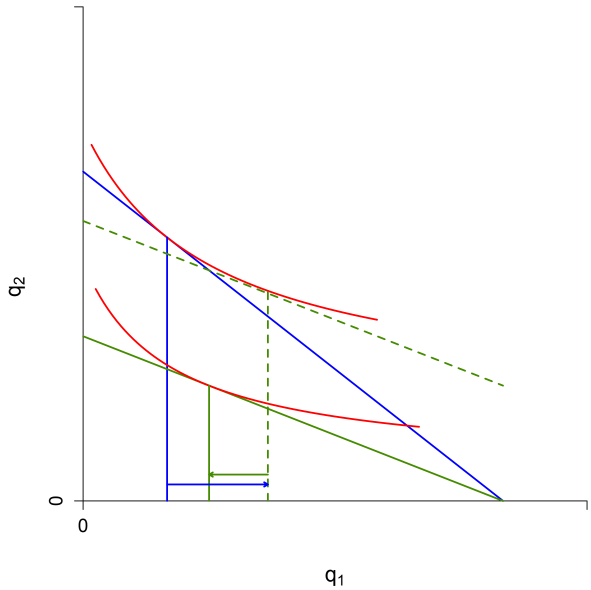


Figure 25: Income and Substitution Effect When the Price of Good 2 Increases: Substitutes

If, as in Figure 23, the demand for some good goes up as the price of the other good increases, then we say that the former good is a substitute for the other, and if it is true that both goods are substitutes for each other, then we simply call the goods substitutes. For example cheese cake may serve as a substitute for an apple tart. You may just wish to eat something sweet, and as the apple tart becomes too expensive, you might just eat the more heavy cheese cake. Thus, cheese cake is a substitute for apple tart.

The observation concerning complements that I mentioned above is also true for substitutes: one good may be a substitute for another, but not the other way round. But again, we shall ignore this possibility here.

The concepts of complements and substitutes as I have defined them here are often also referred to as "gross complements," and "gross substitutes." There are also concepts of "net complements" and "net substitutes" in the economics literature. This is a subtle distinction that we shall ignore in this class. Indeed, with two goods we can't illustrate the difference very well. This is a drawback of our restriction to the case of two goods.

Perfect Complements

In economics we also sometimes call two goods perfect complements or perfect substitutes. They are extreme cases of complements and substitutes. The extreme cases that we study here are sometimes realistic. We are also studying them because considering these extreme cases sometimes helps us to develop some economic intuition.

Whereas the definitions of complements and substitutes refer to properties of the demand function, the definitions of perfect complements and perfect substitutes refer to properties of preferences. But we shall calculate for preferences that correspond to perfect complements the demand function, and then we shall check that this demand function has the properties that define complements. And we shall do the same for substitutes.

Sometimes, two consumption goods need to be consumed in some particular proportions. For example, left shoes, and right shoes, really, for most people, must be consumed in the ratio 1:1. If you have two left shoes, and only one right show, you really don't derive any particular benefit from the additional left shoe, unless, of course, you lack a leg. We say in cases such as this that the two goods "complement" each other, indeed in the case of left and right shoes we say that they are perfect complements. Other goods come close to being perfect complements. The ratio is not always 1:1. It could be 1:2, for example, if you want to buy a bicycle frame, and wheels for the bicycle.

What do indifference curves look like when you need to consume goods in a particular proportion to each other? Let's consider the simplest case: left shoes and right shoes, and let's think about the indifference curve that passes through the point (1,1), i.e. the consumption bundle that consists of 1 left shoe and one right shoe. If I give you one more left shoe, that doesn't benefit you at all, and therefore, you will be indifferent between having one left and one right shoe, and having two left shoes, but still only one right shoe. In other words, the point (2,1) is on the same indifference curve as the point (1,1). In fact, any point $(n, 1)$ with $n > 1$ will be on the same indifference curve as (1,1). Of course, we have to have $n > 1$, because you are not indifferent if I take a left shoe away from you. By the same argument, also all points $(1, n)$ are on the same indifference curve as (1,1). In Figure 24, I show the points on the indifference curve that we have discussed so far.

I have cheated a little bit by drawing not just the points we discussed above,

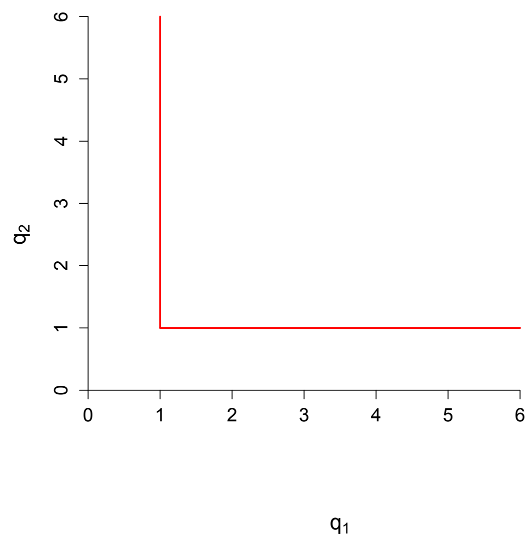


Figure 26: Indifference Curves for Left and Right Shoes

but also connecting those points by straight lines. This yields the rectangle in Figure 24. We shall, actually, pretend in this section that quantities are continuous variables, although that doesn't make much sense with shoes. But there are some complications when quantities have to be integers that we don't want to deal with.

The points that we see in Figure 24 are actually *all* points on the indifference curve through (1,1). There are no other points, because all other points give the consumer either more of both goods, or less of both goods. Moreover, all other indifference curves have the same shape as the indifference curve in Figure 24, with the corner lying on the 45° line. This is because on the 45° line the ratio of the quantities of the two goods is exactly 1:1, which is what the consumer wants. If all indifference curves look as in Figure 24, then we say that the goods are perfect complements.

Can we think of a utility function that has indifference curves as in Figure 24? Here is one:

$$U(q_1, q_2) = \min\{q_1, q_2\}$$

That is, the utility from any bundle (q_1, q_2) is the smaller of the two numbers q_1 and q_2. If these numbers are same, then utility is equal to that number. So, for example, the utility from $(1, 2)$ is 1. Thus, this utility function represents that the smaller of the two numbers determines how many pairs of shoes you have, and that is your utility. This function has indifference curves as in Figure 24. What are the marginal rates of substitution for this function? Well, the slope of the indifference curve is either $-\infty$ (negative infinity), when the indifference curve is vertical, or it is 0, when it is horizontal. In the corner, the slope is not well-defined. Mathematically speaking, in that point the indifference curve is not "differentiable." We defined the marginal rate of substitution to be the negative of the slope of the indifference curve, and thus in the case of perfect complements it is either ∞, zero, or not-defined. You can see that the marginal rate of substitution cannot equal the price ratio in any case, and therefore it will not be very useful in determining the optimal consumption bundle below.

We mentioned before that the ratio at which the consumer wants to consume complements need not be 1:1. For example, the consumer might wish to consume 4 units of good 2 for every unit of good 1. Then all indifference curves are rectangles, but with the bottom left corner on the line that passes through (0,0) and through, say, the point (1,4). That is, the corners are on a straight line that is much steeper than the 45° line. I show this case in Figure 25.

For this example, a utility function that can represent these preferences is:

$$U(q_1, q_2) = \min\{q_1, \frac{q_2}{4}\}.$$

For example, when the consumer has 2 units of good 1 and 9 units of good 2, then the utility is 2: she can consumer 2 units of good 1 and 8 units of good 2, and the remaining unit of good 2 is waste.

What does the demand function for perfect complements look like? Let's consider as an example the last case: the consumer wants to consume the goods in a ratio 1:4. Suppose the price of good 1 is $p_1 = 5$, the price of good 2 is $p_2 = 2$, and the income is $y = 39$. I won't do any calculation to "prove" what the demand function is. I think it is obvious once I have explained the logic that leads to the demand function in words. The consumer will want to buy the goods only exactly in the ratio in which she needs them. Any other purchase would waste units of one of the two goods. So, if she buys 1 unit of good 1, she will buy 4 units of good 2. The price of that bundle is: $5 + 4 \cdot 2 = 13$.

Hopefully, I am right in thinking that.

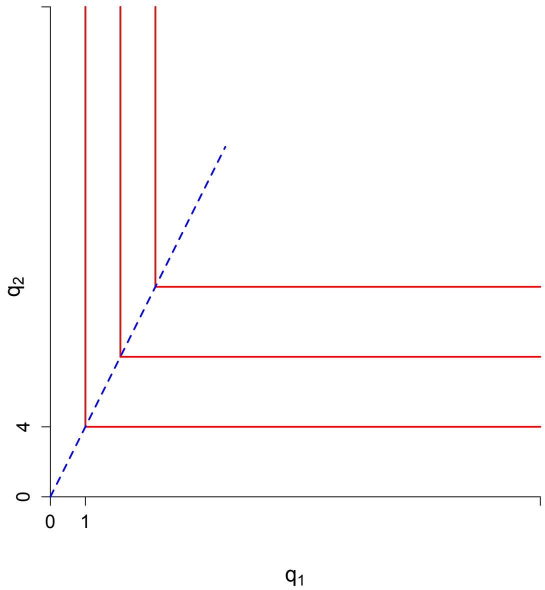


Figure 27: Indifference Curves for Cars (q_1) and Tires (q_2)

Thus, the consumer will think of this as if it was just a single good at price 13. The number of units that she can buy is therefore:

$$\frac{39}{5+2\cdot 4} = 3.$$

She will buy 3 units of good 1 and $3 \cdot 4 = 12$ units of good 2. In general, she will buy:

$$D_1(p_1, p_2, y) = \frac{y}{p_1 + 4p_2} \text{ units of good 1 and } D_2(p_1, p_2, y) = 4\frac{y}{p_1 + 4p_2} \text{ units of good 2.}$$

These are the demand functions. You can probably see how this generalizes if the ratio is $1 : n$ for some $n > 0$ rather than $1 : 4$. Note that demand for each good decreases as the price of the other good goes up. Therefore, perfect complements are indeed a special case of substitutes.

What is the substitution effect on the demand for good 1 as the price of good 1 increases? As you would expect for perfect complements, the substitution effect is zero. Whenever you want to achieve some given utility level, say 3, the expenditure minimizing consumption bundle is always the same: it is 3 units of good 1, and $3n$ units of good 2. This does not depend on the prices. The expenditure minimizing bundle is always the bottom left corner of the indifference curve. Thus, when you are compensated to maintain your utility level, you keep consuming the same quantities. As the substitution effect is zero, the income effect must equal the total effect of a price increase.

Perfect Substitutes

Now we treat the opposite case. Suppose two goods have different brand names but are exactly the same, say Starbucks coffee, and Zingerman's coffee. (This example is chosen to offend at least one of these two companies.) What will indifference curves look like? Let us think of the indifference curve through the point (3,3), which corresponds to 3 cups of Starbucks coffee, and 3 cups of Zingerman's coffee. On the same indifference curve will also be the points (0,6), (5,1), (2,4), etc., up to (6,0). This is because all these consumption bundles offer the same number of, essentially identical, coffee cups. Therefore, the indifference curve will be a straight line through the point (3,3) with slope -1. All other indifference curves look the same. I show the indifference curves in Figure 26. When indifference curves have this shape, we call the two goods perfect substitutes.

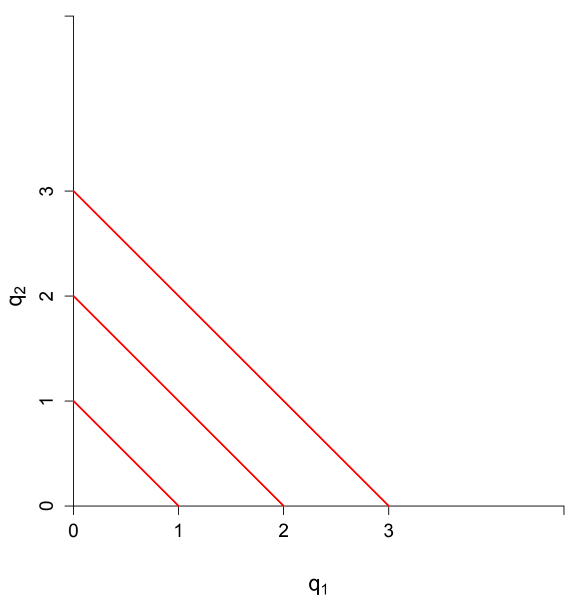


Figure 28: Indifference Curves for Two Types of Coffee That Taste the Same

Which utility function could represent these preferences? This one:

$$U(q_1, q_2) = q_1 + q_2.$$

We actually talked about this utility function and its indifference curves in an earlier Topic.

Suppose coffee cups at Zingerman's were half as large as coffee cups at Starbucks. Then two coffee cups from Zingerman's would be a perfect substitute for 1 coffee cup from Starbucks. What would be a utility function for this case?

$$U(q_1, q_2) = q_1 + \frac{q_2}{2}.$$

That is, if I offer you 2 cups of coffee from Starbucks, and another 6 from Zingerman's, we just divide those from Zingerman's by 2, so that everything is expressed in Starbucks coffee cups, and then add them up. Your utility is:

$$2 + \frac{6}{2} = 5.$$

The indifference curves in this case are parallel straight lines, but with a slope of -2, thus steeper than the indifference curves in Figure 26. We can do the same exercise replacing the number 2 by any other number n, and get steeper or flatter indifference curves than in Figure 26.

What is the demand function for perfect substitutes? Sticking with the last example, the consumer should compare prices: if Starbucks price is less than twice as much as Zingerman's, she should buy from Starbucks, otherwise from Zingerman's. Only if the price at Starbucks is exactly equal to twice the price from Zingerman's, she is exactly indifferent. We can write this demand function in the following form:

$$D_1(p_1, p_2, y) = \begin{cases} \frac{y}{p_1} & \text{if } p_1 < 2p_2 \\ \text{indifferent} & \text{if } p_1 = 2p_2 \\ 0 & \text{if } p_1 > 2p_2 \end{cases}$$

$$D_2(p_1, p_2, y) = \begin{cases} 0 & \text{if } p_1 < 2p_2 \\ \text{indifferent} & \text{if } p_1 = 2p_2 \\ \frac{y}{p_2} & \text{if } p_1 > 2p_2 \end{cases}$$

Note that, as the price of one good goes up, the demand for the other good either remains the same, or, if we cross a threshold, it suddenly jumps. For example, consider the demand for good 1. if p_2 is low, that is $2p_2 < p_1$, or, equivalently: $p_2 < p_1/2$, then demand for good 1 is zero. When p_2 crosses the threshold: $2p_2 = p_1$, or, equivalently: $p_2 = p_1/2$, then demand for good 1 jumps from 0 to y/p_1. And finally, when p_2 is high, that is, $2p_2 > p_1$, or, equivalently, $p_2 > p_1/2$, then demand for good 1 is y/p_1, and thus does not change as p_2 rises further.

Recall that in the previous section we defined substitutes to be goods such that demand for one good rises as the price of the other good increases. Is that true for perfect substitutes? In a somewhat degenerate sense it is. As we saw in the previous paragraph, there is one jump "upward" in the demand for good 1 when p_2 increases, although, aside from this jump, demand for good 1 does not change as the price of good 2 changes. This somewhat degenerate sense in which demand for good 1 rises as the price of good 2 increases is enough: perfect substitutes are a special case of substitutes.

Note that, if it was n rather than 2 cups of Zingerman's coffee that can substitute for one cup of Starbucks coffee, then in our calculations we would just have to replace the "2" by "n." All these cases are still referred to as cases of perfect substitutes. They are also special cases of substitutes.

Let's briefly look at changes in demand as a good's own price changes. Suppose a good's own price increases, say good 1. Either, we are in the region where demand is y/p_1, in which case demand drops continuously as p_1 increases, or we cross the threshold, in which case demand drops discontinuously to zero, or we are in the region where demand is anyway zero for good 1, and therefore good 1's demand doesn't change. In the former case, the effect of the price increase is a pure income effect, when we cross the threshold, the effect of the price increase is a pure substitution effect, and if demand remains zero, both income and substitution effects are zero. Can you find the explanation for this yourself?

Topic 9: Rational Choice of Labor Supply and Savings

In this section, we shall show how the simple two good model of consumption that we have studied so far can also be used to model consumers' choices of how much to work, i.e. their labor supply, and how much to save. We shall do this, though, in two separate models, one for each topic. This is because if we wanted to study labor supply and savings in the same model, and also include at least one consumption good, we would have to draw three-dimensional pictures, and we aren't good at that. By separating the two subjects, and studying them in separate models, we continue to use essentially the same framework as in the previous topics. We shall reinterpret the model of the previous sections, but the basic ideas and intuitions remain the same.

Labor Supply

So, let us begin with labor supply. We shall construct a model in which working only has drawbacks, no advantages. That is, hopefully, not how everyone experiences their work life. But it is our starting point. On the other hand, there will also not be any pain that consumers experience at work in our model. Rather, the only reason people do not like to work will be that they could, instead, spend the time on the beach, or watching movies, i.e., they could spend their time as leisure. Now, of course, the reason why people want to spend time as leisure could be that, in fact, work is also pleasant, but leisure is more pleasant, or work might be painful, whereas leisure is not painful. Be that as it may, in our model we only describe the opportunity cost of work, that is, leisure.

My daughter, when she was very small, once said to me that she did not want to grow up, because grownups' lives seemed to be just "work, work, work." I must have looked unhappy at the time with how things were going at work for me. But sometimes my work is enjoyable.

There will also be a reason why people work: it is to earn a wage. Leisure seems to many people pointless unless it offers an opportunity of consumption. People earn a wage so that they can afford the consumption good which makes their leisure time attractive. So, we shall also include a consumption good in our model. This turns it into a two-good model: the two goods that people will derive utility from are leisure, and consumption.

Let us denote in this Topic leisure by $\ell \geq 0$, and consumption by $q \geq 0$. People will have a utility function $U(\ell, q)$. Of course, only ordinal preferences matter. We shall assume that the utility function is increasing in both arguments. Sometimes, we shall draw pictures in which the indifference curves are convex. Does this make sense? It seems somewhat plausible: you might prefer having a mixture of leisure and consumption over having too much leisure, or too much consumption.

What is the budget constraint? Let us denote the wage by w, and the price of consumption by p. There is no given income y. It depends on how much people work. But so far we only have explicit notation for leisure. Let's also introduce a symbol for how much people work: L.You must love my notation. Two ells. This is because both labor and leisure are words that begin with the letter ell. The two ells are connected: suppose there are 24 hours in the day, and you either spend them on labor or on leisure (sleep counts as leisure, eating counts as leisure, ...). Then the two ells have to satisfy:

$$L + \ell = 24.$$

We can solve this equation for labor: $L = 24 - \ell$. Thus, if the consumer chooses to have ℓ hours of leisure, then we also know that she will choose to work for $24 - \ell$ hours, thereby earning an income of $w(24 - \ell)$.

We are now ready to write the budget constraint:

$$pq \leq w(24 - \ell).$$

This says that what the consumer spends on consumption cannot be more than she earns from labor. Let us re-write it a little bit:

$$pq + w\ell \leq 24w.$$

The reason I re-wrote the budget constraint in this way is that now on the left hand side I have a sum that looks a lot like: $p_1 q_1 + p_2 q_2$, which we had on the left hand side of the budget constraint in previous topics. Now we have on the left hand side what the consumer spends on consumption, pq, but also some indirect spending on leisure: to buy ℓ units of leisure, the consumer has to give up $w\ell$ units of potential income. On the right and side is the potential income: $24w$. That is what the consumer would earn if she worked 24 hours per day.

Thus, the model can be described using this metaphor: The consumer writes an employment contract that says that she is going to work for 24 hours per day, and in return gets paid $24w$. But, if she wants to have some leisure, she has the option to "buy back" the leisure. She has to buy it at a price of w. With this metaphor, our model looks a lot like the model that we have studied in previous topics, with the two goods being consumption and leisure, their prices being p and w, and the only difference from previous models is that income is no longer y, which was in previous topics just a given, but $24w$. We can study the optimal consumption and leisure demand of the consumer, and the labor supply will then simply be given by $24 - \ell$, that is, the remaining time. *Explicitly*, we study leisure demand, but *implicitly*, this is really about labor supply.

Figure 27 shows an example of optimal consumption and leisure demand in our model. The intersection of the budget line with the horizontal axis is at the point $(24, 0)$. This is because at this point the consumer spends all his time on leisure, does not earn any income, and therefore does not consume anything. The intersection of the budget line with the vertical axis is at the point $\frac{24w}{p}$. This is because at this point the consumer does not earn have any leisure, thus works 24 hours per day, and earns a wage of $24w$. From this, she can buy $\frac{24w}{p}$ units of the consumption good. You can read off the figure the optimal leisure and consumption demand, but also the optimal labor supply. Labor supply equals the distance between optimal leisure, and 24 hours, i.e. $L = 24 - \ell$. I have indicated in the figure the optimal leisure demand in blue and optimal labor supply in green.

It is interesting to write down explicitly the necessary condition for interior (i.e. not corner) solutions:

$$\frac{\frac{\partial U}{\partial \ell}}{\frac{\partial U}{\partial q}} = \frac{w}{p}.$$

This is just our standard condition, marginal rate of substitution equals price ratio, but now for the two goods leisure and consumption. In our current context, the right hand side has an interesting interpretation. We can interpret the right hand side as the "real wage." The nominal wage is w. But of course, anybody who considers their wage should not look at the nominal number, but they should think about what they can buy for it. A nominal wage of w buys $\frac{w}{p}$ units of the consumption good. This is what should really matter to the consumer. The same nominal wage is worth less in Manhattan than it is in a small

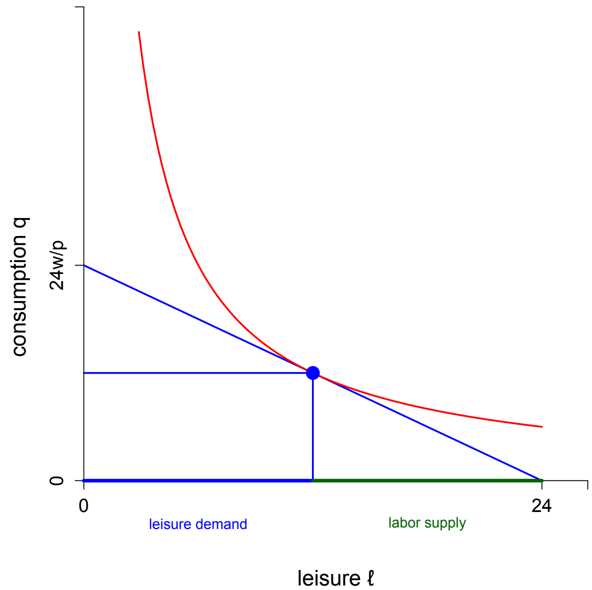


Figure 29: Optimal Choice of Consumption and Leisure

town in central Michigan. The optimality condition says that the marginal rate of substitution between leisure and consumption should be equal to the real wage.

Now let us do an interesting exercise: Let us consider how labor supply changes as the wage goes down. The following discussion is illustrated by Figure 28. First: how does the budget line change? Note that the intersection point on the horizontal line is determined purely by the length of a day, not by wage. Only the intersection point of the budget line with the vertical axis depends on w. Obviously, as w goes down, it decreases. Thus, if w goes down, the budget line rotates inwards around the point $(24, 0)$, and correspondingly the optimal choice of consumption and leisure changes, and also therefore implicitly labor supply. I show this in Figure 28. The figure shows the old budget line as a blue line, and the new budget line as a green line. You can see in the figure that in this example as the wage drops the leisure demand increases. This means that labor supply decreases, because, remember, leisure and labor have to add up to 24. Labor supply decreases by the distance between the green and the blue vertical lines. This is, perhaps, what you would expect: lower wages implies less labor supply. But let's take a closer look.

I shall now decompose, as we did in previous Topics, the effect of a price

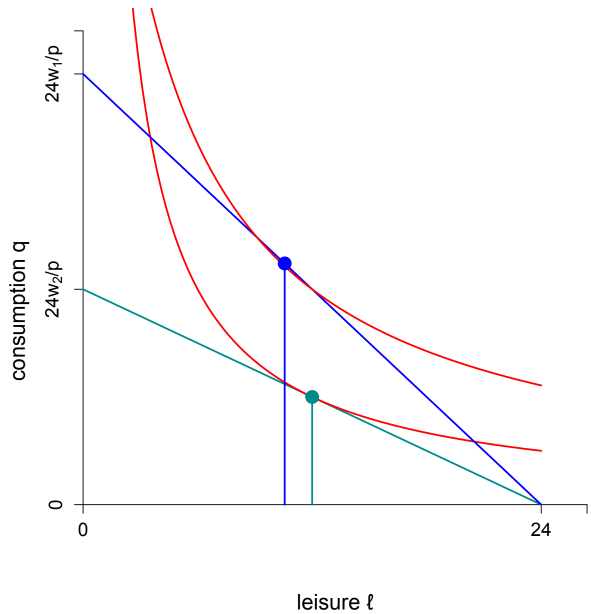


Figure 30: The Effects of a Wage Drop

change, here a wage drop, into an income and a substitution effect. Let me just draw this decomposition in the same way as we have done in previous Topics. The result is shown in Figure 29. The substitution effect is represented by the movement on the original indifference curve to the point where the dashed budget line, which is parallel to the new budget line, is a tangent of the indifference curve. You see in Figure 29 that indeed the substitution effect of a wage drop on leisure demand is positive, because the opportunity cost of leisure has gone down. This means that the substitution effect on labor supply is negative. But what about the income effect? If leisure is a normal good, as in Figure 29, then the income effect on leisure demand is *negative*, which means that the income effect on labor supply is *positive.* As wage goes down, you get poorer. Therefore, you demand less leisure, and instead work more. So, substitution effect and income effect go into opposite directions. Whether labor supply goes down, as in Figure 29, or up, depends on which of the two effects is stronger. You can easily draw a figure similar to Figure 29 in which the income effect is stronger than the substitution effect, and as a result the labor supply goes up when the wage drops. Of course, if leisure is an inferior good, that is, as you get richer you demand less leisure, then it is true that labor supply goes down as the wage drops. But that does not seem a very

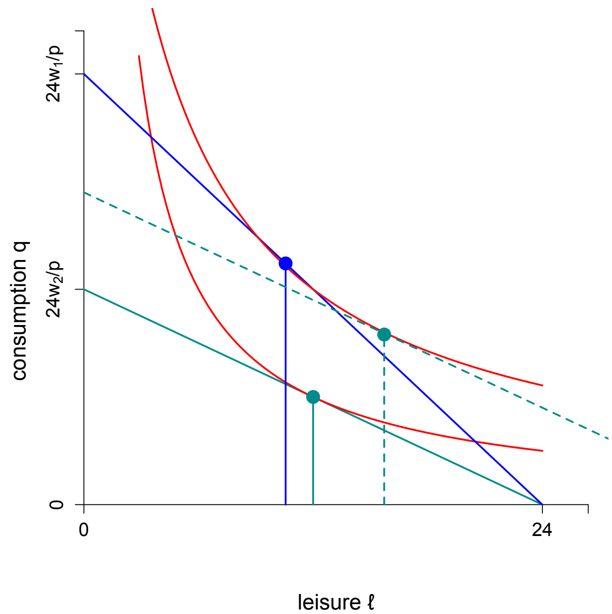


Figure 31: Income and Substitution Effects of a Wage Drop

plausible case.

This is a surprising implication of our rational choice model of labor supply. The effects of a wage increase on labor supply are, in the normal case, indeterminate. This is contrary to common wisdom that people work more when they earn more per hour. One way of decreasing wages is by increasing income tax. Our model does *not* unambiguously predict that an income tax rise decreases labor supply, as is commonly said in political discussions.

There is one very counter-intuitive aspect to Figure 29. Where does the green dashed budget line intersect with the horizontal axis? I haven't drawn the intersection point. But it will be somewhere to the right of 24. But a day has only 24 hours, and the intersection point with the horizontal axis is the length of a day! Well, an income increase in this model, if it is to correspond to a parallel shift of the budget line to the right, must necessarily correspond to an increase in the length of the day. Let's just accept that. The story is that God comes along and generously decides to make days longer. If you earn your income by working, and never fatigue, this makes you richer. It is a bit hypothetical, I admit. But it is the only story I can come up with. In short, in Figure 29, when we compensate the consumer for the drop in wages, we must play God, and increase the length of the day. This is what the dashed green line represents.

There are other effects that we could study in this model. But the effect of a wage drop is the most interesting. So, we shall leave it at that.

Savings

We shall think about savings again in our two good model. Therefore, we shall think in a model where there are only two time periods: $t = 1, 2$. Let us also imagine that there was income y only in period 1. We shall take this income now again as given. As I warned you before, we cannot study both labor supply and saving in the same model if we only want to draw two-dimensional diagrams. To keep things simple, in each period there will be a single consumption good. Let us denote consumption in period 1 by q_1, and consumption in period 2 by q_2. The consumer will have a utility function $U(q_1, q_2)$. It is natural that it is increasing. It also seems not implausible that preferences are convex, although we don't need this, of course, for our analysis. Here, convexity means that you prefer to have similar consumption in both periods over consuming a lot in period 1 but starving in period 2, or vice versa.

Let us denote the prices of consumption in periods 1 and 2 by p_1 and p_2. What is the consumer's budget constraint? Well, to obtain income in the second period she must save. There is no given income in the second period. What are her savings? If she has income y in period 1, and spends $p_1 q_1$ on consumption in period 1, then she has $y - p_1 q_1$ left over at the end of period 1. Let us denote her savings by s:

$$s = y - p_1 q_1.$$

Suppose there is an interest rate r, such that saving s today gives us income $(1 + r)s$ tomorrow. So, if the interest rate is 5%, then we should set $r = 0.05$. So the budget constraint can be written as the constraint that what the consumer plans to spend tomorrow must not be more than she earns through her savings:

$$p_2 q_2 \leq (1 + r)(y - p_1 q_1),$$

where I have already substituted $y - p_1 q_1$ for s.

Let us re-write the budget constraint a little bit:

$$\begin{aligned} p_2 q_2 &\leq (1+r)(y - p_1 q_1) \Leftrightarrow \\ \frac{p_2}{1+r} q_2 &\leq y - p_1 q_1 \Leftrightarrow \\ p_1 q_1 + \frac{p_2}{1+r} q_2 &\leq y. \end{aligned}$$

Now it looks again like our normal budget constraints have looked in past topics. On the right hand side we have the given income. On the left hand side we have the expenses on consumption in periods 1 and 2. But the price of consumption in period 2 is not p_2, it is $\frac{p_2}{1+r}$, that is, less than p_2. Why is that? If you save $\frac{p_2}{1+r}$ in period 1, then in period 2 you have, including interest: $(1+r)\frac{p_2}{1+r} = p_2$, that is, to buy one unit of the consumption good in period 2, you have to save $\frac{p_2}{1+r}$ in period 1. The budget constraint is written from period 1's perspective. from period 1's perspective consumption of one unit good 2 is a little less expensive than p_2, because the consumer has to save a little less than p_2 in period 1, earn interest, and then have p_2 available in period 2.

In Figure 30 I show an example of a budget constraint and an optimal choice. There is nothing very surprising in this graph. But please look briefly

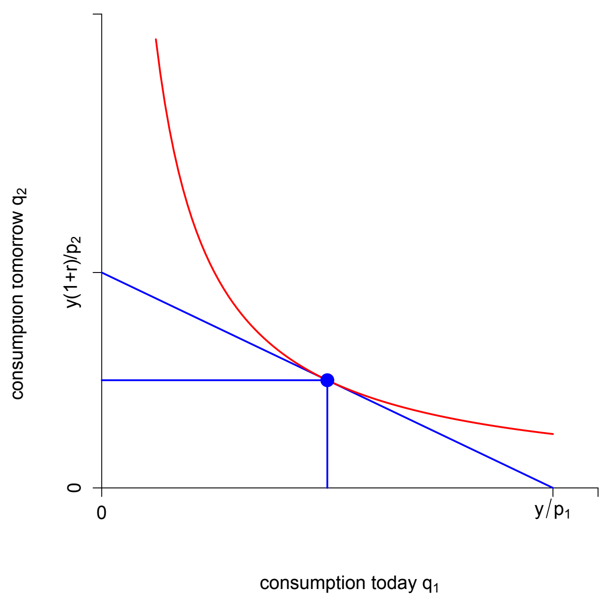


Figure 32: Optimal Consumption in Periods 1 and 2

at the intersection point of the budget line with the vertical axis. It is at:

$$\frac{y(1+r)}{p_2}.$$

That is intuitive: If you save all income in period 1, your income grows to $y(1+r)$ in period 2, and thus the maximum period 2 consumption is: $\frac{y(1+r)}{p_2}$.

We wanted to study optimal saving. Where is optimal saving in this graph? There is no “distance” in this graph that shows savings. But perhaps most naturally we should look at the consumption in period 2, because that must be financed through savings. Savings are this consumption, times p_2. But these are “nominal” savings. The “real” savings are just the period 2 consumption. So, maybe the optimal choice of period 2 consumption is what we should focus on.

As in the previous section, let’s do one simple exercise: Let us suppose the interest rate went up. What happens to savings? Well, the maximum possible consumption in period 1 does not change, but the maximum possible consumption in period 2 increases. Thus, the budget line rotates outwards around its intersection point with the horizontal axis. This resembles a price decrease in the price of consumption of good 2, and it can be analyzed just as a price decrease is analyzed in the model with two consumption goods. The substitution effect says that the consumer wants to consume more in period 2, as saving for period 2 has become more attractive.

The income effect corresponds to an income increase: a higher interest rate makes the consumer in real terms richer. Therefore, if period 2 is a normal good, period 2 consumption also goes up due to the income effect. Thus, the effect of an increase in interest rate on “real” savings is unambiguously positive if period 2 consumption is a normal good. This seems very intuitive.

We don’t have to do much more to study the savings model. It is like the static model with two consumption goods. All that we have to keep in mind is that the effective price of period 2 consumption is $\frac{p_2}{1+r}$, and is therefore affected both by the nominal price of period 2 consumption p_2, and by the interest rate r. If we keep that in mind, then what we learned in previous topics about the two consumption good model applies to savings with no change.

Topic 10: Revealed Rationality and Revealed Preferences in Consumption

The Weak Axiom of Revealed Preferences

We are now going to ask one more time: "What is consumer theory for?" We are going to adopt the perspective of an empirical researcher, who only observes behavior, but who does not know whether the consumer maximizes utility. Nor does the researcher know what the consumers' utility function is if she indeed maximizes utility.

To see as clearly as possible which issues the empirical researcher has to deal with, let's start with a very stylized example. Let us imagine we observed the same consumer in three different weeks making consumption choices. For the purpose of the example, suppose there are two goods. Each row in the table below lists observations of the prices of the two goods, of the income, and of the quantities demanded.

Table 1: Fictional Consumption Data

p_1	p_2	y	q_1	q_2
1	1	6	5	1
1	2	8	2	3
1	1	7	3	4

What do we learn from these observations about the consumer's utility function? Well, for example, the consumer's choice of the consumption bundle $(q_1, q_2) = (5, 1)$ when prices are $p_1 = p_2 = 1$, and the consumer's income is $y = 6$, tells us that she prefers this consumption bundle over others that she could have afforded, for example $(3, 3)$. It may also be that she is indifferent between the bundle that she has chosen, and some of the other bundles. If we knew that what the consumer chose was actually her only optimal choice, then we could even conclude that she strictly prefers $(5, 1)$ over $(3, 3)$. In fact, for this topic let us assume that whatever choice we observe is the unique optimal choice. We can now conclude that $U(5, 1) > U(3, 3)$ for this consumer.

Before we go further, though, let's throw out the utility function. The utility function was very useful when we studied the mathematics of rational consumer behavior, but it loses its usefulness in this section. This is because all we can ever observe when we see a consumer choose consumption bundles is what the consumer prefers. We can never observe the consumer's "utility." There is no observation that could reveal information about utility as a number. But remember from Topics 2 and 3, utility functions are just representations of preferences, and for any given preference there are infinitely utility functions that may equivalently be used to represent this preference. For a rational consumer's choices only the preferences matter. Even when we call the consumer "utility maximizing" we actually just mean that the consumer has a preference that can be represented by a utility function, and that she chooses what is best according to this preference. So, instead of asking "is the consumer maximizing utility?" we shall ask in this section"Is the consumer maximizing a preference?" And instead of asking "What is the consumer's utility function?" we shall ask in this section: "What is the consumer's preference?"

Now let us return to the consumer whose choices are shown in Table 1. Instead of saying that we have concluded that $U(5,1) > U(3,3)$, we shall say that we have observed that $(5,1) \succ (3,3)$. The notation "$\succ$" for preferences was introduced in Topic 2. It should be read as: "is strictly preferred to." Now, continuing, from the choices that we have observed we can conclude much more, namely that the consumer strictly prefers $(5,1)$ over all consumption bundle that are in her budget set when prices are $(1,1)$, and income is 6, i.e.:

$$(5,1) \succ (q_1,q_2) \quad \text{for all } (q_1,q_2) \text{ such that } q_1+q_2 \leq 6 \text{ and } (q_1,q_2) \neq (5,1).$$

Similarly, the second choice shows us that:

$$(2,3) \succ (q_1,q_2) \quad \text{for all } (q_1,q_2) \text{ such that } q_1+2q_2 \leq 8 \text{ and } (q_1,q_2) \neq (2,3),$$

and the third choice shows us that:

$$(3,4) \succ (q_1,q_2) \quad \text{for all } (q_1,q_2) \text{ such that } q_1+q_2 \leq 7 \text{ and } (q_1,q_2) \neq (3,4).$$

We seem to have learned quite a lot about the consumer's preferences. Maybe we can even go one step further. For example, we see that the third choice tells us, in particular, that: $(3,4) \succ (2,3)$, and the second choice tells us, in particular, that $(2,3) \succ (8,0)$. Using the assumption that preferences are transitive, that implies that: $(3,4) \succ (8,0)$. If we continue like this, and collect all preferences directly revealed through the consumption choices, as well as fill in all gaps that we can fill in using the assumption of transitivity, we get quite a lot of information about the consumer. We still miss out on some information. For example, there is no way that we can know whether the consumer prefers $(2,1)$ over $(1,2)$. This is because for all budget sets that the consumer faced, both $(2,1)$ and $(1,2)$ were included in the budget set, but never chosen. There is no information about the comparison of these two choices. But whatever information we can glean from the observed choices we call the consumer's revealed preferences.

But, there is actually a problem hidden in Table 1. The first choice shows us, in particular, that $(5,1) \succ (2,3)$, because $(2,3)$ is in the budget set when prices are $(1,1)$ and income is 6. The second choice tells us, in particular, that $(2,3) \succ (5,1)$, because $(5,1)$ is in the budget set when prices are $(1,2)$ and income is 8. But now we have a contradiction! We have concluded:

$$(5,1) \succ (2,3) \quad \text{and} \quad (2,3) \succ (5,1).$$

These two cannot be true at the same time! In fact, the revealed preference is inconsistent! This consumer does not make rational choices.

I illustrate the situation in Figure 31, where I have drawn budget sets, and choices, for all three rows in Table 1. The problems result from the two choices $(2,3)$ and $(5,1)$. Each is in the interior of the budget set when the other consumption bundle is chosen. So, each is revealed preferred over the other. This is the contradiction. Note that there is no problem with the choice $(3,4)$. This choice is in neither of the other budget sets.

We call a situation as in Table 1 and Figure 31 a violation of the Weak Axiom of Revealed Preferences. This axiom says:

> If consumption bundle $\hat{q}$ is revealed preferred over consumption bundle q, then consumption bundle q must *not* be revealed preferred over consumption bundle $\hat{q}$.

Sometimes, this axiom is given the acronym WARP. When there are two goods only, then a well-known mathematical result says that there exists a preference such that a consumer's choices are optimal given that preference, i.e.

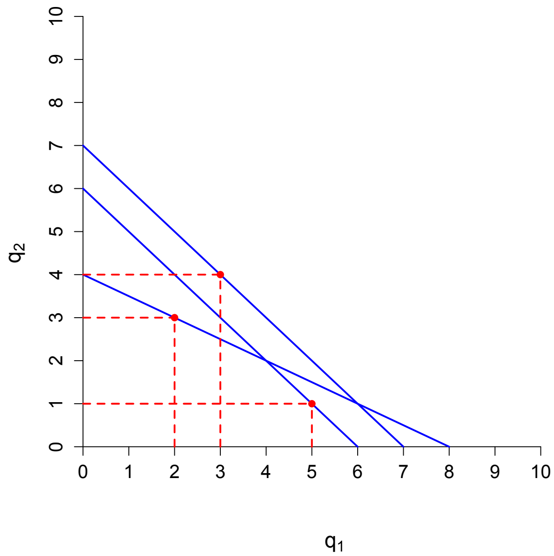


Figure 33: Fictional Consumption Data

the consumer is rational in the sense of economics, if and only if the consumer satisfies WARP. It is obvious that a rational consumer must satisfy WARP. But it is much more complicated to show that if a consumer satisfies WARP there is a preference that she maximizes. We shall skip that step.

Checking WARP looks at first sight difficult, because each choice that the consumer makes reveals infinitely many strict preferences, namely the strict preferences of the chosen consumption bundle over all others in the budget set. Perhaps, an observed choice also reveals further preferences if we consider those preferences that are implied by transitivity. It is, however, easy to see that it is sufficient to only check whether WARP holds for the consumption bundles that were at least once chosen by the consumer. All other consumption bundles can be neglected. In our example, only three consumption bundles were chosen. So, we really needed to check WARP only for pairs made up of two of those three consumption bundles. Thus, checking WARP would have quickly lead us to detect the contradiction that we pointed out here, even if we didn't know about it.

Why is this true? WARP involves two consumption bundles each of which is revealed preferred over another one. But a consumption bundle can only be revealed preferred if it is chosen at least once. Therefore, both consumption bundles in WARP have to have been chosen at least once.

WARP is a black and white criterion: it determines either that a consumer is rational, or that she is not. There is no grey area. Another question that we could ask is not whether people sometimes violate the conditions for rationality, but whether these violations are "serious." But what is "serious"? This is a good question, but one that we have to leave open here.

Testing Rationality in Practice

How do people choose their consumption in practice? Are they rational? Many surveys, run by governments or research institutes, collect data on households' consumption behavior. These data have been used to perform tests of rationality in consumption, based on requirements, such as the WARP explained in the previous section. Typically, in practice, not WARP, but GARP, is used, where the "G" stands for "generalized." GARP is the right condition in a world with more than two goods, i.e. the world of real world data.

Let's study what this research typically finds. We'll base our discussion on a paper by Timothy Beatty and Ian Crawford with the title "How Demanding is the Revealed Preference Approach to Demand?" that was published in the *American Economic Review* in October 2011, pages 2782-2795.

I present a very small part of the results of the paper. What I present is not really the focus of the paper, but just a brief empirical application.

Their work studies the consumption behavior of Spanish households between 1985 and 1997, using a government collected expenditure survey. They have observations on 3,134 households. The data report expenditures not for individual goods, but for "aggregates," such as "food." Beatty and Crawford regard these as "goods."

The first result that Beatty and Crawford report is that 95.7% of the observed households pass the rationality test. That is an extraordinary large number. The result matches observations in other studies. Beatty and Crawford write that previous research has found predominantly rational behavior among New York dairy farmers, Danish consumers, children, and capuchin monkeys. The paper goes further, though. It asks, *why* it seems that everyone in the world is rational.

Here is their second finding: They compare the rational proportion among real the Spanish households with that of a similar, fictional group of households, who have the same budget sets as the Spanish households, but who randomly pick their consumption. This group of households can just be simulated on the computer. They discover that those households would also pass the rationality test with a high rate, only 4.5 % lower than the observed pass rate of the Spanish households. Beatty and Crawford interpret this result as a caution: Don't interpret the high number of people passing the rationality test as a "triumph of economic theory."

Why do they have the finding that even random households behave rationally? It is because in the real world budget sets are rarely aligned as shown in Figure 31. Most of the time budget sets are nested, as they are in our example in observations 1 and 3. That is, one budget set is just larger than the other. This is because price fluctuations in their set of observations are small, but income variations are large. If one budget set is larger than another, whatever choice the household makes satisfies the axiom of revealed preferences.

Beyond the fact that even random choices from real world budget sets are often rational, the authors offer one more finding that should make us cautious. They find that 133 households violate the rationality axiom GARP. The next question to ask is: how "bad" are the violations? Beatty and Crawford propose a measure of how bad the violations are. We cannot go here in to the details of this measure. But the bottom line is that those 133 households that do violate the rationality axiom, violate it "badly."

What does all of this teach us? Two lessons: One is: approaching data through the lenses of the theory that we have studied is likely to be successful. The second is: that does not tell us much about the "true" rationality of people.

But is it useful to rely on the economic model of consumer behavior when studying consumer choices in practice? Or should we perhaps adopt a random choice model? One advantage of the economic approach is that it allows us to attribute preferences to people. Those preferences then allow us to assess how well the market satisfies these preferences. We shall discuss later in detail what we mean by this question. But having an answer to this question is extremely useful when thinking about public policy and the market. Economists proceed in this way all the time. You might worry that the preferences that we attribute to consumers are not really those that they reveal to us through their choices. But what else do we have to work with? In the next section, we shall discuss in a stylized example *how* to recover preferences.

Estimating Preferences

Suppose we observed the same consumer four times, making the choices shown in Table 2. Each row corresponds to one choice. Perhaps, each row corresponds to the purchases of the consumer in a given week There are no violations of the strong axiom of revealed preferences in this table. One can check this, but we won't. So, we know from the previous section that we can interpret this consumer's choices as preference maximizing. Suppose now we wanted to predict how the consumer would respond, say, to an increase in income from 30 Dollars to 40 Dollars, keeping the prices the same as in the last row? To do this, we would have to know which preference the consumer maximizes. But how can we do this? There are many preferences which imply that the observed choices are optimal. Using language introduced in Topic 2, we can say that the observed choices reveal the preferences only incompletely.

Table 2: More Fictional Consumption Choices

p_1	p_2	y	q_1	q_2
1	1	20	15.6	4.4
5	1	20	2.5	7.7
4	2	25	4.5	3.6
3	2	30	7.4	3.8

One common way of proceeding is to restrict attention to preferences that can be represented by a utility function of some particular form. Here is an example of such a form:

$$U(q_1, q_2) = \alpha q_1^\beta + (1-\alpha) q_2^\beta.$$

where α and β are parameters strictly between 0 and 1. We have encountered this utility function before in the special form: $U(q_1, q_2) = \sqrt{q_1} + \sqrt{q_2}$, which is the special case that $\alpha = \beta = 0.5$. With these parameters you get: $U(q_1, q_2) = 0.5\sqrt{q_1} + 0.5\sqrt{q_2}$. But if you then multiply by 2 you obtain: $U(q_1, q_2) = \sqrt{q_1} + \sqrt{q_2}$, and multiplying by 2 leaves, of course, the preferences unchanged.

Here, $0 < \alpha < 1$ and $\beta > 0$ are required for the utility function to be monotonically increasing, and $\beta < 1$ is needed to make sure that the preferences are convex.

Notice that we are bringing back utility functions, which we had just thrown out in the previous section. This is not because we give the utility numbers any particular meaning. Using a particular form of utility functions really just means that we are using a particular type of preferences. For each value of α and β there is a preference that is represented by the utility function, and, in this case, also for different values of α and β the represented preferences are different. Thus, if we use our language very carefully, we can say that have restricted attention to a subset of possible preferences, and we have introduced parameters, α and β such that each preference in our subset corresponds to one, and only one, pair of values of α and β. At this point it makes sense to say that our objective is to recover the values of α and β from the observations that we make.

The demand functions for a consumer with a utility function of the form shown above are quite tedious to derive. Just for completeness, I give you the result. But you don't have to do worry about the calculation that leads to this result. Moreover, you can read the remainder of this topic without going through the following equations in detail. You need those equations only if you want to verify by yourself some of what I say below.

$$\begin{aligned} D_1(p_1, p_2, y) &= \frac{\alpha^{\frac{1}{1-\beta}} p_2^{\frac{\beta}{1-\beta}}}{\alpha^{\frac{1}{1-\beta}} p_2^{\frac{\beta}{1-\beta}} + (1-\alpha)^{\frac{1}{1-\beta}} p_1^{\frac{\beta}{1-\beta}}} \quad \frac{y}{p_1} \quad \text{and} \\ D_2(p_1, p_2, y) &= \frac{(1-\alpha)^{\frac{1}{1-\beta}} p_1^{\frac{\beta}{1-\beta}}}{\alpha^{\frac{1}{1-\beta}} p_2^{\frac{\beta}{1-\beta}} + (1-\alpha)^{\frac{1}{1-\beta}} p_1^{\frac{\beta}{1-\beta}}} \quad \frac{y}{p_2}. \end{aligned}$$

If we assume that the consumer has a utility function and demand functions of the form described in the previous two paragraphs, then we can try to use our observations to infer the values of α and β. Trying to do that for the data shown in Table 2 is actually very tedious. If you have great programming skills, you may try and program some way of doing so. If you want to do so, don't read on, but first get to work.

For the lazy rest of us, let me give the answer away. I generated the data in Table 2 by maximizing a utility function of the form shown above. As parameters, I have chosen:

$$\alpha = 0.7 \quad \text{and} \quad \beta = \frac{1}{3}.$$

You can check my calculations using the formulas for demand that I showed you above.

But now let us suppose that our consumer had been observed to make the choices shown in Table 3. It turns out that there is no value of α and β that predicts exactly these consumption choices. But actually, they are not so far from the choices in Table 2. We might argue that the values for α and β that we had before are pretty good explanations of what we have seen here as well. Maybe the consumer maximized utility for those values of α and β, but also made some random errors.

Table 3: Yet More Fictional Consumption Data

p_1	p_2	y	q_1	q_2
1	1	20	14.9	5.1
5	1	20	2.4	8
4	2	25	4.3	4.0
3	2	30	7.5	3.7

But before we jump to that conclusion, what should try to figure out the *best* values of α and β to explain the in Table 3 choices? There are many ways of formalizing this question. Let us take a simple one: Let us suppose we wanted to know what values of α and β predicted choices that came "closest" to the observed choices. What do we mean by "closest"? We could measure the distance between a predicted quantity consumed and an observed quantity consumed by taking the absolute value of the difference between the two quantities. An often used alternative is to use the square of the difference between the two quantities. This latter measure of distance attaches more importance to large deviations than to small deviations. For illustration purposes, we shall use here the absolute value of the difference as our measure of distance.

I have written a little program that discovers the values of α and β that do best at predicting the choices observed in Table 3. They are:

$$\alpha = 0.681 \quad \text{and } \beta = 0.292.$$

The total prediction error for these parameter values is just 1.11. That is, perhaps, not such a large deviation from the observed quantities. The largest prediction error is in the fourth row, where the best guess parameters predict that only 7.1 units of good 1 are consumed, but 4.3 units of good 2, which is not quite what we observe. I show all predictions in Table 4.

So, now we know what we would infer about α and β if we observed the

p_1	p_2	y	q_1	q_2
1	1	20	14.9 (14.9)	5.1 (5.1)
5	1	20	2.4 (2.4)	8 (8)
4	2	25	4.3 (4.3)	4.0 (3.9)
3	2	30	7.5 (7.1)	3.7 (4.3)

Table 4: Yet More Fictional Consumption Data (Predictions in Red)

choices in Table 3 and knew nothing about how those choices came about. But here is a secret: I have generated the data in Table 3 using the same parameter values as for Table 2, but adding some random noise. You can see that my best guess parameter values are pretty close to the true ones. I could try to use my best guess to make a prediction. For example, if we observed a period that is identical to the fourth row, but in which the price of good 2 went up to 4, then my prediction would be that the consumer chooses the consumption bundle $(7.7, 1.75)$. But the truth is that he would choose $(8.0, 1.5)$, plus some random error. But I would not be totally far off.

In this section, I have used the word "best guess" as a placeholder for "estimate." When you learn proper statistical methods, you will learn how to obtain "best guesses" in a much more sophisticated way, and then they deserve to be called "estimates." In the following, I shall sometimes use the language of "estimation."

Estimating Demand Functions

Instead of estimating a utility function, one could also simply estimate a demand function directly. In the previous section, had we just started with demand function, without ever talking about the utility function, we would have estimated α and β in exactly the same way. We could have written down the demand functions directly. It would be a "two parameter" model of demand. All that we achieved through our utility maximization calculation was that we were sure that our model of demand is compatible with the economic theory of rational demand.

A modern approach is to start from demand functions. Moreover, a modern approach is "non-parametric." What this means is that, instead of writing down a function such as the demand function in the previous section, where parameters"α" and "β" were the only things non specified, we just let the function be *anything*. Modern statistics has developed "non-parametric" methods that allow the researcher to estimate functions with arbitrary shapes.

But how can we check whether the demand function that we obtain can be derived from utility maximization? One approach would be to search the set of *all* utility functions, and to investigate whether one of them, when maximized, implies the demand function that we observe. But that is an impossible undertaking. We could alternatively check whether WARP is satisfied, but that does not quite make sense because a demand function specifies for *every* list of prices and income what the consumer purchases. Prices and income are continuous variables, and thus we would have to conduct infinitely many checks of WARP.

An alternative approach is indirect, using our Slutsky equation: "substitution effect + income effect = total effect." This equation implies: "total effect - income effect = substitution effect." Now recall that the substitution effect has to be negative if a consumer maximizes a utility function. Therefore, we find that a condition that every demand function that maximizes preferences must satisfy is:

$$\text{“total effect - income effect”} \leq 0.$$

If we have data rich enough that allow us a non-parametric estimation of a demand function of the form $D_1(p_1, p_2, y)$, then we can actually check the above condition. And, it turns out, that a generalization of the above condition is *necessary and sufficient* for a demand function to result from utility maximization. Sometimes, the condition is referred to as the *Slutsky condition*.

Thus, when estimating demand functions, we might wish to restrict attention to demand functions that satisfy the Slutsky condition. This is what Richard Blundell, Joel Horowitz, and Matthias Parey did in a paper entitled "Measuring the Price Responsiveness of Gasoline Demand: Economic Shape Restrictions and Nonparametric Demand Estimation" that was published in a journal called *Quantitative Economics* in 2012, pages 29-51.

These authors use data from something called the "National Household Travel Survey," a survey conducted by the United States government in 2001 and 2002. They find the "optimal guess" at a demand function, under the restriction that they only consider demand functions that satisfy the Slutsky restriction. Aside from that, no restriction is imposed on the shape of the demand function. In Figure 32, I show their estimates of gas demand function for different income groups. Be careful in looking at this table: Blundell and his co-authors place prices on the horizontal axis, and quantities on the variable axis, contrary to the economics textbook convention that we have also adopted in these notes.

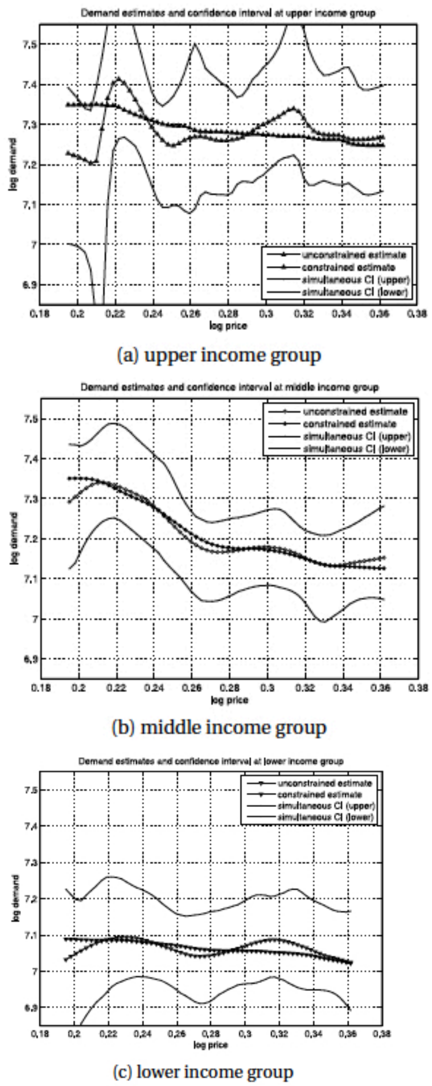


FIGURE 2. *Continued.*

Figure 34: Estimated Gas Demand

There are many lines in the diagrams. Focus on the lines that are in the middle, and not on the outer lines. The outer lines just indicate the extent of the uncertainty of the estimation. We'll ignore this issue here. The lines in the middle show two possible estimates of the demand function. One restricts attention to demand functions that satisfy the Slutsky restriction. The other estimate is shown for comparison purposes. It is constructed using the same methods, but not imposing the Slutsky restriction.

The estimated demand functions are not too far apart in either of the three figures. That is perhaps a reflection of the fact that most people, as we noted above, are not too far from rationality in their consumption behavior. But, note that the demand function estimated without imposing the Slutsky restriction is not always downward-sloping, i.e. we seem to find that gas can be a Giffen good. When imposing the Slutsky condition, by contrast, we find downward sloping demand functions. Note that this is not by assumption: the Slutsky restriction does not rule out Giffen goods, as you know! It just happens to be so in these data.

One other observation is that the demand function for the middle income group is much steeper than the demand function for the upper and lower income groups. That may seem intuitive. If you are poor, you just have to drive, you have no alternative. If you are rich, you don't care about gas prices. But if you are in the middle, perhaps you are rich enough to live in an area where there are alternative modes of transport. So, these households have more price elastic gas demands.

Here we end our long journey through the theory of consumer demand. Its details are sometimes tedious. But the big picture, economists' attempt to understand consumption behavior through the lenses of utility maximization, is, if nothing else, an impressive piece of scientific research. The research is ongoing. The project has its pros and cons. But it is still one of the core pieces of economic theory.

Topic 11: Production Functions

A Stylized Image of the Firm

Firms are organizations that produce goods or services. But how would we define an "organization"? One definition is that organizations are a "network of contracts." In the case of firms, this network of contracts consists of the contracts with suppliers, the managers' contracts, the workers' contracts, the contracts with customers, and the contracts with lenders and shareholders. All these different agents contract with each other, thus creating a network of contracts, and this network "is" the firm. The firm is not any individual on its own.

Once the contracts are written, lenders and shareholders provide money, suppliers provide inputs, workers and machines turn the inputs into outputs, and then the outputs are sold to customers. Which output is produced depends on the firms' technological knowledge, that is, its knowledge of how to turn inputs and labor into output. Which output is produced also depends on the incentives created by the contracts. For example, managers might use their time playing computer games rather than managing, if their contracts don't create incentives for effective management, and little output might be produced as a consequence of the lack of effective management.

The incentives created by contracts form a large part of the modern theory of the firm. We shall study incentive contracts later in this course. For the moment, we shall abstract from all issues related to contracts, and construct a much more stylized image of the firm. It is simple: firms have a certain given technological knowledge about how to transform inputs into outputs. Among all the technologically feasible production plans, firms then chooses, and carry out, a profit maximizing production plan. All other problems are assumed away: Nobody shirks. Nobody plays computer games. No suppliers deliver low quality. Etc. This is an unrealistic starting point. But it allows us to introduce some ideas that are also essential for understanding the more advanced theory of the firm, and also for understanding the theory of markets.

We shall also assume that firms are price-takers, that is, they observe the prevailing prices in the market, and they consider no alternative but to buy and sell at those prices. They don't think that they can choose the prices for their products, for example. We made the assumption of price taking behavior also for consumers in our discussion of consumer theory. For consumers, this assumption seems plausible in many contexts. For firms, it is much less plausible, although it is perhaps not too far-fetched that in some markets, say the market for hotel rooms, there is an understanding of a range of prices from which a hotel cannot deviate too far. We sometimes also refer to the assumption of price taking behavior as the assumption that there is perfect competition among firms. What we are currently studying, but not in all subsequent sections, is therefore the theory of perfectly competitive firms.

Production Functions

Production functions describe the technological knowledge of a firm, that is, its knowledge about how to make outputs from inputs. For our study of production functions, and also actually for the remainder of this course, we shall make two simplifying assumptions. The first is that the firms that we consider produce only a single output. The second is that the firms that we consider use only two inputs, labor and capital. Both assumptions together allow us to draw simple graphs. Some of what we are going to say will easily generalize to the case of more than two inputs and more than one output. But, to be honest, I should admit that some other parts of the theory need to be modified quite substantially, in particular if one wants to accommodate multiple outputs.

Let us denote by q the output quantity that the firm produces. Let us denote by L and K the quantities of labor and capital respectively that the firm uses as its inputs. When referring to capital we have in mind machines, factories, computers, etc. Firms will choose q, and they will choose L and K. But, of course, they are constrained by what is feasible. For example, it is not feasible to produce 10 units of most types of output from zero units of capital an zero units of labor. Nobody knows how to do that. The production function that we are about to introduce describes what is feasible. More precisely, it describes what the firm knows how to do. In this sense it describes the firm's technological knowledge.

I apologize for using the same letter L here for the labor input as I used in Topic 9 for labor supply by consumers.

The production function f assigns to every combination L and K of non-negative input quantities a number q of feasible output quantities:

$$q = f(L, K)$$

A good way to interpret this is that the production function describes the *largest possible output* that can be produced with a given input combination. Firms in our theory will never conside to produce less than the largest possible output, because this would not be profit maximizing. Therefore, we just take as given that if the firm uses input quantities L and K, then it will produce output $q = f(L, K)$, and not less.

Much of what follows will be built on the similarity of production functions and utility functions. Utility functions assign to two quantities, q_1 and q_2, the utility $U(q_1, q_2)$. This is similar to production functions that assign to two quantities, L and K, the output $f(L, K)$. The similarity between these two functions will make some of what comes very easy for you. There is, however, one difference that is very important to keep in mind. The interpretation of utility functions that we explained earlier is *o*rdinal, that is, the utility number itself has no meaning. Only the comparison between two utility numbers is meaningful. Utility functions are just convenient representations of preferences. This is different for production functions. For a production function, the actual value matters. The actual value is a quantity of outputs, that is, for example, $q = 10$ cars. We have no difficulty understanding what that means. By contrast, if someone tells us that the utility from some consumption is 10 utils, if we are honest, we have no idea what this is supposed to mean.

Isoquants

We shall now introduce a graph that helps us to visualize a production function. We shall construct a diagram that has the quantity of labor used on its horizontal axis, and the quantity of capital used on the vertical axis. We shall draw into this diagram the location of all combinations of labor and capital that make it possible to produce some given output q. Note that this is just what in consumer theory we had called an indifference curve. Here, we shall call these curves isoquants.

The "iso" stands for "equal." The "quant" stands for "quantity."

There is an isoquant for every quantity that could possibly be produced. That is, there is an isoquant for $q = 1$, $q = 3.5$, etc. Let us denote by $\bar{q}$ some arbitrary fixed quantity, then the isoquant corresponding to $\bar{q}$ is the graph of all pairs (L, K) that satisfy:

$$\bar{q} = f(L, K).$$

Figure 35 shows three isoquants. Each corresponds to a different quantity of output. The non-negative orthant is completely covered by isoquants. It would not make sense to try to draw all of them, however.

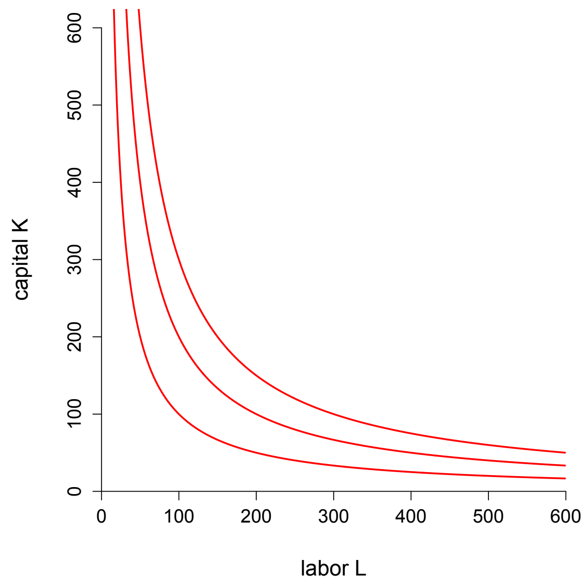


Figure 35: Isoquant

Note that in Figure 35 any given output can be produced with very little labor, but a lot of capital, but also with very little capital, and a lot of labor. That might not be always realistic. We shall discuss later in this topic some cases in which a certain output can only be produced if, say, a certain amount of capital is available.

For utility functions, the diagram showing all indifference curves, together with an indication into which direction utility was increasing (typically: indifference curves further away from the origin correspond to higher utility levels), indicated for us all that we needed to know about a given utility function. This is because the indifference curve diagram provided complete information about the consumer's preferences that were represented by the utility function. The diagram did not tell us what the utility numbers were, but we didn't care about those numbers anyway. The situation is different for production functions.

We do care which precise output quantities correspond to each isoquant, and therefore the isoquant diagram gives us only incomplete information about the production function. Really, the production function should be represented by a 3-dimensional diagram, where the third dimension shows the quantity produced. But 3-dimensional diagrams are harder to produce, and for some of our discussion below the 2-dimensional diagram will be enough.

Marginal Rates of Technical Substitution

We shall now introduce some assumptions for production functions. The first assumption is: f is monotonically increasing in K and L. What that means is simple: the more capital we use, the larger is the output that we can produce. Or at least it does not get smaller. The same is true for labor. Mathematically:

$$\text{If } \hat{L} > L, \text{ then } f(\hat{L}, K) \geq f(L, K) \text{ and if } \hat{K} > K \text{ then } f(L, \hat{K}) \geq f(L, K).$$

If f is monotonically increasing, then isoquants must be downward sloping. This is because they show the combinations of labor and capital that produce *the same* output quantity. If an isoquant were increasing, then a combination of more labor and more capital would produce the same output as a combination of less less and less capital. But that contradicts monotonicity.

Now let's look at how steeply an isoquant is decreasing, that is, at the slope of the isoquant. We shall call the absolute value of this slope the marginal rate of technical substitution. It is completely analogous to the marginal rate of substitution in consumer theory. It tells us how much capital the firm can give up if it hires one more (infinitesimally small) unit of labor, provided that the output that it wants to produce does not change. We are going to use the acronym MRTS for the marginal rate of substitution. Because it is an absolute value, it is non-negative.

We can calculate the MRTS for a production function in the same way as we could calculate the MRS for a utility function. Let us define the marginal product of labor to be:

$$\frac{\partial f}{\partial L}.$$

It indicates by how much the output goes up if the firm hires one more (infinitesimally small) unit of labor. Similarly, the marginal product of capital is:

$$\frac{\partial f}{\partial K}.$$

It shows how much the output goes up if the firm buys one more (infinitesimally small) unit of capital. The marginal rate of substitution can be calculated as the ratio of the marginal products:

$$\text{MRTS} = \frac{\frac{\partial f}{\partial L}}{\frac{\partial f}{\partial K}}.$$

Let us look at an example, that will look familiar from consumer theory: If the production function is:

$$f(L, K) = L^{0.25}K^{0.25},$$

then the marginal rate of technical substitution is:

$$\text{MRTS} = \frac{\frac{\partial f}{\partial L}}{\frac{\partial f}{\partial K}} = \frac{0.25L^{-0.75}K^{0.25}}{0.25L^{0.25}K^{-0.75}} = \frac{K}{L}.$$

We shall refer to production functions of this form as Cobb-Douglas production functions. They are thus analogous to Cobb-Douglas utility functions in consumer theory.

In the above example, the MRTS is decreasing as L increases and K decreases, that is, as we move along an isoquant. If that is the case, then the isoquants are convex. The intuitive interpretation is that the amount of capital that the firm can give up without reducing the output decreases as it has more labor and less capital. It becomes more and more reluctant to give up capital. Another interpretation, as in the case of consumer theory, is that if two input bundles produce the same output, that is, are on the same isoquant, then a mixture of these two input bundles produces more output than each of them.

When we discussed convex indifference curves, I explained in some detail different versions of the "law of diminishing marginal utility." I distinguished versions which make sense from versions which do not make any sense, because they implicitly assume that utility numbers by themselves are meaningful. For production functions, the situation is much easier. A firm has a diminishing marginal product of labor if the marginal product of labor decreases as the labor input increases, holding capital fixed. A firm has a diminishing marginal product of capital if the marginal product of capital decreases as the capital input increases, holding labor fixed. Together, diminishing marginal product of labor and diminishing marginal product of capital imply that the isoquants are convex. This is hopefully plausible, but I won't prove it here, because the proof needs a little bit of knowledge of multi-dimensional calculus.

Returns to Scale

In some industries, one can imagine that the production function is such that much more output can be produced if the firm operates at a sufficiently large scale. For example, if two cooks work in parallel, having two pots to cook with, they may be able to produce more than twice of what each cook can produce working alone, each having just one pot to cook with. In this section, we shall define this property formally. It has important economic consequences.

In our example, let us think of the cooks as labor, and of the pots as capital. The statement of the previous paragraph about the productivity of cooks then says that $f(2L, 2K)$ is more than twice $f(L, K)$, that is:

$$f(2L, 2K) > 2f(L, K).$$

Of course, we may study the same comparison for three cooks, and three pots, etc. In general, if

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

"λ" is the Greek letter pronounced "lambda."

is true for all $\lambda > 1$ and all combinations of L and K, then we say that the firm has increasing returns to scale. Why do we only consider $\lambda > 1$? It is because only multiplying the input quantities by a number greater than 1 corresponds to "scaling up" production.

In the case of constant returns to scale, we need not restrict attention to $\lambda > 1$. In fact, if $f(\lambda L, \lambda K) = \lambda f(L, K)$ for all $\lambda > 1$ and all (L, K) then it is also true for all $\lambda < 1$. Why? Maybe you can prove it yourself.

Let us introduce two other cases. We say that a firm has constant returns to scale if

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

for all $\lambda > 1$ and all combinations of L and K, and we say that a firm has decreasing returns to scale if

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

for all $\lambda > 1$ and all combinations of L and K.

Which of these cases is more plausible? Well, it surely must depend on the industry that we are considering. Moreover, it is surely an empirical question, and therefore we should refrain from discussing it without evidence, which is what we shall do here.

Note that not every production function implies increasing, constant, or decreasing returns to scale. It may well be true that $f(\lambda L, \lambda K) > \lambda f(L, K)$ for some λ, L and K, but $f(\lambda L, \lambda K) < \lambda f(L, K)$ for some other λ, L and K.

Let us consider in our numerical example whether it has increasing, constant, or decreasing returns to scale, or perhaps none of those. Recall that the production function was:

$$f(L, K) = L^{0.25} K^{0.25}.$$

Multiplying all inputs by λ, we get:

$$f(\lambda L, \lambda K) = (\lambda L)^{0.25} (\lambda K)^{0.25}.$$

Now we can factor out λ as follows:

$$(\lambda L)^{0.25} (\lambda K)^{0.25} = \lambda^{0.25} L^{0.25} \lambda^{0.25} K^{0.25} = \lambda^{0.5} L^{0.25} K^{0.25}.$$

Now notice that in this last expression the last two factors equal $f(L, K)$. Therefore, we can write:

$$f(\lambda L, \lambda K) = \lambda^{0.5} f(L, K).$$

That is, if both input quantities get multiplied by λ, then output gets multiplied by the square root of λ. But when $\lambda > 1$, then the square root of λ is smaller than λ, and therefore:

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

which means that this production function has decreasing returns to scale.

Suppose alternatively the production function were:

$$f(L, K) = LK.$$

This function is a monotone transformation of the function in the previous paragraph, and therefore, in consumer theory, we would say that it presents exactly the same preferences. But in producer theory, of course, it is a different production function. Indeed, using the same calculations as before, you can find:

$$f(\lambda L, \lambda K) = \lambda^2 f(L, K),$$

which implies, when $\lambda > 1$:

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

and we have increasing returns to scale, the opposite of what we had before.

Complements and Substitutes Among Factors of Production

Everyone's nightmare: you are a perfect substitute for a machine; or perhaps even an inferior substitute. ("Inferior substitute" is not part of our formal terminology.)

Sometimes, a precise combination of labor and capital is needed to produce one unit of output. For example, to make one omelet, you may need exactly ten minutes of labor, and two eggs. In this case, we might say that the two factors of production are perfect complements. Perfect complements in production are really completely analogous to perfect complements in consumption. And similarly, we can also speak of perfect substitutes in production, if, say, every worker can just as well be replaced by one machine. In Figure 36, I show what isoquants look like if we have perfect complements, or perfect substitutes. The diagram will look familiar from consumer theory.

What about imperfect complements and imperfect substitutes? As in consumer theory, we shall drop the adjective "imperfect," and just refer to complements and substitutes. In consumer theory we defined these ideas only after we had introduced the demand function. In producer theory, though, we can use the fact that output numbers, unlike utility numbers, have a meaning, and offer a definition that only refers to the production function. We shall say that the two factors of production are complements if:

An interesting theorem of multi-dimensional calculus says that if the first line in the definition of complements holds, then the second line is automatically true. The same is true for the definition of substitutes.

$$\text{the marginal product of labor, } \frac{\partial f}{\partial L}, \text{ is increasing in } K,$$

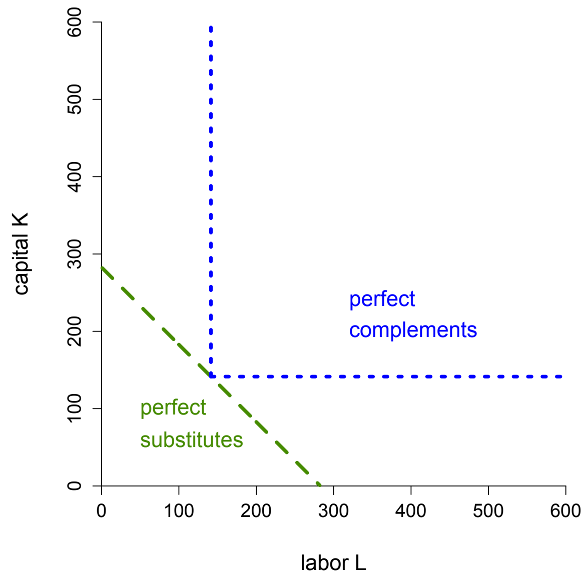


Figure 36: Perfect Complements and Perfect Substitutes

and

$$\text{the marginal product of capital, } \frac{\partial f}{\partial K}, \text{ is increasing in } L.$$

What does this mean in words? The extra output that the firm gets from hiring one more worker is the larger the more capital it has. And the extra output that the firm gets from one extra unit of capital is the larger the more workers the firm has. Conversely, we say that the two factors of production are substitutes if

$$\text{the marginal product of labor, } \frac{\partial f}{\partial L}, \text{ is decreasing in } K,$$

and

$$\text{the marginal product of capital, } \frac{\partial f}{\partial K}, \text{ is decreasing in } L.$$

Probably, these definitions make intuitive sense to you without further explanation.

Let us return to our numerical example. In this example, are the factors of production complements or substitutes? We can calculate, for example, the marginal product of labor as:

$$\frac{\partial f}{\partial L} = 0.25 L^{-0.75} K^{0.25}.$$

The right hand side increases as K increases, and therefore, capital and labor are complements. The calculation for the marginal product of capital shows the same result.

Topic 12: Cost Functions

Recall that our objective is to study firms that maximize profits taking prices as given. A useful first step towards studying firms' profit maximization problem is to study, for any given quantity q, the cheapest combination of labor L and capital K that is on the isoquant corresponding to q. The expenditure needed to buy that combination of labor and capital is called the "cost of producing q," which we shall also denote by $C(q)$. We shall call the function C that assigns to every quantity q the lowest expenditure required to produce q the cost function. The cost function obviously depends on the production function f, but also by the prices of labor and capital. We shall denote the price of labor by w, for "wage," and the price of capital by r, for "rent" (as if the firm rented its capital; but actually, it will buy it). We shall complete the study of the firm's maximization problem in the next topic by considering its optimal choice of output q.

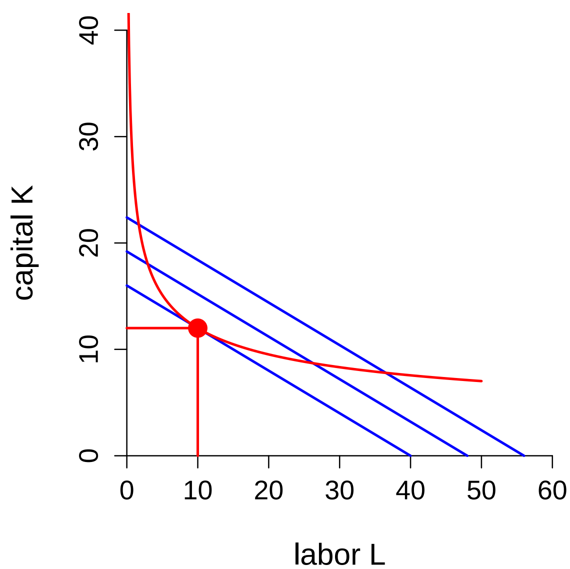


Figure 37: Cost minimization

Our discussion of cost functions can be short. The problem of finding on a given isoquant the cheapest combination of L and K is exactly analogous to the expenditure minimization problem of consumer theory, where we sought on a given indifference curve the cheapest combination of consumption quantities.

In fact, the cost function is the exact equivalent of the expenditure function. Figure 37 is a familiar diagram that illustrates the necessary condition for cost minimization. Constant cost lines are lines the slope of which is the negative of the factor price ratio. Lower constant cost lines correspond to lower expenditure. The cost minimizing input combination is the one for which the MRTS equals the factor price ratio.

I show in Figure 37 only the case in which the isoquant is convex, and in which the optimal factor combination is interior. But corner solutions are possible, and the isoquants need not be convex. But these are cases that you have all already seen in consumer theory. Everything that we said for expenditure minimization in consumer theory also applies here.

Let us briefly consider the condition that the MRTS must equal the factor price ratio at interior solutions. It can be re-written as follows:

$$\begin{aligned}
\begin{array}{r}\text{marginal rate of}\\ \text{technical substitution}\end{array} &= \text{factor price ratio} \Leftrightarrow \\
\frac{\frac{\partial f}{\partial L}}{\frac{\partial f}{\partial K}} &= \frac{w}{r} \Leftrightarrow \\
\frac{\frac{\partial f}{\partial L}}{w} &= \frac{\frac{\partial f}{\partial K}}{r} \Leftrightarrow \\
\begin{array}{r}\text{marginal product per \$1}\\ \text{spent on labor}\end{array} &= \begin{array}{l}\text{marginal product per \$1}\\ \text{spent on capital}\end{array}
\end{aligned}$$

Thus, we have exactly the analogue of the condition in consumer theory that the marginal utility per dollar for each good must be the same.

One brief side-remark on factor price changes. Suppose that, say, the wage falls. If we ask *for a fixed output level* how the cost minimizing combination of input factors changes, then we are moving along an isoquant, and therefore, in the language of consumer theory, we only have a substitution effect, but no income effect. Therefore, we can determine the sign. For example, if the wage goes down, labor input increases. Another example: if the price r of capital increases, the capital input decreases. But note that this is only true for a fixed output level. In response to a change in factor prices, the firm may of course also wish to change how much it produces. This will be discussed in the next section. Figure 38 illustrates the effect of a factor price change.

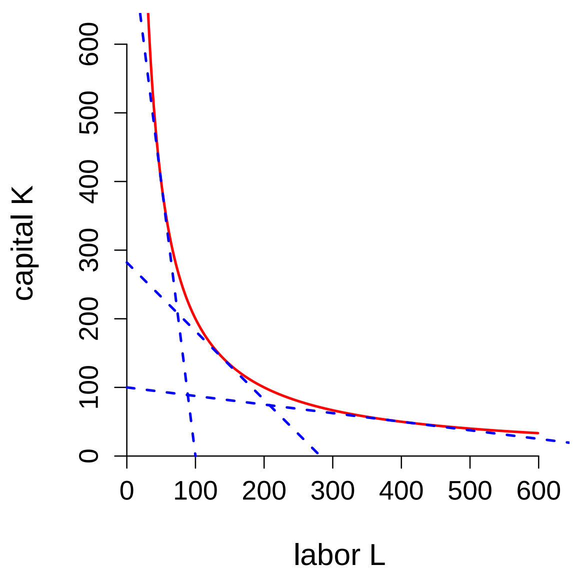


Figure 38: Labor input increases as the wage falls, and therefore the constant cost lines become flatter.

Now recall that we are interested in the cost function C that tells us for *every* quantity q how much it costs to produce that quantity. To find C we determine for every q the cost minimizing combination of labor and capital that should be used to produce q, then we find how much we have to pay for that combination, and then this expenditure is $C(q)$. If we connect the expenditure minimizing input combinations for different output levels, we obtain a curve that one calls the expansion path. An example is shown in Figure 39, where the expansion path happens to be a straight line. It need not be a straight line. The cost function is obtained by evaluating cost along the expansion path.

A Numerical Example

Suppose the production function were: $f(L, K) = L^{\frac{1}{3}} K^{\frac{1}{3}}$, and the factor prices were: $w = 1$ and $r = 8$. We shall determine the cost function. We can easily calculate that:

$$\text{MRTS}=\frac{K}{L}.$$

I omit this calculation. We have a Cobb-Douglas production function, and we have done similar calculations many times in this class. Note that the MRTS is decreasing as we increase L and decrease K, so that we have convex isoquants.

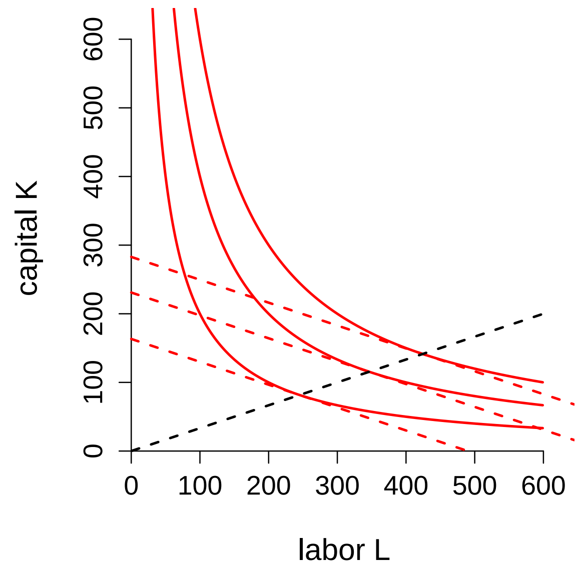


Figure 39: The Expansion Path

Moreover, the isoquants do not ever intersect with the axes.

The Cobb Douglas production function implies that nothing can be produced if you have zero of one factor, however much you have of the other factor.

The cost minimizing combination of labor and capital is thus determined by the necessary condition:

$$\text{MRTS= price ratio} \Leftrightarrow \frac{K}{L} = \frac{1}{8} \Leftrightarrow K = \frac{1}{8}L.$$

Our other necessary condition is that we have to be on the isoquant that corresponds to the quantity q that we are considering:

$$f(K,L) = q \Leftrightarrow L^{\frac{1}{3}}K^{\frac{1}{3}} = q.$$

Substituting the first condition into the second, we find:

$$L^{\frac{1}{3}}\left(\frac{1}{8}L\right)^{\frac{1}{3}} = q \Leftrightarrow L^{\frac{2}{3}}\frac{1}{2} = q \Leftrightarrow L = (2q)^{\frac{3}{2}},$$

and substituting back, we get:

$$K = \frac{1}{8}(2q)^{\frac{3}{2}}.$$

We have thus found for given q the optimal choice of L and K that minimizes cost. To find the cost function itself, we have to multiply L by w, and K by r, and add up:

$$C(q) = L + 8K = (2q)^{\frac{3}{2}} + 8\frac{1}{8}(2q)^{\frac{3}{2}} = 2(2q)^{\frac{3}{2}}.$$

Marginal and Average Cost

Two concepts derived from the cost function play an important role in the firm's profit maximization problem that we study in the next section. One is the concept of the marginal cost. We denote the marginal cost by MC. This is defined as:

$$\mathrm{MC}(q) = \frac{dC}{dq}.$$

Thus, the marginal cost represent the extra cost of producing one (infinitesimally small) extra unit of output. The second concept is the concept of the average cost, denoted by AC.

$$\mathrm{AC}(q) = \frac{C(q)}{q}.$$

Thus, the average cost represent the cost per unit produced. In our numerical example we have:

$$\begin{aligned} C(q) &= 2(2q)^{\frac{3}{2}}, \\ \text{and therefore} \quad \mathrm{MC}(q) &= \frac{dC}{dq} = 6(2q)^{\frac{1}{2}} \\ \text{and} \quad \mathrm{AC}(q) &= \frac{C(q)}{q} = 2(2)^{\frac{3}{2}} q^{\frac{1}{2}}. \end{aligned}$$

These are already simplified expressions. I skip the calculations. Hopefully, you can verify them by yourself.

There is something interesting about this calculation. We find that average cost are smaller than marginal cost. There is an interesting intuition for this. But first let's check that it is true:

$$\mathrm{AC}(q) < \mathrm{MC}(q) \Leftrightarrow 2(2)^{\frac{3}{2}} q^{\frac{1}{2}} < 6(2q)^{\frac{1}{2}} \Leftrightarrow 2 < 3.$$

Here, in the last step, I have divided both sides by $2 \cdot 2^{\frac{1}{2}} \cdot q^{\frac{1}{2}}$. Now why is it that average cost are less than marginal cost? The key is that in this example the marginal cost increase as q gets bigger. You can verify this in the expression for marginal cost. Now let us see the consequences of increasing marginal cost for average cost. Let's take a simple example in which units are discrete. The analogue of marginal cost is then the cost of producing one extra unit where the extra unit is not "infinitesimally small." Suppose the first unit to produced costs 3 Dollars, and the second unit produced costs an extra 4 Dollars. What are the average cost? If only one unit produced, the average cost is, of course, 3 Dollars, but if two units are produced, the average cost are $(3+4)/2 =$ 3.5 Dollars, and thus less than the marginal cost of the second unit, which is 4 Dollars. Thus, the average cost, 3.5 are less than the marginal cost, 4 Dollars. If marginal cost continue increasing, average cost will remain below the marginal cost. This is because the average cost are the average of the marginal cost of the first, the second, the third, ..., the nth unit, and therefore less than the largest term in this sum. But the nth term is the largest term in the sum. Thus, for any given quantity, the average cost of producing this quantity is smaller than the marginal cost of producing the last unit in that quantity.

Notice that, although average cost are below the marginal cost, they are also increasing. This is because every extra unit that is added to the average is more expensive than the previous one.

Increasing marginal cost reflect that it becomes more and more expensive to raise production by one more unit, perhaps because the extra hours of work become less and less productive, but all cost the same. But marginal cost may also decrease. For example, it seems plausible that perhaps when a firm has increasing returns to scale the marginal cost decrease. This is not precisely true, and we will discuss the details of this in the next section. For the moment let's consider one more example. Let's look at the cost function shown in Figure 40. For that cost function, I show in Figure 41 the average and the marginal cost.

The example shown in these two figures is popular in many elementary microeconomics textbooks. I am not sure that there is much evidence to show that this is what actual firms' cost curves look like in the real world. We shall treat it as just another example. The main point about this example is that the marginal cost curve in Figure 41 is U-shaped. You'll see that. This means that marginal cost are first decreasing, and then increasing. What does this imply for average cost?

Let's look in some detail at Figure 41. Note that average and marginal cost

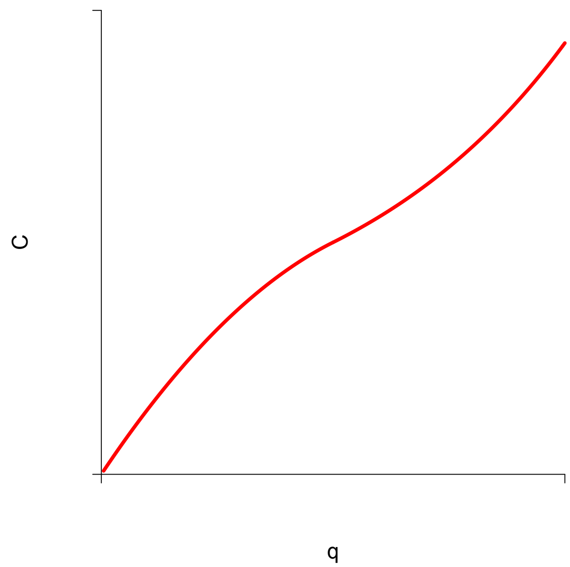


Figure 40: A Cost Function

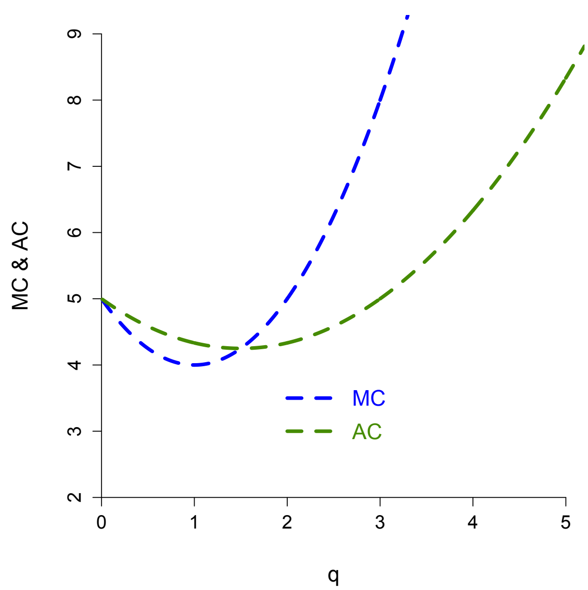


Figure 41: Average and Marginal Cost

are the same when $q = 0$. Intuitively, this corresponds to our observation earlier that when you produce just one unit, then average and marginal cost are the same. When quantity is continuous, this argument applies to the case when q is (approximately) zero.

Next, we observe that as marginal cost are decreasing, average cost are above marginal cost. This is true for exactly the same argument that we used above to show that when marginal cost are increasing, the average cost are below the marginal cost. We just have to reverse that argument.

In fact, even when marginal cost are increasing, initially, the average cost are above the marginal cost. This is because the average cost still average over the earlier, higher marginal cost. At some point then, as marginal cost are increasing, though, the marginal cost catch up with the average cost, and the two curves intersect.

Up to the intersection point, average cost are decreasing. This is because every extra unit added has smaller cost than the average unit produced that far, and therefore it pulls the average down.

After the intersection point, the average cost are below the marginal cost, and both are increasing. This is similar to our numerical example, that we discussed earlier.

Note that the average cost are minimized when they equal the marginal cost. We can show this mathematically. Recall that average cost are given by:

$$AC(q) = \frac{C(q)}{q}.$$

When they are minimized, the first derivative of $AC(q)$ has to be zero. Let us calculate that first derivative using the quotient rule, and set it equal to zero:

$$AC'(q) = \frac{qC'(q) - C(q)}{q^2} = 0.$$

The fraction in this equation is zero if and only if the numerator is zero. Let's set the numerator equal to zero:

$$qC'(q) - C(q) = 0.$$

Re-arranging this equation, we get:

$$C'(q) = \frac{C(q)}{q}.$$

This says exactly that marginal cost, on the left hand side, equal average cost, on the right hand side. Thus, in the minimum of average cost, average cost must equal marginal cost.

When are Marginal Cost Increasing?

A firm has increasing marginal cost if as production increases every additional ("infinitesimally small") unit of output cost more than the previous one. In other words, it gets more and more costly to expand production. But how can this be? If, for example, the firm has constant returns to scale, that is, it can multiply output by some other factor $\lambda > 1$ if it uses λL units of labor, and λK units of capital, it seems intuitive, and indeed one can prove, that marginal cost cannot be increasing. Indeed, one can show that if the production function has constant returns to scale, then the cost function must be linear:

$$C(q) = cq \text{ for some } c > 0.$$

In this case, of course the marginal cost are equal to c, and thus are constant; they don't depend on q.

Are marginal cost increasing if the production function has decreasing returns to scale? Not necessarily. The reason is that, as the firm tries to expand production by a factor λ, there are may be intelligent ways of doing that than multiplying both factors by some constant. Maybe the firm could change the mixture of capital and labor that it is using, and thereby reducing marginal cost? But returns to scale are only about what happens when all inputs are multiplied by the same number, and thus the input proportions stay the same.

There is one special case in which decreasing returns to scale do indeed imply increasing marginal cost. This special case is the case that there is some number $r < 1$ such that

$$f(\lambda L, \lambda K) = \lambda^r f(K, L), \text{ for all } \lambda > 1K, L.$$

This is a special case of decreasing returns to scale because. We have decreasing returns to scale because, for $\lambda > 1$, and $r < 1$, we have that $\lambda^r < \lambda$. In this special case we say that the production function is homogeneous of degree r. For such production functions, if $r < 1$, it is indeed true that marginal cost are increasing. The following proposition can be proved:

If there are decreasing returns to scale, and the production function is homogenous of degree $r < 1$, then marginal cost are increasing.

I shall not give here a mathematical proof of this proposition. But I shall briefly explain the intuition. Start with the following thought experiment. The firm currently produces output q with factor combination $\bar{L}$ and $\bar{K}$. Then, the marginal rate of technical substitution is:

$$\frac{\frac{\partial f(\bar{L},\bar{K})}{\partial L}}{\frac{\partial f(\bar{L},\bar{K})}{\partial K}}.$$

Now suppose that the firm raised all inputs by proportionally the same factor λ, thus using now $\lambda\bar{L}$ units of labor, and $\lambda\bar{K}$ units of capital. What happens to the marginal rate of technical substitution? Well, let's calculate numerator and denominator for the marginal rate of substitution separately. We begin calculating the marginal product of labor:

$$\frac{\partial f(\lambda\bar{L},\lambda\bar{K})}{\partial L} = \frac{\partial \lambda^r f(L,K)}{\partial L} = \lambda^r \frac{\partial f(\bar{K},\bar{L})}{\partial L}.$$

Here, in the first equality, we use the homogeneity of degree r of the production function. The second step just applies the rules of calculus. We do the same calculation for the marginal product of capital:

$$\frac{\partial f(\lambda\bar{L},\lambda\bar{K})}{\partial K} = \frac{\partial \lambda^r f(L,K)}{\partial K} = \lambda^r \frac{\partial f(\bar{K},\bar{L})}{\partial K}.$$

Next, we divide the two marginal products, to find the marginal rate of technical substitution:

$$\frac{\frac{\partial f(\lambda\bar{L},\lambda\bar{K})}{\partial L}}{\frac{\partial f(\lambda\bar{L},\lambda\bar{K})}{\partial K}} = \frac{\lambda^r \frac{\partial f(\bar{K},\bar{L})}{\partial L}}{\lambda^r \frac{\partial f(\bar{K},\bar{L})}{\partial K}} = \frac{\frac{\partial f(\bar{L},\bar{K})}{\partial L}}{\frac{\partial f(\bar{L},\bar{K})}{\partial K}}.$$

Thus, the marginal rate of technical substitution, after we have multiplied both inputs by λ, is exactly the same as it was before.

In geometric terms, the factor combinations $(\lambda\bar{L},\lambda\bar{K})$ all lie on the same line through the origin with slope $\frac{\bar{K}}{\bar{L}}$. Thus, along this line, the marginal rates of technical substitution, and hence the slope of the isoquants, are all the same.

Now suppose the ratio of factor prices were given: $\frac{w}{r}$. The firm aims to produce any quantity at a point where the slope of the isoquant is equal to this ratio. This is how it minimizes cost. But as the slope is the same along a straight line through the origin, the firm will always choose the input combination as some point along this line. In other words, as the firm expands output and minimizes cost, it will hold the ratio of labor and capital constant, and multiply both inputs by the same constant. The expansion path is linear.

Decreasing returns to scale now means that, along this linear expansion path, the output grows more slowly than the inputs. Therefore, for given and fixed factor prices, as we grow output along this expansion path, we have increasing marginal cost. This concludes my attempt to explain the intuition for the result that if the production function is homogeneous of degree $r < 1$, then marginal cost are increasing.

There is also another form of homogeneity. We say that it a production function homogenous of degree $r > 1$ if:

$$f(\lambda L, \lambda K) = \lambda^r f(K, L), \text{ for all } \lambda > 1.$$

Homogeneity of degree $r > 1$ is a special case of increasing returns to scale.

Topic 13: Profit Maximization and Supply

The firm seeks to maximize profits, that is, its revenue, i.e. price times quantity, minus its costs, that is, its expenses on inputs: wage times labor plus rent times capital. Let us denote profit by π:

The Greek letter "π" is pronounced as "pi."

$$\pi = pq - wL - rK.$$

The firm chooses q, L and K. But it has to respect these constraints:

$$L \geq 0, K \geq 0, \text{ and } q = f(L, K).$$

We want to study the solution to this maximization problem. The solution will depend on three prices: p, w, and r. When we solve the maximization problem we obtain the supply function, which shows us the optimal q, K and L as a function of the three prices. We call these the output supply and the factor demand of the firm.

There are several ways of approaching this maximization problem. We shall begin with an approach that relies on the cost function derived in the previous section.

Optimal Supply when Marginal Cost are Increasing

Let us focus on the optimal choice of the output quantity q. Assuming the firm chooses for given level of q the cheapest factor combination that let's the firm produce q, then the firm's choice problem is to choose a $q \geq 0$ that maximizes:

$$\pi(q) = pq - C(q).$$

We have used the same symbol π as before for profits, emphasizing that π now only depends on q, and not on L and K. In comparison to our previous formulation, we have now replaced the expression $wL + rK$ by $C(q)$.

The familiar first order condition for maximizing a function of one variable says that the first derivative of π with respect to q must be zero:

$$\pi'(q) = 0.$$

Calculating the first derivative, we get:

$$p - C'(q) = 0 \Leftrightarrow p = C'(q).$$

This first order condition has the simple interpretation that price must be equal to marginal cost. If price were above marginal cost, it would pay to increase q a little bit, and if price were below marginal cost, it would pay to reduce q a little bit.

Of course, this is only a necessary condition. Moreover, being more precise by taking care of the potential boundary solution $q = 0$, we replace the above first order condition by:

$$\text{if } q = 0: \quad p \leq C'(q) \quad \text{and} \quad \text{if } q > 0: \quad p = C'(q).$$

The left hand side says that quantity can be zero if the price is below the marginal cost at zero. In that case, there is no point in producing anything

These conditions are necessary and sufficient if the profit function is concave. This means that the second derivative of profit must be non-positive:

$$\pi''(q) \leq 0.$$

Taking the second derivative of the profit function, we see that it is minus the second derivative of the cost function, and therefore, we get:

$$-C''(q) \leq 0 \Leftrightarrow C''(q) \geq 0.$$

Thus, our conditions are necessary and sufficient if marginal cost are non-decreasing, that is, if the firm's cost function is convex. This is the reason why, in the previous topic, we spent some time discussing sufficient conditions for the production function that imply that marginal cost are non-decreasing.

In Figure 42 I show a case of increasing marginal cost. The quantity is on the horizontal axis, and the marginal cost are on the vertical axis. The marginal cost curve is shown in blue. To determine the optimal supply at a given price, we can locate the price on the vertical axis, and then draw a horizontal line until we hit the marginal cost curve. The corresponding quantity is the optimal supply. We can draw a curve that shows for every possible price the corresponding optimal output quantity. This curve is called the firm's supply curve.

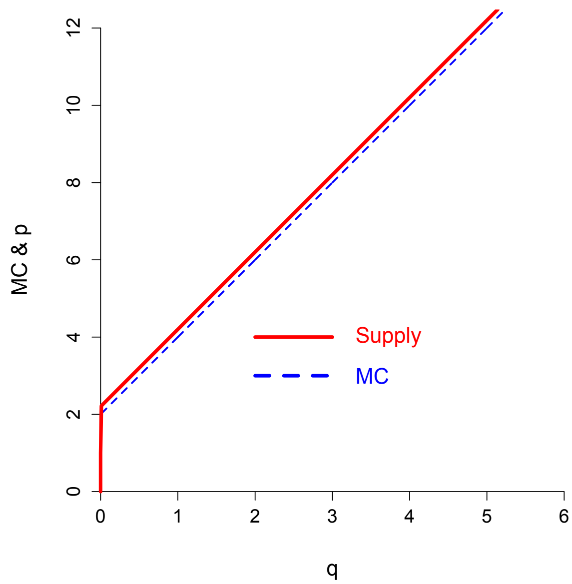


Figure 42: Increasing marginal cost and supply

But note that this means that the supply curve is equal to the marginal cost curve. I have drawn the supply curve in Figure 42 as a red line. It is geometrically the same curve as the marginal cost curve. There is a difference in interpretation, though. For the marginal cost curve, the independent variable is on the horizontal axis, and the dependent variable is on the vertical axis. For the supply curve, the independent variable, the price, is on the vertical axis, and the dependent variable, the quantity, is on the horizontal axis. The only point where the two curves don't coincide is when supply is zero. This happens when the price is below the marginal cost at zero. For those prices, the supply curve appears in Figure 42 as a vertical red line.

Let us go back to our numerical example from Topic 12. In that example, we had found the marginal cost function:

$$\mathrm{MC}(q) = 6(2q)^{\frac{1}{2}}.$$

This is increasing in q, and therefore we have increasing marginal cost. Note that the marginal cost at zero are zero in this case. Therefore, *for all prices*, optimal supply is determined by the condition that price equals marginal cost. We obtain:

$$\begin{aligned} p &= 6(2q)^{\frac{1}{2}} \Leftrightarrow \\ (2q)^{\frac{1}{2}} &= \frac{p}{6} \Leftrightarrow \\ 2q &= \frac{p^2}{36} \Leftrightarrow \\ q &= \frac{p^2}{72}. \end{aligned}$$

Here, we have solved for optimal supply for arbitrary price of the output q, but for specific input prices. We had assumed in Topic 12 in our calculation that $w = 1$ and that $r = 8$. Had we calculated the cost function for general values of w and r, we would have obtained here the optimal supply as a function of p, w and r.

What about factor demand? We don't only want to know what the firm's optimal output is, but also which inputs it is going to use. But recall form Topic 12 the optimal input combination:

$$L = (2q)^{\frac{3}{2}} \text{ and } K = \frac{1}{8}(2q)^{\frac{3}{2}}.$$

Now that we have found the optimal q, we could plug our result into these formulas, and get labor and capital demand. It would be a tedious calculation, and we skip it here. Observe that we would get L and K only as a function of p, the output price, not as a function of the factor prices w and r. Again, this is because we assumed specific values for w and r in Topic 12. Had we calculated the cost function for general values of w and r, we would have obtained here the factor demand as a function of p, w and r.

Optimal Supply when Marginal Cost are U-shaped

Let us now give a little bit of attention to the case that marginal cost are not increasing. Suppose that we have collected all the solutions to the necessary conditions. What can we say about those solutions? I'll make two simple points.

The first is that if a solution satisfies:

$$C''(q) < 0,$$

then it can not be an optimum. The reason is this: $C''(q) < 0$ means that the marginal cost are decreasing. Whereas the given point satisfies the necessary condition, i.e. price equals marginal cost, it will be possible to increase profit by expanding production, because in this case marginal cost fall, and therefore, for every additional unit of output, we earn more than this unit costs. Thus, if $C''(q) < 0$, a solution to the necessary conditions cannot be a maximum.

The second simple observation is as follows. If a solution to the necessary conditions leads to negative profits (which is possible!, see below), then a better solution is to choose $q = 0$, because then profits will be zero. Thus, such a solution to the necessary conditions cannot be a profit maximum. When does a solution to the necessary conditions lead to negative profits? Well, if:

$$pq - C(q) < 0 \Leftrightarrow p < \frac{C(q)}{q} = AC(q),$$

that is, if the price is below average cost.

These are two very elementary observations. They are not complete instructions for how to solve the case when marginal cost are not increasing. But they give some first insight. We shall now consider a very special case in which marginal cost are not increasing, namely the case in which marginal cost are U-shaped, as in Figure 6 and 7. Consider prices above the marginal cost at zero. At those prices, there is exactly one quantity that satisfies the first order condition. At this quantity, price equals marginal cost, and this quantity is indeed the optimal choice of quantity, and thus supply, as indicated by the red curve in Figure 43.

Next, consider very low prices, specifically, prices that are even lower than the minimum of marginal cost. For those prices, only the quantity zero solves the necessary conditions for a profit maximum, and indeed this quantity is the optimal choice. I have indicated in Figure 43 with the red line that is vertical that supply at these prices is zero.

Now consider prices below the marginal cost at zero, but above the minimum of marginal cost. For those prices, exactly three quantities satisfy the necessary conditions for profit maximization, namely, the quantity zero, and

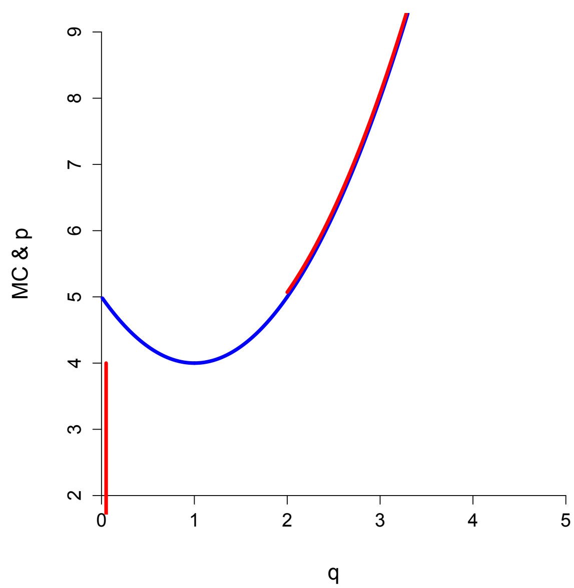


Figure 43: Optimal Supply When Marginal Cost are U-Shaped (MC in blue, Supply in red)

the two intersection points of a horizontal line through the price, parallel to the q-axis. This horizontal line intersects the U-shaped marginal cost in two points. These two points correspond to the two additional solutions of the necessary conditions. I show the situation in Figure 44.

At the smaller of the two intersection points, marginal cost are decreasing, and, as our discussion at the beginning of this section showed, this cannot be an optimum. This leaves for us to choose as optimal solutions either the quantity 0, or the larger quantity at which price equals marginal cost. But which one is optimal?

As we mentioned earlier, it depends on whether price is larger, or less, than average cost at the larger quantity at which the horizontal price line intersects the marginal cost curve. Therefore, to determine optimal supply, it will be helpful to enrich Figure 44 and to draw also an average cost curve into the figure. I show the resulting graph, and the optimal supply, in Figure 45.

Which of our two candidates for optimal supply, zero, and the second of the two points where price equals marginal cost maximizes profits? The latter point is optimal if at that point price is at least average cost. Otherwise zero is optimal. The resulting supply curve is shown in Figure 45.

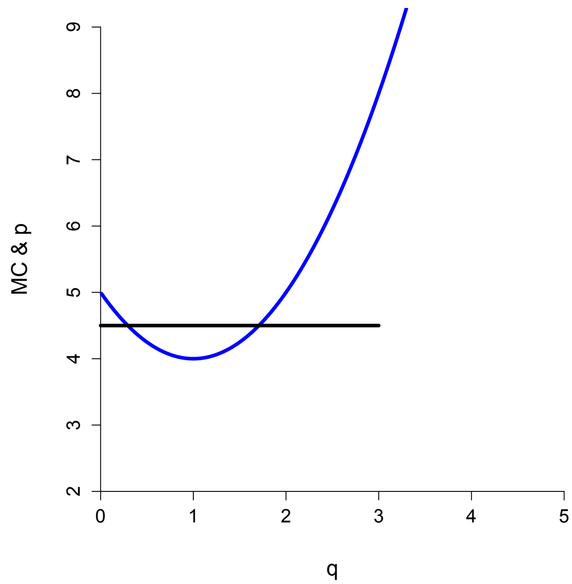


Figure 44: Three Solutions To The Necessary Conditions (MC in blue, horizontal line through the price in black.

In the range in question, zero is the optimal choice if and only if the price is below the minimum of average cost. One can see that it is exactly then that the price, at the second intersection point of price and marginal cost, is below average cost. If the price is higher, the second intersection point is the optimal supply. In summary, if the price is below the minimum of average cost, supply is zero. If it is above the minimum of average cost, then the supply curve coincides with the marginal cost curve.

Note that we have a jump in the supply function. It occurs at the minimum of average cost. If the price is below that minimum, then supply is zero. It is above that minimum, supply is positive, and indeed it is significantly above zero. This reflects the idea that, if the firm produces at all, it needs to produce at a minimum scale. This seems plausible in practice, and it is perhaps for this reason the the case of U-shaped marginal cost curves is so popular in the textbooks.

We have focused on the determination of the optimal q, but what about L and K? Well, to construct the cost function we always need to first find the optimal input combinations for every output value q. We have not done this in this section, but instead just assumed that the cost function was given. But, had we derived it, we would be able to substitute the optimal supply q into our

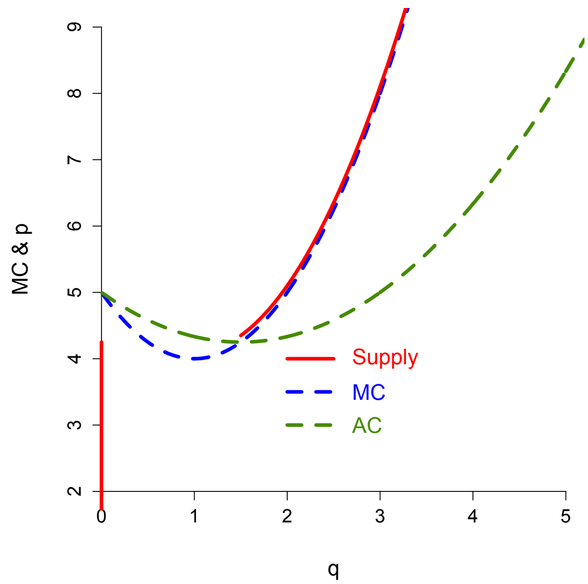


Figure 45: Optimal Supply When Marginal Cost are U-Shaped (MC in blue, AC in green, supply in red)

results, and find labor and capital demand.

An Alternative Approach to Profit Maximization

Recall that the firm's profit maximization problem is to maximize:

$$\pi = pq - wL - rK.$$

subject to:

$$L \geq 0, K \geq 0, \text{ and } q = f(L, K).$$

We now consider an alternative approach to solving this problem, an approach that does not make use of the cost function. This approach eliminates the choice variable q from the problem. The choice variable q can be eliminated because once the firm has chosen K and L, the equation $q = f(K, L)$ determines the firm's choice of q. Mathematically, we can replace q by $f(L, K)$ in the expression for profits, and we obtain that the firm's problem is to maximize

$$\pi = pf(L, K) - wL - rK.$$

subject to the constraints:

$$L \geq 0 \text{ and } K \geq 0.$$

A familiar condition for profit maximization is that the first derivative of the function that we maximize should be zero. This condition is called a "first order condition." In our case, we have two choice variables: K, and L, and so, taking partial derivatives, we can write down two first order conditions:

$$\frac{\partial \pi}{\partial L} = 0 \text{ and } \frac{\partial \pi}{\partial K} = 0.$$

Let's calculate the two partial derivatives, and then re-arrange these conditions a little bit:

$$\begin{aligned} \frac{\partial \pi}{\partial L} &= p\frac{\partial f}{\partial L} - w = 0 \Leftrightarrow \frac{\partial f}{\partial L} = \frac{w}{p}, \quad \text{and} \\ \frac{\partial \pi}{\partial K} &= p\frac{\partial f}{\partial K} - r = 0 \Leftrightarrow \frac{\partial f}{\partial K} = \frac{r}{p}. \end{aligned}$$

Consider the first of these two conditions. On the left hand side, we have the marginal product of labor, $\frac{\partial f}{\partial L}$. On the right hand side, we have an expression that we encountered before, when discussing labor supply: $\frac{w}{p}$. Recall that we can interpret this expression as the "real" wage, that is, the wage, expressed in units of output rather than in Dollars. Therefore, our first condition says that the marginal product of labor must equal the real wage, and similarly our second equation says that the marginal product of capital must equal the real rent.

The two first order conditions have an intuitive interpretation. If the first condition, for example, were violated, then one unit of labor could produce more (or less) output than the firm would have to pay for that one unit of labor, as expressed output units. Therefore, the firm could raise profits by increasing (or reducing) the amount of labor it puts into the production process. The second condition has a similar interpretation.

We want to be a little more careful in studying these first order conditions. First, like all first order conditions, they are only necessary, not sufficient for profit maximization. Second, they are really necessary only if $K > 0$ and $L > 0$, because only then are small increases of labor or capital input, and also small decreases possible. Our intuitive argument for these $L = 0$, or $K = 0$, or both. That is, the cases in which the firm makes no use of one or the other of the two inputs. Therefore, if we are more careful, our claim is just that the necessary conditions for profit maximization are that:

$$\text{if L>0: } \frac{\partial f}{\partial L} = \frac{w}{p}, \quad \text{if L=0: } \frac{\partial f}{\partial L} \leq \frac{w}{p}, \qquad \text{and}$$

$$\text{if K>0: } \frac{\partial f}{\partial K} = \frac{r}{p}, \quad \text{if K=0: } \frac{\partial f}{\partial K} \leq \frac{r}{p}.$$

In summary, when the firm maximizes profits, it should only look at pairs of K and L that satisfy the two necessary conditions. Not all of them will maximize profits, but if there is a pair that maximizes profits, it will be among the pairs that satisfy these necessary conditions.

We shall not go into sufficient conditions as they apply to this approach. But let us observe that if the firm has chosen the optimal values of K and L, it can determine its optimal output simply from the equation $q = f(K, L)$, and thus, we have obtained the firm's supply function as well as the firm's demand function for labor and capital. Note that the necessary conditions only involve the ratios of the prices, specifically, the ratios $\frac{w}{p}$ and $\frac{r}{p}$. If we multiplied output and input prices by a constant, say 2, then the necessary conditions would not change. This is not just true for the necessary conditions, it is also true for the maximization problem overall: if we multiply all prices by a constant, then the solution to the profit maximization problem does not change. Intuitively, whether we maximize profit, or whether we maximize two times profit, we must find the same solution. Another way of putting this observation is: supply of output and demand for factors only depends on real wage, and real rent.

Firms' Supply and Demand Functions

When we studied consumer theory, we introduced explicitly the consumer's demand functions, that is, optimal consumption as a function of prices and income. We can do the same here. In our model, a firm has a demand function for two factors, labor L and capital K, and it has a supply function for output. All of these depend on the prices of the inputs, w and r, and on the output price p. Thus, we can write these functions as:

$$L(w, r, p) \text{ and } K(w, r, p) \text{ and } q(w, r, p).$$

Emphasizing that q is supply, we can also write for $q(w, r, p)$:

$$S(w, r, p).$$

Using the same ideas as in Topic 7, we can define various elasticities of factor demand and output supply.

Although producer theory is in many respects analogous to consumer theory, you will be relieved to hear that there is no analogue of the income effect in producer theory, simply because there is no concept of a given income of a firm. In consumer theory, the income effect could cause a demand function to be upward sloping, although it sounds plausible that demand functions should be downward sloping. For supply functions, it sounds plausible that they should be upward sloping. Because there is no income effect, this is indeed *always* true. More precisely:

> Supply $S(w, r, p)$ stays the same or goes up as p goes up, holding w and r constant.

The supply curves we constructed earlier in this topic showed how quantity changes as price varies. In the cases that we showed, we kept the cost function, and therefore implicitly w and r, fixed. You can check that all supply curves were either constant, or upward sloping.

Topic 14: The Law of Supply and Demand

Figure 46 shows what is perhaps the most famous diagram in economics: an upward sloping supply line, a downward sloping demand line, and their intersection marked as the equilibrium values of price and quantity. In Figure 46, the equilibrium price equals 6, and the equilibrium quantity also happens to be 6.

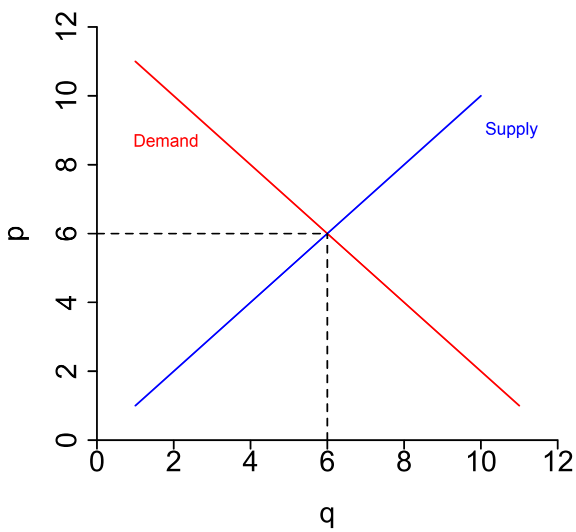


Figure 46: Supply and Demand

In Figure 46, the price is on the vertical axis, and the quantity is on the horizontal axis. Price is the independent variable: the quantity demanded and the quantity supplied are functions of the price. It is a little bit against the conventions of mathematics to put the independent variable on the vertical axis, but it has a long tradition in economics, and we shall follow this tradition here.

Figure 46 illustrates the law of supply and demand: The price that prevails in a market is the price where supply and demand are equal. Most readers will have seen the diagram in Figure 46 before. It has somewhat glorified status in economics. Our task here is to make the familiar look fresh again. I shall try to make clear in this section that the law of supply and demand is by no means obvious. I then want to raise and discuss the question how economists know that the law is true.

If it is true.

But first, let's clarify the precise meaning of the law. To start with we ask: how do the demand and supply functions in Figure 46 relate to the demand and supply functions that we studied in the previous chapters? Earlier, when considering a two good world, we derived, for example a demand function for good 1 that we denoted by $D_1(p_1, p_2, y)$, indicating that the demand for good 1 depends on the price of good 1, but also the price of good 2, and on income. A demand function like the one in Figure 46 shows demand for some particular good, say good 1, as a function of the price of that good only, that is, keeping the other prices, say p_2, and income, y, fixed. Moreover, Figure 46 shows not just one consumer's demand function, but it shows the sum of all consumer's demand, that is, it shows for every price of a given good the sum of demand from consumer 1, consumer 2, up to consumer 250,000,000, if that is how many consumers buy in this market.

The same for supply: Earlier, we had determined the supply of a good as a function of that good's price p, but also of the wage w, and the price r of capital. If we fix w and r, then the supply only depends on price. That is what the figure shows. Moreover, it shows not just the supply from one firm, but from all firms that operate in the market to which the figure applies.

So, what is the claim of the law of supply and demand? Suppose we look at any particular market, say the market for some particular type of light bulb. In the real world, we find stores selling these light bulbs at some particular price, and we find people buying them. What the law of supply and demand claims is that there is a simple regularity behind the price that people pay for light bulbs. The regularity is this: Suppose you went to buyers, and asked every buyer, for every conceivable price, how many light bulbs they would like to buy if this was the prevailing market price. You need to make clear when you ask the question that the person should regard the price as independent of her choice, and that she should feel free to plan to buy as many or as few light bulbs as she wants, as if there were no constraints. Sum the numbers of lightbulbs that people want to buy for every price, and draw the demand curve.

Do the same for sellers, asking them for every price how many lightbulbs they would want to sell at that price, and draw the supply curve. Note that we have constructed completely hypothetical functions. The claim of the law of supply and demand is that, if we draw the graphs of these functions, and check where they intersect, then, magically, they intersect exactly at the price that we observe as the actual price in the market, and at this price people buy and sell exactly what they have told us they would like to buy or sell. If supply and demand were as in Figure 46, then all sellers of light bulbs would charge price 6, and the quantity of light bulbs sold and bought, would also happen to be exactly equal to 6.

It would be an interesting discovery, and a little magical indeed, if we could do this for a variety of markets, and if we discovered, each time, that it worked. We would have discovered a hidden regularity behind one of the most common human activities, and a regularity that most of us would not be aware of if nobody pointed it out to us. It would be as if an invisible hand guided buyers and sellers like marionettes to the equilibrium.

The metaphor of the market as an "invisible hand" is attributed to Adam Smith, a philosopher of the Scottish enlightenment.

But is it really true? Which observations have we made that convinces us that this is true? If it is a natural law of human life, then presumably it is supported by much evidence, not just by its superficial plausibility. Physicists don't just believe the law of gravity because it sounds plausible - they believe it because you can measure gravity, and find the law confirmed again and again. What would we measure to verify, or falsify, the law of demand and supply?

We would not really have to ask all those hypothetical questions listed in above, and construct complete demand and supply curves. Actually, it would be enough to ask people whether, at the price at which they all trade, they can trade exactly the quantity that they would want to trade if they assume that they can trade any arbitrary quantity at this price without affecting the price level. If everyone affirms this, then we are in an equilibrium of supply and demand.

There are still some problems, even with this simplified procedure. In most markets, prices are not quite the same in every transaction. So, we would not quite know which price to pick. But even if we knew, would we really just ask people, and trust that they give us the correct answer? The large light bulb makers might not even have ever considered this question. They might always have taken it for granted that they themselves choose the price. Our claim is that, although they think they choose it, if they examined their own choices, they would discover that they can be predicted by the the law of supply and demand. But how could we persuade them to consider, and honestly address, the fictional scenario that we want them to consider, i.e., a world in which they have to take prices as given.

It is perhaps not surprising, then, that there are very few papers in the economics literature that I could point in which the law of supply and demand would be subjected to a rigorous test. Of course, in every day conversation people, including economists, often claim that the price of such and such a good went up because, once again, the law of supply and demand was operating. But that is anecdotal evidence at best. Perhaps, we see the law of supply and demand only because we want to see it, and we don't have many alternative mental models. In short, how do we know that the most fundamental law of economics, that of supply and demand is indeed such a universal regularity of human interactions in the economic domain?

Experimenting With the Law

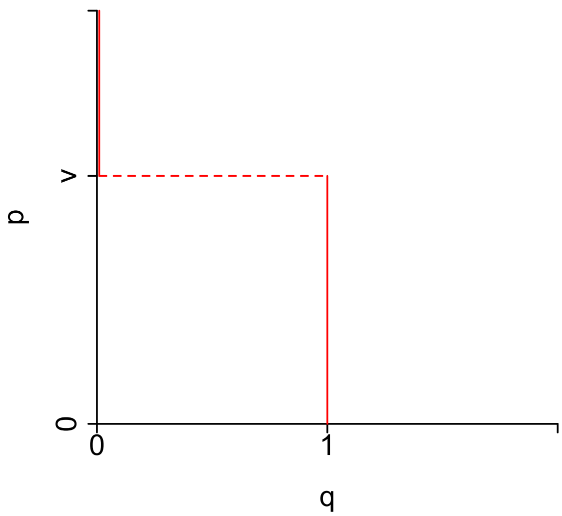
To investigate whether the law of supply and demand predicts well what people do in the real world economists have often resorted to choice experiments with human subjects. You may think of these experiments as if they were psychology experiments. The subjects of these experiments are often students, just because the experimenters work in universities, and because students don't yet have to be paid a lot for giving up their time and participate in an experiment. The students are often invited into something like a computer laboratory. They are given instructions, and then invited to make choices. In laboratories the environment can be a little bit better controlled than in classrooms. But some experiments are also simply conducted in classrooms.

We'll now recount an experiment specifically designed to test the law of supply and demand. Subjects in this experiment are divided into two groups: buyers and sellers. The market is for a fictional good, but we'll refer to this good as something really existing, just for concreteness. We'll settle for calling the good "apples."

Because the good is fictional, the subjects don't really have demand for it, or a supply. Instead, the demand, or the supply, are artificially created by the experimenter. And this is an *advantage* of the experiment. Because, if the experimenter creates demand and supply curves, then the experimenter knows them, and he can check, after the experiment, whether trading outcomes confirm the law of supply and demand.

It would be too complicated, however, to create demand, or supply curves of the continuous form that we have studied so far. Experiments have to be simple. Instead of saying that a buyer demands a quantity q of apples, where q is a real number, a buyer either demands either 1 or 0 apples in the experiment. No other quantity is allowed. What is the demand curve that we are creating for buyers? Well, if the “curve” is “downward sloping,” it has to be that at high prices the buyer does not want to buy an apple, and at low prices the buyer does want to buy an apple. Thus, for prices above some threshold demand will be zero, and for prices below this threshold demand will be one. Let us call the threshold “v.” It represents the value the buyer attaches to the apple: the maximum amount he is willing to pay. The demand “ curve” is shown in Figure 47.

Figure 47: Single Unit Demand

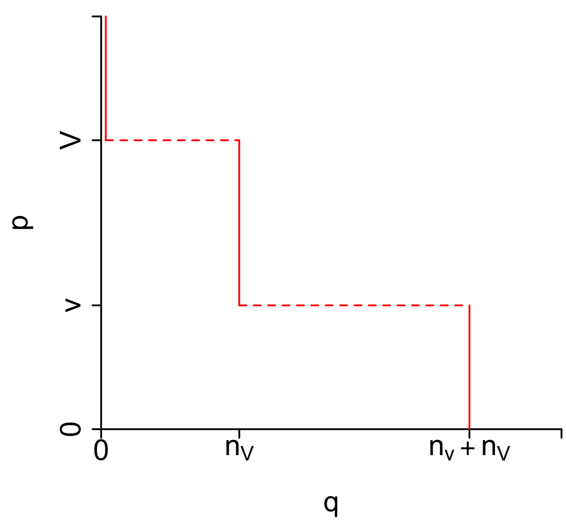


How can the experimenter induce a buyer to have this demand curve? The method used in the experiments that I am referring to here is to make a promise to the buyer before the experiment: “If you buy an apple in this artificial market, then you can bring the apple after the experiment to me, the experimenter, and I will pay you an amount of v Dollars.” If he trusts this promise, the buyer knows exactly what the value of the apple is to her, and that it is worth paying up to v Dollars for an apple, but not more. She has the demand function that we are trying to achieve.

Of course, we want many buyers to be in our market. Therefore, we need many subjects in the experiment to whom we will assign the role of buyers, all of whom have instructions of the form described above. Things are only interesting if not all buyers have the same value v. Let’s make it very simple:

Let's assume that every buyer either has the low value v, or the high value V, where thus $V > v$. Let's suppose we assigned the low value v to n_v buyers, and the high value V to n_V buyers. Then at prices below v, both types of buyers would want to buy an apple, and therefore the number of apples demanded would be $n_v + n_V$. At prices between v and V, only the buyers who have been assigned a high value would want to buy an apple. Therefore, the number of apples demanded would be n_V. At prices above V, demand would be zero. In Figure 48, I show you the market demand curve that I have just described in words.

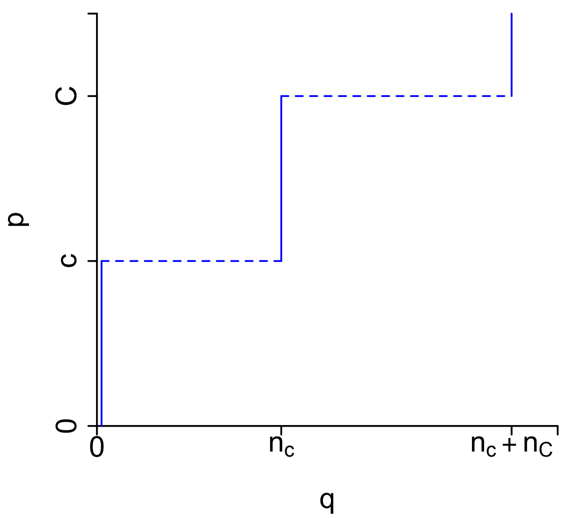
Figure 48: Market Demand



Note the horizontal dashed lines in Figure 48. These correspond to the case in which the price is exactly equal to V, or equal to v. In those cases, there are groups of buyers in the market, namely the high value or the low value buyers, who are indifferent between buying and not buying an apple. For example, if the price is equal to V, then all high value buyers are indifferent between buying and not buying. Therefore, at this price, the number of buyers can be anything between 0 and V. Thus, the dashed lines show indifference.

Our market also needs a supply side. We assign to some subjects the role of an apple seller. An apple seller can "grow" at most one apple. The cost of growing an apple are c for an apple seller with low cost, and C for an apple seller with high cost. Apple sellers are told: "if you find a buyer with whom you can agree on a price, then I, the experimenter, will pay you the difference between that price, and your cost." This induces a supply function where an apple seller is unwilling to sell at a price below cost, but is willing to sell an apple at a price above cost. Let us suppose the number of apple sellers with low cost is n_c, and the number of apple sellers with high cost is n_C. We can then construct the supply curve that is shown in Figure 49. Dashed lines again show points of indifference.

Figure 49: Market Supply



With these, somewhat unfamiliar looking supply and demand curves, we can now find the equilibrium price. I shall focus on two numerical examples. In both examples I set: $v = 20$, $V = 40$, $c = 10$ and $C = 30$. In the first example, I also choose $n_v = 16$, $n_V = 8$, $n_c = 15$ and $n_C = 8$. In the second example, I choose $n_v = 8$, $n_V = 16$, $n_c = 8$ and $n_C = 15$. Supply and demand diagrams for both cases are shown in Figures 5 and 6. In both figures, I have marked the market equilibrium with a circle.

We can see in Figure 50 that in Example 1 the equilibrium price is 20, and the equilibrium quantity is 15. At price 20, all 15 sellers with cost $c = 10$ want to sell, but the 8 sellers with cost $C = 30$ do not want to sell. Thus, the quantity supplied is 15. What is demand? All 8 buyers with value $V =$

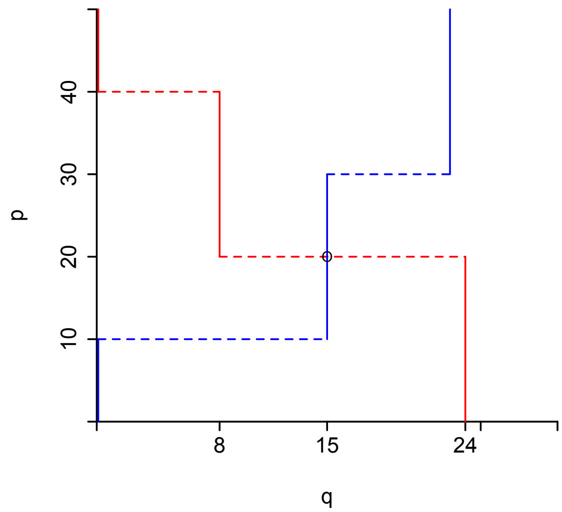


Figure 50: Market Equilibrium in Example 1

40 clearly want to buy. But what about the 16 buyers with value $v = 20$? The price is exactly equal to their valuation. Whether or not they trade at the equilibrium price, their net utility is zero. They are indifferent between trading and not trading. But the only equilibrium is that exactly 7 of these buyers trade, and the remaining 9 buyers do not trade. Only with this split is demand, namely 8+7, exactly equal to supply, namely 15. How do the buyers know which of them are supposed to trade, and which are not? We'll remain silent on this question, and just accept the equilibrium as it is.

Figure 51 for Example 2 shows that in that market the equilibrium price is 30, and the quantity traded in equilibrium is 16. all 16 buyers with high values want to buy. All 8 sellers with low cost want to trade. But sellers with high cost are indifferent between selling, and not selling. Exactly 8 of them sell, and the remaining 7 do not sell.

I have given these two examples because for these particular scenarios, experiments have actually been conducted, and we know what happens in those markets. The rules were as follows: every buyer was told a value. Every seller was told a cost. Everyone knew only their own value or cost, and not anybody else's values or costs. Thus, nobody knew the demand curve, or the supply curve, or the equilibrium price. Therefore, it would be magical if people did in-

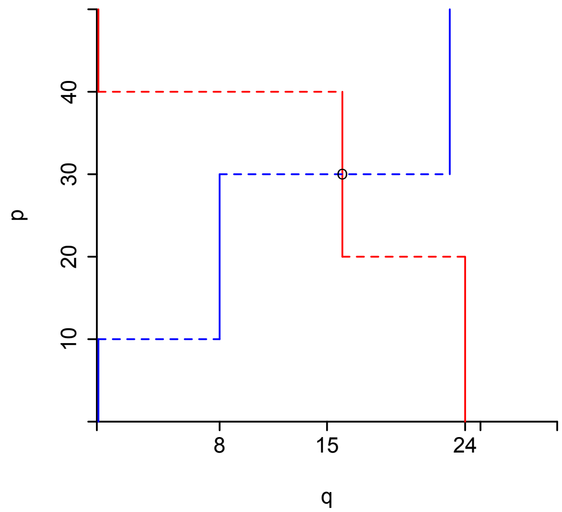


Figure 51: Market Equilibrium in Example 2

deed end up trading at the equilibrium price. It is this magic that we want to see. The true supply and demand curves, that were only known to the experimenter, were those shown in Figures 50 and 51.

Having been assigned their value, the experiment subjects could walk around the room, looking for trading partners. Communication was informal, and not controlled by the experimenter. Whenever a buyer found a seller, or a seller found a buyer, and the two could agree to a price at which to , they went to the experimenter and told him. The experimenter wrote the price at which the two had traded on the board so that everyone could see it, and made the promised payments. The experiment ended once nobody wanted to continue. Did the magic of the invisible hand do its work?

The Invisible Hand At Work

I shall now describe the results of the experiments that I described in the previous paper. I shall rely on the paper "Bazar Economics" by John Miller and Michele Tumminello that was published in the *Journal of Economic Behavior & Organization*, Vol 119, 2015, page 163-181. I shall use their tables to represent the results. The tables shown in Figure 52 are for Example 1, and the tables shown in Figure 53 are for Example 2. Let's first focus on Example 1. Recall that in this example the equilibrium price is 20. Figure 52 appears complicated, but we shall focus only on some of the information provided. There are 8 panels. In each panel we have on the horizontal axis prices on which a pair buyers and sellers might agree, and on the vertical axis is the number of pairs of sellers and buyers who actually traded at these prices. For example, in the top left panel in Figure 52, you see a little bar at price 15, which corresponds to, perhaps, 10 trades.

The top four panels show results from one researcher's experiments, the bottom four panels show results from some other researcher's experiments. For each researcher's experiments, there are four panels, corresponding to the prices at which (1) low cost sellers traded with high value buyers, (2) low cost sellers traded with low value buyers, (3) high cost sellers traded with high value buyers, (4) high cost sellers traded with low value buyers. If one follows each set of four panels clockwise, starting in the top left corner, then one sees first case (1), then (2), then (3) and then (4). For example, in the experiment conducted by the first researcher, low cost sellers traded with low value buyers most frequently at the price 15. This is shown int he top right panel.

Miller and Tumminello analyze not their own, but other researchers' data.

Now recall that the law of supply and demand predicts that in Example 1 low cost sellers, i.e. sellers who have a cost of 10, and high value buyers, i.e. buyers with a value of 40, should always trade when they meet, and they should trade at the equilibrium price of 20. Focusing on the top left panel for the first researcher's experiments, we can see that this prediction is violated and instead the majority of the subjects traded at price 25 rather than at price 20. Many other prices were also chosen, however, and among those other prices the price 20 was the most popular. In the second researcher's experiments, low cost sellers and high cost buyers traded mostly at price 20, although many also traded at price 25, and many other prices were also popular. This you can see in the panel in Figure 48 that is in the third row on the left hand side.

What do we make of these findings? Do they confirm or falsify the law of demand? As is so often the case in economics, the answer is not a clean "yes," or "no." Let us postpone judgment, and continue to look at data.

Considering next the low cost sellers and low value sellers, the equilibrium prediction is that only some of these will trade, and that, if they trade, they too should trade at the price of 20. This price is the value of the low value buyers.

Figure 52: Results for Example 1

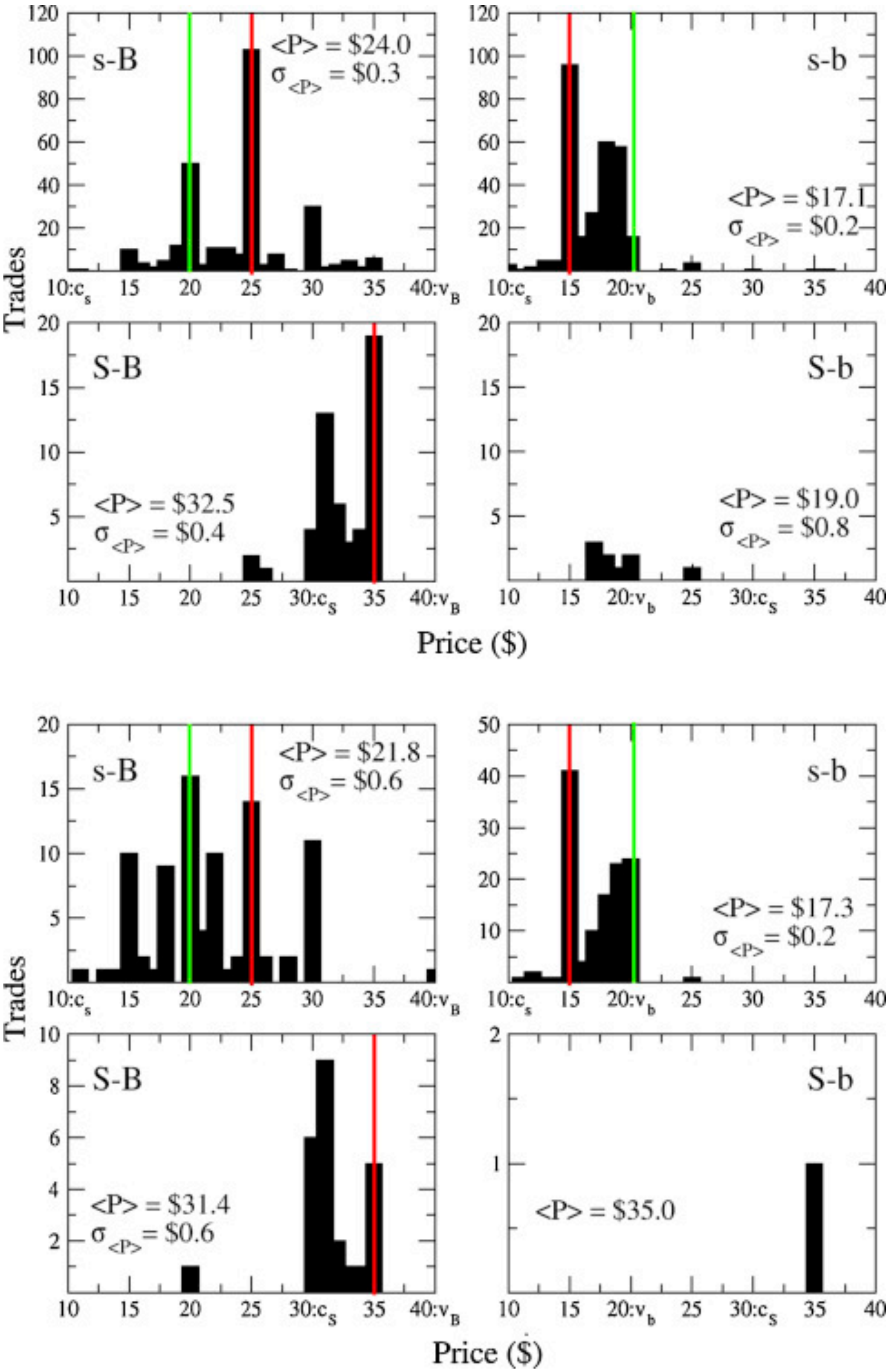


Figure 52 shows that the majority of these pairs of buyers and sellers actually trades at price 15, and not at price 20. Prices between 15 and 20 were quite common.

Another prediction of the law of supply and demand is that none of sellers with high cost will trade. But note in Figure 52 that many of the high cost sellers do find a buyer with a high value who is willing to strike a deal, and indeed such deals are possible. Both sides typically on prices between 30 and 35, quite far away from the equilibrium price. And, somewhat amazingly, even sellers with high cost (30) sometimes trade with buyers with low values (20). Someone must make a loss there.

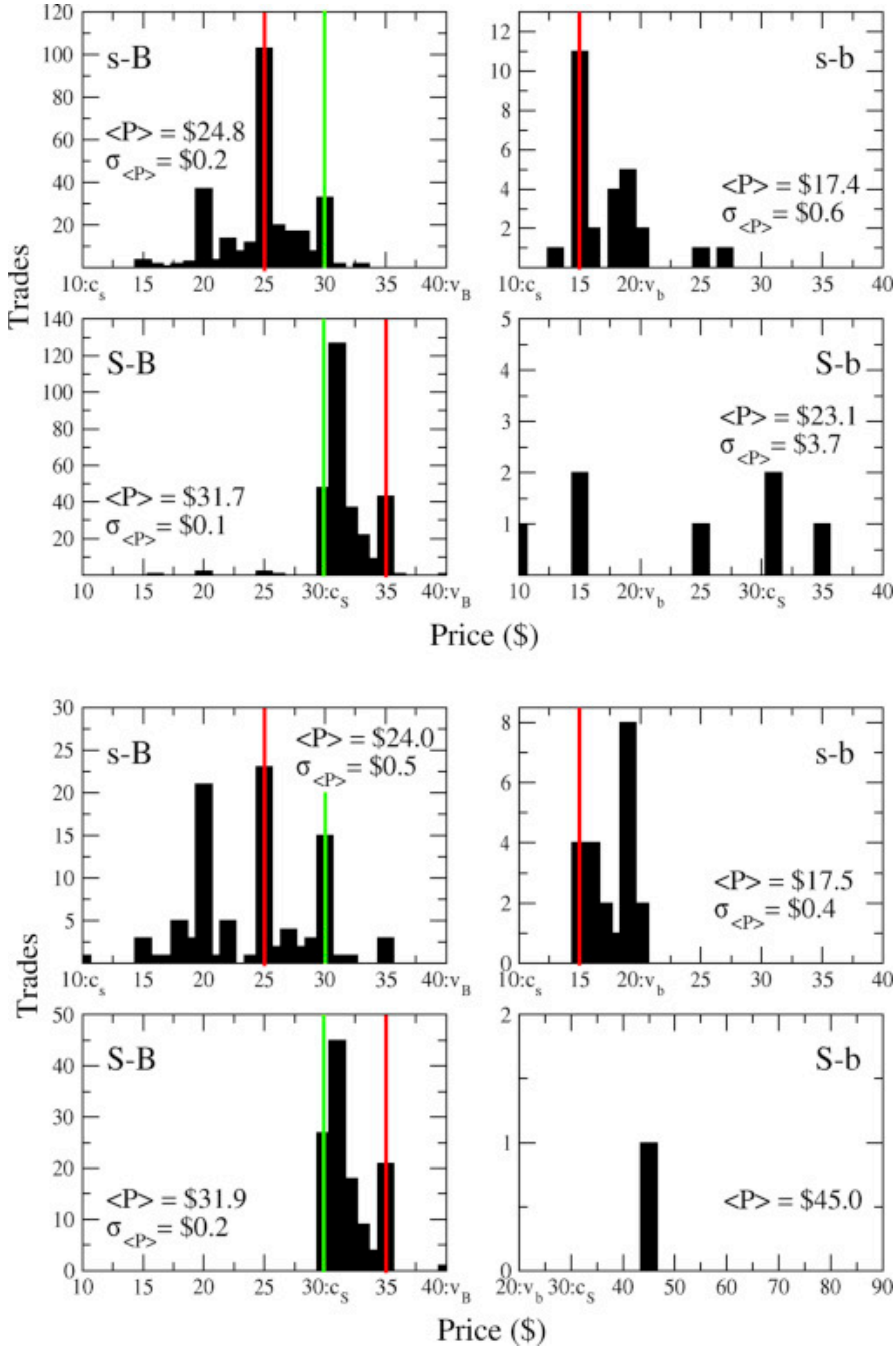
The results for Example 2 are similarly messy. Buyers with low values are not at all supposed to buy, but some of them do find a seller who is willing to sell to them. Buyers with high values are all supposed to trade at price 30, but if they trade with a low cost seller, they often just pay 25, or even 20, rather than 30. Only if they trade with a high cost seller do they pay 30, or sometimes more.

We should not overlook the positive results though. If you pay attention to the units on the vertical axes, in Figure 52, as predicted, most of the trades involved low cost sellers, as supply and demand predicts. The number of trades involving high cost sellers is small. In Figure 53, by contrast, both low and high cost sellers trade with high frequency, but many more high value buyers than low value buyers get to trade, as supply and demand predicts. Finally, we are able to see that in Figure 52 the low cost sellers mostly traded at prices between 15 and 25. In Figure 53, the high value buyers mostly traded at prices between 25 and 30. Thus, as the law of supply and demand predicted, prices in Example 2 were higher than they were in Example 1.

What Can We Conclude?

What is going on? Let us develop some intuition for the equilibria in Examples 1 and 2. In Example 1, shown in Figure 50, there are many sellers with very low cost: just 10 Dollars. Ideally, each of these would like to sell to a high value buyer, and charge this buyer a high price. But lots of sellers compete for few high value buyers. Once buyers realize that not all sellers can sell to high value buyers, but some must sell to low value buyers, they will not accept a

Figure 53: Results for Example 2



deal that is worse than a deal that the low value buyers get. Of course, high value buyers must realize that there are few of them. So they have to observe the market quite carefully!

But if high value buyers are not willing to get treated worse than low value buyers, it is as if all buyers had low values. Therefore, high cost sellers can't sell. And there are overall more buyers than low cost sellers. So, now the market power is on the side of the sellers. They can exploit the buyers' fear to find nobody to buy from, and charge them (almost) all of their willingness to pay, i.e. 20 Dollars.

What happens in the experiment? The high value buyers often end up paying more than they should. Maybe this is because they don't realize that there are only few of them, and that the sellers all compete for them. The low value buyers often pay less than their willingness to pay. Perhaps it is because the sellers don't really want to exploit their market power too much? The most common price, 15, is just in the middle between the buyers' values and the sellers' cost. Perhaps they are splitting the difference equally?

There are some indications in these experiments that things change if the experiment is repeated with the same subjects. High value buyers realize slowly that they don't really have to pay that much. Sellers realize that they can exploit low value buyers without getting rejected. It seems as if over time the law of supply and demand was reinforced. But I am not aware of enough evidence, and so this remains a conjecture rather than becoming a fact that I could teach.

In Experiment 2, the situation is the reverse of that in Example 1. In Example 2, as shown in Figure 51, there are many buyers with high value: 40 Dollars. They would all like to trade with the low cost sellers, and persuade those sellers to accept a low price. But low cost sellers are rare, and they can try to exploit this fact to push the price up, so that they are paid as much as high cost sellers. Once we are at that point, the bargaining power is on the side of the buyers. There are now more eager sellers than buyers. This pushes the equilibrium price down to the high cost value, namely 30. In practice, the low cost sellers get most often a price of 25 rather than 30. Maybe they don't realize how rare they are. The high cost sellers, though, as Figure 53 indicates, are typically pushed down right to their cost value. Buyers in Figure 53 don't show as much mercy as sellers in Figure 52.

This very informal discussion of the experimental results suggests that one might suspect two factor to be important in these markets. One is the transparency of the market: how much do the market participants know about supply and demand side? The other is: how much are participants willing to exploit their bargaining power to the limit, and how much are fairness considerations stopping them from this? High transparency, and ruthlessness, seem to favor the law of supply and demand. But these remain conjectures. If I were not writing these notes right now, I would, perhaps, tomorrow start an experimental research project designed to investigate the conjectures more carefully. But instead ... I shall continue to work on these notes tomorrow.

Topic 15: Equilibrium Consequences of Economic Policy

Let's now put aside our doubts about the law of supply and demand. The insights that we discuss next remain relevant, even if the law of supply and demand does not hold exactly. The subject of this section is a simple but important principle: If economic policy affects the market system, it is important to not only anticipate the immediate consequences of the policy, but the equilibrium consequences, i.e., one has to ask: If I choose such and such a policy, how does the equilibrium of the market change?

An analogy might be helpful: suppose we want to study how pesticides affect some particular eco system. We might find that animals or plants of some particular species are killed by the pesticide. But that observations is just a first step when thinking about the consequences of the pesticide. Other consequences might follow, because some other species perhaps preys on the species that is directly affected, and therefore, indirectly, might also be threatened in its survival. Or, conversely, some species competing for food might benefit from the extinction of their competitors. Eco systems, like markets, have "equilibria," and any interferences with an ecosystem will have not just immediate, but also equilibrium effects. We shall illustrate this principle in some examples.

Tax Incidence

Suppose a market for a good is in equilibrium, as shown in Figure 46 in Topic 14, and suppose the government decides to tax the good in the following way: Whenever a seller sells a unit of the good, the seller has to pay τ Dollars to the government. What will happen? Consumers have to pay τ Dollars more? Firms will lose τ Dollars for each unit sold? Neither of these conjectures, although often heard in policy debates in the real world, is a prediction that follows from the law of supply and demand. The law of supply and demand predicts something different.

The tax will affect in the first place what firms are willing to supply. If the market price, i.e. the price that consumers are paying to the firms, is p, then what firms are getting net is $p - \tau$. Therefore, firms' optimal supply at price

p is the same as what it was previously at price $p - \tau$. That is, to construct new supply at price p, we go to the old supply curve, look up supply at price $p - \tau$, and go τ units up vertically. That is a point on the new supply curve. Therefore, the new supply curve is the old supply curve, shifted up by τ units. I show this in Figure 54.

Figure 54: Equilibrium Effects of a Tax on Supply

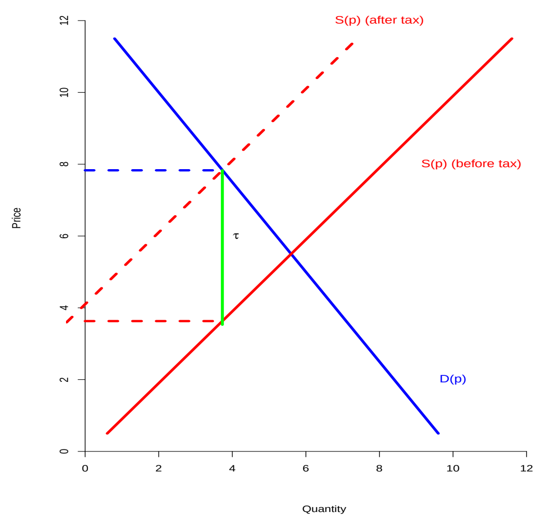


Figure 54 shows that the shift in the supply curve, which is the immediate effect of the policy, leads to a change of the equilibrium. The new equilibrium is given by the intersection of the new supply with the demand curve, which, of course, has not changed. We see that the equilibrium price has risen, but not by the tax τ, but by less. Let us call the equilibrium price before the tax increase by $p^*(0)$, where the "0" indicates that this is the price when tax is zero. The equilibrium price after the tax increase is then: $p^*(\tau)$. What we have just established is that $p^*(\tau) - p^*(0)$ is less than τ. Recall that the price that we are talking about is the price that consumers hand over to firms. So, the law of supply and demand says that this price goes up, but it goes up by less than τ.

A consequence is, of course, that the firms get a price that is lower than what they had before. They get: $p^*(\tau) - \tau$ (because τ goes to the government) and this is less than $p^*(0)$. In other words, we can split the effect of the tax into two parts: $p^*(\tau) - p^*(0)$, which is the loss to the consumers, and $p^*(0) - (p^*(\tau) - \tau))$ which is the loss to firms. If we add up these two terms, we get exactly τ:

$$p^*(\tau) - p^*(0) + p^*(0) - (p^*(\tau) - \tau)) = \tau$$

Thus, in a sense the proportion of the tax that becomes a price increase for consumers, and the proportion of the tax that becomes a price decrease for firms is:

Check for yourself that these two proportions add up to 1.

$$\text{consumers: } \frac{p^*(\tau) - p^*(0)}{\tau} \qquad \text{firms: } \frac{p^*(0) - (p^*(\tau) - \tau))}{\tau}$$

One also says that these terms represent the tax incidence, which is the answer to the question: "Who bears which proportion of the burden of the tax?" Of course, the above formula only looks at price changes. One might also look at quantity changes, or one might look at "utility" changes, but remember that utility has only ordinal meaning, so the latter calculation would be difficult to interpret.

There are two questions that we shall pursue for the remainder of this section. One is: Can we say something about the determinants of tax incidence? The second is: How would things be different if consumers rather than firms have to pay the tax?

To answer the first question we shall do a little calculation. We shall focus on the share of just one market side, because the two shares add up to 1, and once we know one side's share we also know the other side's share. Let's focus on consumers. Thus, the fraction in which we are interested is:

$$\frac{p^*(\tau) - p^*(0)}{\tau}.$$

We shall calculate this fraction for the case that τ is infinitesimally small. This allows us to use calculus. We shall obtain a very clean and simple result. Thus, we shall calculate:

$$\lim_{\tau \to 0} \frac{p^*(\tau) - p^*(0)}{\tau}.$$

But note that this is just the derivative of p^* at zero. To simplify notation, we shall drop the *. Then we can write the number in which we are interested as:

$$p'(0).$$

Now let's start the calculation. We shall first write down the equation that indicates how the function $p(\tau)$ is defined:

$$S(p(\tau)-\tau) = D(p(\tau))$$

This just says that supply at price $p(\tau)-\tau$ must equal demand at price $p(\tau)$. Let's differentiate both sides of this equation. Because the two sides are equal, their derivatives also have to be equal:

$$S'(p(\tau)-\tau))(p'(\tau)-1) = D'(p(\tau)p'(\tau).$$

Here, I have used on both sides the chain rule. Our objective is to solve for $p'(\tau)$. Multiplying out the left hand side, we get:

$$S'(p(\tau)-\tau))p'(\tau) - S'(p(\tau)-\tau) = D'(p(\tau))p'(\tau).$$

We can re-arrange this as:

Sorry for the many brackets. There is nothing I can do about them.

$$\Big(S'(p(\tau)-\tau)) - D'(p(\tau))\Big)p'(\tau) = S'(p(\tau)-\tau).$$

And now we can solve for $p'(\tau)$:

$$p'(\tau) = \frac{S'(p(\tau)-\tau)}{S'(p(\tau)-\tau) - D'(p(\tau))}.$$

Now recall that we want to evaluate this derivative when $\tau = 0$. We conclude:

$$p'(0) = \frac{S'(p(0))}{S'(p(0)) - D'(p(0))}.$$

Note that the last term in the denominator on the right hand side, $D'(p(0))$, is negative. So, the denominator really is the sum of the derivative of supply and the absolute value of the derivative of demand. The numerator is the derivative of supply. Thus, the ratio must be a number between 0 and 1, which it should be, because we are calculating the fraction of the tax burden that goes to the consumers.

Economists like to express everything in elasticities. Recall that we obtain the elasticity of, say, supply by multiplying its derivative by $\frac{p}{S(p)}$. If we don't want to change the value of the fraction that we have obtained for $p'(0)$, we shall multiply both its numerator and its denominator by $\frac{p(0)}{S(p(0))}$. This gives us:

$$p'(0) = \frac{S'(p(0))\frac{p(0)}{S(p(0))}}{S'(p(0))\frac{p(0)}{S(p(0))} - D'(p(0))\frac{p(0)}{S(p(0))}}.$$

Now we have the elasticity of supply both in the numerator and the denominator. But what is the second term in the denominator? It is actually the elasticity of demand. This becomes clear if we remember that:

$$S(p(0)) = D(p(0)),$$

because $p(0)$ is an equilibrium price when tax is zero. Therefore, in the denominator, we can replace $S(p(0))$ by $D(p(0))$. Our final equation for $p'(0)$ is:

$$p'(0) = \frac{S'(p(0))\frac{p(0)}{S(p(0))}}{S'(p(0))\frac{p(0)}{S(p(0))} - D'(p(0))\frac{p(0)}{D(p(0))}}.$$

We can write this in simpler notation. Let's write ε_S for the price elasticity of supply, and ε_D for the price elasticity of demand. Then:

Aren't you relieved that our equation finally looks simple?

$$p'(0) = \frac{\varepsilon_S}{\varepsilon_S - \varepsilon_D}.$$

Thus, for infinitesimally small price changes, the share of the tax burden that consumers bear is the elasticity of supply, divided by the elasticity of supply plus the absolute value of the elasticity of demand. The larger the elasticity of demand the smaller this share is. The larger the elasticity of supply is, the larger the share the consumers bear is.

What is the firms' share? It must be 1 minus the consumer's share:

$$1 - p'(0) = 1 - \frac{\varepsilon_S}{\varepsilon_S - \varepsilon_D} = \frac{-\varepsilon_D}{\varepsilon_S - \varepsilon_D}.$$

Whose share is smaller? It is the consumer's share if and only if:

$$\begin{aligned} \frac{\varepsilon_S}{\varepsilon_S - \varepsilon_D} &< \frac{-\varepsilon_D}{\varepsilon_S - \varepsilon_D} \Leftrightarrow \\ -\varepsilon_D &> \varepsilon_S. \end{aligned}$$

This is the bottom line. Consumers are better off than firms if demand (in absolute terms) is more elastic than supply. Otherwise, firms bear the larger part of the tax burden. You can check this by drawing diagrams of supply and demand, and shifting the supply curve upwards. If demand is elastic, for example, meaning that it changes a lot as the price changes, which means that in the supply and demand diagram the demand curve is a pretty flat line, then consumers bear a small portion of the tax burden. I show this in Figure 55.

Remember to bend your neck when you look at the supply and demand diagram, so that the vertical price axis becomes the "x-axis."

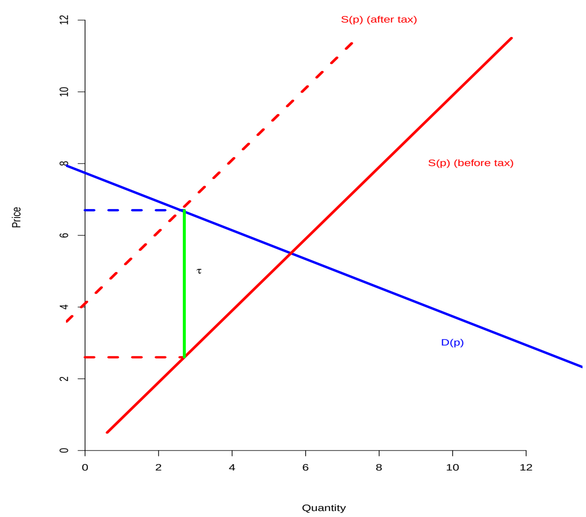
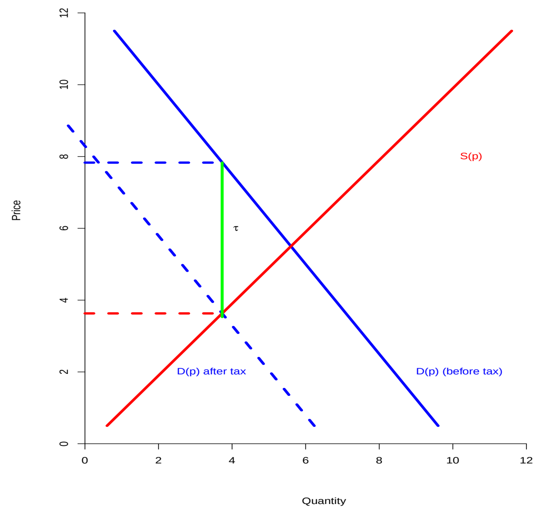


Figure 55: Equilibrium Effects if Demand is Very Elastic

I mentioned (as it might seem: a long time ago) a second question we wanted to study in this section. It is: which difference does it make whether firms, or consumers pay the tax? The surprising answer to this question is that the law of supply and demand predicts that it makes no difference whatsoever. Consider Figure 56, which shows what happens if consumers pay the tax. What is on the vertical axis is, as in Figures 9 and 10, the price in Dollars interpreted as the payment that consumers make to firms. The demand function before the tax is the blue unbroken line, and the demand function after the tax is the blue dashed line. The demand function after tax is constructed as follows. At any price p the consumers anticipate that they don't just pay the firms p Dollars, but that in addition they have to pay the tax τ to the government. Therefore, to construct demand after tax at a price p, we check the quantity demanded before the tax at price $p + \tau$. That will be the demand at price p

after the tax is introduced. This means that the demand function shifts parallel downwards by an amount τ when the tax is introduced.

Figure 56: Equilibrium Effects of a Tax on Demand



Now compare Figure 56 to Figure 54. In both figures the equilibrium quantity is reduced; geometrically, it has moved to the left. It has moved to exactly the point where the vertical distance between the old demand and the old supply function is τ. Thus, the quantity produced and consumed in equilibrium is the same in both diagrams. Moreover, the price that firms get net is the price on the old supply function that corresponds to the equilibrium quantity, and the price that consumers have to pay in total is the price on their old demand function that corresponds to the equilibrium quantity. Thus consumers pay, and firms get, the same amounts of money in Figures 9 and 11. The only difference between the two figures is *who* transfers the tax τ to the government: in Figure 54 it is the firms, and in Figure 56 it is the consumers. But that is just a bureaucratic detail. The end result for consumers and firms is the same in both graphs. Thus, who pays the tax ultimately makes no difference.

Subsidizing Perfect Complements

So far, we have focused on a single market. But often several markets are connected with each other. We shall now investigate the effects of a government intervention in a market on the equilibrium not only in that market, but also in a related market. In this section, the related market is the market for a perfect complement of the good traded in the market in which the government interferes. In the next section, we shall then turn to the case of perfect substitutes.

Recall that if two goods are perfect complements, a simple representation of their preferences is given by the utility function:

$$U(q_1, q_2) = \min\{q_1, q_2\}.$$

Let's assume that all consumers have this utility function. Then each consumer's demand is given by:

$$D_1(p_1, p_2) = \frac{Y}{p_1 + p_2} \text{ and } D_2(p_1, p_2) = \frac{Y}{p_1 + p_2}.$$

Actually, we can also think of these expressions as market demand. We just have to interpret Y as the sum of the incomes of all consumers in the market.

Both of the goods have a supply function. Let's denote the supply of good 1 by S_1, its price by p_1, the supply of good 2 by S_2, and the price of good 2 by p_2. To make things very simple we postulate these supply functions:

$$S_1(p_1) = p_1 \text{ and } S_2(p_2) = p_2.$$

Before we introduce some form of government policy, let's calculate equilibrium in these two markets without government policy. Both markets are in equilibrium if supply equals demand in both markets simultaneously. The equilibrium prices therefore must satisfy:

$$D_1(p_1, p_2) = S_1(p_1) \quad \text{and} \quad D_2(p_1, p_2) = S_2(p_2) \Leftrightarrow$$

$$\frac{Y}{p_1 + p_2} = p_1 \quad \text{and} \quad \frac{Y}{p_1 + p_2} = p_2$$

Let's solve these two equations for p_1 and p_2. Note that the terms on the left hand side in both equations are the same. Therefore, also the terms on the right hand side need to be equal to each other:

$$p_1 = p_2,$$

That is, the two goods have to have the same price. Let's denote the price by p. Then both equations are equivalent to:

$$\frac{Y}{2p} = p.$$

We can solve for p as follows:

$$\frac{Y}{2p} = p \Leftrightarrow 2p^2 = Y \Leftrightarrow p^2 = \frac{Y}{2} \Leftrightarrow p = \sqrt{\frac{Y}{2}}.$$

Thus, in equilibrium the price of both goods is $\sqrt{\frac{Y}{2}}$. The quantities supplied are equal to the price. Therefore, also the quantities demanded will be equal to the price. That is because we have solved for an equilibrium of supply and demand.

Now let's introduce some government policy. To mix things up, we shall now consider a subsidy, not a tax. The analysis of a tax is similar. Specifically, we shall assume that whenever a firm in the market for good 1 sells a unit at price p_1, then the government steps in and pays the firm, from government revenue, another p_1 Dollars. The effect of this on the supply of good 1 is this. When the market price, i.e. the amount that consumers hand over to the firms, is p_1, then the firms understand that they will get another p_1 Dollars from the government, and therefore, their supply is now what it was before at price $2p_1$. Thus, the new supply is:

After you have read this section, check whether you can do the analysis of a tax yourself.

$$S_1(p_1) = 2p_1.$$

There is no change in the market for good 2. We don't assume that the government also subsidizes good 2.

The equations that say that demand equals supply in both markets are now:

$$\frac{Y}{p_1 + p_2} = 2p_1 \text{ and } \frac{Y}{p_1 + p_2} = p_2.$$

Again, we note that the left hand sides are the same, which implies that the right hand sides have to equal each other:

$$p_2 = 2p_1.$$

Substituting back into the equilibrium condition for good 1, we can solve for the price of good 1:

$$\frac{Y}{p_1 + 2p_1} = 2p_1 \Leftrightarrow Y = 6p_1^2 \Leftrightarrow p_1 = \sqrt{\frac{Y}{6}}.$$

The price of good 2 is:

$$p_2 = 2\sqrt{\frac{Y}{6}} = \sqrt{\frac{4Y}{6}} = \sqrt{\frac{2Y}{3}}.$$

How have prices changed? Before the subsidy both prices equaled $\sqrt{\frac{Y}{2}}$. After the subsidy, the price of good 1 has fallen, but the price of good 2 has gone up! Recall that we are talking about the prices paid by consumers. Firms in the market for good 1 receive a lower price, but they also get the subsidy from the government. But let's focus here on consumers. Because firms are subsidized, consumers in the market for good 1 pay a lower price. This is not surprising. The effect on the price of good 2 is more surprising. The price of good 2 goes up because good 1 has become cheaper, and therefore there is more demand in the market for the complementary good 2. Consumers want to buy the two goods together. But there has been no change in the supply of good 2. Therefore, good 2's price has to go up. This is an indirect consequence of the subsidy. One might overlook it if one doesn't think in terms of equilibrium of supply and demand in both markets.

We might also ask how the utility of consumers has changed. The utility of consumers is equal to the number of units of both goods that they can buy. Because income is fixed, what matters for utility is the sum of the two prices. Before the subsidy this sum was:

$$\sqrt{\frac{Y}{2}} + \sqrt{\frac{Y}{2}} = 2\sqrt{\frac{1}{2}}\sqrt{Y} = \sqrt{2}\sqrt{Y}.$$

After the price increase it is:

$$\sqrt{\frac{Y}{6}} + \sqrt{\frac{2Y}{3}} = \left(\sqrt{\frac{1}{6}} + \sqrt{\frac{2}{3}}\right)\sqrt{Y} = \sqrt{\frac{3}{2}}\sqrt{Y}.$$

I am leaving out a lot of steps in the calculation that proves the last equality. Can you fill them in?

We see that the sum of the prices has fallen. Thus, consumers are better off after the subsidy, even though only the price of good 1 decreases, but the price of good 2 has increased.

Subsidizing Perfect Substitutes

Recall that if two goods are perfect complements, a simple representation of their preferences is given by the utility function:

$$U(q_1, q_2) = q_1 + q_2.$$

Let's assume that all consumers have this utility function. If Y is the sum of consumers' income, then market demand is simple: If good 1 is cheaper, all consumers buy good 1. If good 2 is cheaper, all consumers buy good 2. If the two goods have the same price, consumers are indifferent. Let's assume that the two goods, although perfect substitutes, have two separate supply functions. Let's denote the supply of good 1 by S_1, its price by p_1, the supply of good 2 by S_2, and the price of good 2 by p_2. To make things very simple we postulate these supply functions:

$$S_1(p_1) = p_1 \text{ and } S_2(p_2) = p_2.$$

Like in the previous section we shall start by asking what is an equilibrium if the government does not interfere at all with the market. I'll derive it without equations, just through logical reasoning, because the case of perfect substitutes is a little bit exceptional. First, I'll argue that the prices in the two markets have to be the same. This seems very intuitive. Let's be pedantic, though, about the argument: Suppose one price, say p_1 were higher than the other, i.e. p_2. Then demand for good 1 would be zero. But supply for good 1 would be some positive number. Therefore, there would be more supply than demand in the market for good 1. Therefore, this cannot be an equilibrium. We conclude that the price has to be the same for both goods. Let us denote the price by p.

In each market, the supply is then equal to p. The sum of the supplies in both markets is then $2p$. Consumers are indifferent between the two goods. Their total demand is: Y/p. For equilibrium, the sum of supplies must equal the total demand, that is:

$$2p = \frac{Y}{p} \Leftrightarrow p^2 = \frac{Y}{2} \Leftrightarrow p = \sqrt{\frac{Y}{2}}.$$

The same formula showed up in the previous section as an equilibrium price. This is pure coincidence.

This is the equilibrium price in both markets. It is also the quantity supplied in both markets, because of the simple form of the supply functions. Consumers are indifferent, but the sum of their demands is equal to the sum of supplies in both markets. Thus, if half of the consumers buy one good, and half of them buy the other, we have an equilibrium.

Now suppose the government decided to subsidize good 1 by paying p_1 to firms for every unit sold. As in the previous section, we would then get that the supply of good 1 is:

$$S(p_1) = 2p_1.$$

Prices in the two markets would still have to be the same in equilibrium. Denoting the equilibrium price by p, the sum of supplies is now, however: $2p + p = 3p$. The equation that says that sum of supply in both markets equals sum of demand in both markets is now:

$$3p = \frac{Y}{p} \Leftrightarrow p^2 = \frac{Y}{3} \Leftrightarrow p = \sqrt{\frac{Y}{3}}.$$

In equilibrium, firms in market 1 supply a quantity twice as large as the quantity supplied in market 2. To reach equilibrium, therefore, twice as many consumers have to buy good 2 as buy good 1.

Comparing prices before and after the subsidy, we see that prices in both markets have fallen. The point of this example is its contrast with the previous example: when goods are perfect complements, then, when you subsidize one, the price of the other goes up. When goods are perfect substitutes, then, when the government subsidizes one, the price of the other goes down. These are examples of consequences of government policies that become transparent only after an equilibrium analysis.

Hopefully transparent?

Topic 16: General Equilibrium

After we have studied equilibrium of supply and demand in one market, and then in two interrelated markets, we are now going to study equilibrium of supply and demand in all markets at the same time. Equilibrium in all markets at the same time in an economy is also called a *General Equilibrium*.

The task before us looks a little frightening. If you think about the number of markets in, say, the US economy, surely it must be in the millions, one for each good that we can think of. If we wrote down for each of these markets that supply must equal demand, then we would have millions of equations. This would be impossible for us to handle, at least with the methods of mathematics that we can use in this course.

Solving ever more equations is, however, also not the reason why we study general equilibrium here. Our purpose is very different. Let us recall why we study market equilibrium, or, even more generally, why we study economics: We want to see how markets determine which use the economy will make of its resources. The resources of the economy are the ability of some people to work, maybe some materials that we find in the ground, and also, less tangible, society's knowledge of how to produce outputs from inputs. Our labor time and the material resources can be put to many uses. We could spend our time mostly in leisure, we could make a lot of wooden furniture, or, perhaps, we could decide to produce lots of cell phones. Which of these things happens is determined by markets.

Ultimately, we want to study whether markets do a good job at deciding about how to use resources. But let's postpone this normative question until later. At the moment, we are just asking how markets make these decisions. The purpose of this section is to demonstrate this in simple, numerical examples. But to show such an example, I have to give a complete description of all the resources that the economy has, which outputs could be produced, and which preferences people have. Then I have to determine equality of supply and demand in all markets. Thus, I need to study general equilibrium to be able to give a complete picture of how a market economy directs the economy's resources.

Fortunately, I can do this in models with a very small number of markets. For example, it is enough to have two goods, leisure and consumption, and to just study how the economy decides how much time goes into leisure, and how much time goes into the production of output. A consequence of having only two goods, leisure and consumption, will be that the firms in the following models will only use one factor of production, namely labor. This makes the firms' profit maximization problem much simpler than it was in the models of profit maximization that we have seen earlier, but the treatment of the firms in these examples may look unfamiliar. I hope that it is unfamiliar, but straightforward for you.

We shall study two very simple examples of models in which there is "equilibrium in all markets," and there are actually only two goods. The purpose of these examples is just to make it very clear how the model of supply and demand determines resource allocation.

Inevitably, as we only look at specific examples with particular functional forms, in this section we shall do a couple of calculations. Don't be deterred by these calculations. I have tried to keep them really simple. It just *looks* like a lot of equations. It is not really hard.

Example 1

We shall make this example very simple. There is only one consumption good, but there is also leisure. There is only one firm. The firm can produce the consumption good using labor. The firm's production function is:

$$f(L) = \sqrt{L}.$$

Let us suppose that there are $2n$ consumers. Each of them has an endowment with time that is equal to 24 hours. The consumers have preferences over leisure and consumption. The first n of the consumers, that is, consumers 1,2,...,n, have utility function:

$$U(\ell, q) = \ell q^2,$$

where ℓ is the amount of leisure, and q is the amount of the consumption good. The remaining n consumers, $n+1, \ldots, 2n$, have the utility function

$$U(\ell, q) = \ell^2 q.$$

This economy will have only two markets: the labor market, and the market for the consumption good. Thus, this is a very simple model. But we want to investigate how much of their time the consumers give up to work, and how much output is going to be produced.

We shall determine demand and supply for labor, and for the consumption good. Let us begin with the firm's demand and supply. We can find the firm's cost function easily. To produce q units of output, the firm has to use the number L of hours of labor that solves the equation:

$$q = \sqrt{L}.$$

The solution is:

$$L = q^2.$$

If the price of labor is w, therefore, the cost function is:

$$C(q) = wL = wq^2.$$

The marginal cost are:

$$C'(q) = 2wq.$$

This is increasing in q, and moreover, at $q = 0$, the marginal cost are zero. Therefore, at all prices, the quantity supplied is determined by the equation price equals marginal cost:

$$p = 2wq.$$

If we solve this for q, we get the supply function:

$$S(p, w) = \frac{p}{2w}.$$

We also get a labor demand function, because, as we found out earlier, $L = q^2$. Therefore:

$$L(p, w) = \left(\frac{p}{2w}\right)^2.$$

We now turn to the consumers' side, and determine the consumers' demand for the consumption good and their labor supply. The consumers have Cobb Douglas utility functions. We have seen earlier in the course how the demand functions for consumers with Cobb Douglas utility functions are determined. We shall now use that, assuming you still remember. If you don't, you need to go back in the notes.

Because for consumers $1, 2, \ldots, n$ the coefficient in the utility function for leisure ℓ is 1, and the coefficient for consumption q is 2, these consumers will spend 1/3 of their income on good 1, and 2/3 of their income on good 2. Moreover, recall from the section on labor supply that it is useful to think of the consumer's income as the amount of money that she gets if she sells all her leisure initially, and then later "buys back" leisure. Thus, all consumers have income $24w$. Putting all of this together, we find that the consumers $1, 2, \ldots, n$ demand leisure equal to:

$$\frac{1}{3}\frac{24w}{w} = 8.$$

They will thus supply 24-8=16 hours of labor. These consumers' demand for consumption is:

$$\frac{2}{3}\frac{24w}{p} = \frac{16w}{p}.$$

For consumers $n+1, n+2, \ldots, 2n$, the calculation is essentially the same, except that they have different exponents in their utility functions. We can calculate their demand for leisure as:

$$\frac{2}{3}\frac{24w}{w} = 16.$$

The only offer 24-16=8 hours of work. Their consumption demand is:

$$\frac{1}{3}\frac{24w}{p} = \frac{8w}{p}.$$

We have to introduce one more idea to complete the model. We have to introduce a "capitalist" who owns the firm and therefore gets all its profits. The firm's profits are revenue minus labor cost, that is:

$$\begin{aligned} & p \cdot \frac{p}{2w} - w\left(\frac{p}{2w}\right)^2 \\ = \quad & \frac{p^2}{2w} - \frac{p^2}{4w} \\ = \quad & \frac{p^2}{4w} \end{aligned}$$

Let us suppose that the capitalist is unable to work. Therefore, the profit income is all his income, and he spends it all on consumption. Therefore, his consumption demand is income divided by p:

$$D(p,w) = \frac{\frac{p^2}{4w}}{p} = \frac{p}{4w}.$$

Now let us write down the equation "supply=demand" in the market for the consumption good:

$$n\frac{16w}{p} + n\frac{8w}{p} + \frac{p}{4w} = \frac{p}{2w}.$$

Here, the left hand side is demand, and the right hand side is supply. The left hand side is the sum of three terms: the demand from the first n consumers, the demand from the second n consumers, and the demand from the capitalist.

We also need that supply equals demand in the market for labor:

$$n16 + n8 = \left(\frac{p}{2w}\right)^2.$$

The left hand side is labor supply from the first n consumers, who supply 16 hours each, and labor supply from the second n consumers, who supply 8 hours each.

Now let us solve the two "supply=demand" equations. We want to solve them for prices. We begin with the equation for the consumption good. Subtracting $\frac{p}{4w}$ on both sides, we get:

$$n\frac{16w}{p} + n\frac{8w}{p} = \frac{p}{4w}.$$

Re-writing this a little bit, we get:

$$16n\frac{w}{p} + 8n\frac{w}{p} = \frac{1}{4} \cdot \frac{p}{w}.$$

And now we can add the factors in front of $\frac{w}{p}$ on the left hand side to get:

$$24n\frac{w}{p} = \frac{1}{4}\frac{p}{w}$$

Now let us multiply both sides by $\frac{w}{p}$, and divide both sides by $24n$. We get:

$$\left(\frac{w}{p}\right)^2 = \frac{1}{96n}.$$

Taking square roots on both sides, we obtain:

$$\frac{w}{p} = \sqrt{\frac{1}{96n}},$$

which we can simplify a little bit to get:

$$\frac{w}{p} = \frac{1}{4}\sqrt{\frac{1}{6n}}.$$

This is how far we can go. We have obtained an equation for the ratio between wage and price, which we have also called the "real wage." The right hand side of our equation includes n. But that is just the number of consumers of each type. It is a variable that we want to keep in our model, so that we can later see what happens as n increases.

Let us now look at the second equation, supply equals demand for labor. If we add the two terms on the left hand side we get:

$$24n = \left(\frac{p}{2w}\right)^2.$$

We can take the factor $\frac{1}{4}$ out on the right hand side to get:

$$24n = \frac{1}{4}\left(\frac{p}{w}\right)^2.$$

Then, if we multiply both sides by $\left(\frac{w}{p}\right)^2$, and also divide by $24n$, we get:

$$\left(\frac{w}{p}\right)^2 = \frac{1}{96n}.$$

This is the same equation that we had earlier, and the solution is:

$$\frac{w}{p} = \frac{1}{4}\sqrt{\frac{1}{6n}}.$$

Thus, by solving supply=demand for labor, we have obtained exactly the same equation as when we solved supply=demand for the consumption good. Moreover, we have not solved separately for w and p. We have only got an equation for the ratio of w and p. There are no equations left. This is all we get.

The calculation actually illustrates two insights that are not just true in this example, but very generally. Firstly, in a general equilibrium model the equations supply=demand can only be solved for price ratios, not for a unique value for each price. This is for the following simple reason: If we multiply all prices by a positive constant, then no consumer changes their behavior, because both the prices they have to pay, and their labor income, have been multiplied by the same constant. Thus, their real budget constraint has not changed. Also, for firms, if we multiply all prices by a positive constant, then the firms' profit maximizing choices don't change. it makes no difference whether a firm maximizes profits, or a constant times profits.

This was the first insight that we got from the above calculation. The second insight is: We could have stopped after solving the equation "demand=supply" for the consumption good. We did not have to solve "demand=supply" for labor. We did not get any additional information from solving the second equation. The general version of this insight is: if we have a certain number of equations "supply=demand," then, in a general equilibrium model, if we solve all but one of these equations, we can forget about the last one. It will automatically be satisfied by the solution to the other equations. Unfortunately, it would go too far to attempt to explain here why this is true.

Both insights are only true in *general* equilibrium models, not in *partial* equilibrium models. This is very important to keep in mind. In a partial equilibrium model, we keep some prices in other markets fixed. This makes both of the above insights invalid.

Now we return to our example. How does the economy, in a general equilibrium of all markets, use its resources? Well, the resources were 24 hours of time for each worker. Some of them offer 16 hours of labor, the others only 8, depending on how strong their preference for leisure is. The ones who work a lot consume in the equilibrium:

$$D(w,p) = \frac{16w}{p}.$$

We know the equilibrium value of $\frac{w}{p}$. So we can plug it in:

$$D(w,p) = 16 \cdot \frac{1}{4} \cdot \sqrt{\frac{1}{6n}} = 4\sqrt{\frac{1}{6n}}.$$

The consumers who work less consume in equilibrium:

$$D(w,p) = 2\sqrt{\frac{1}{6n}}.$$

That is half as much as the other consumers get. Finally, the capitalist does not supply any labor, but he consumes:

$$D(w,p) = \frac{1}{4}4\sqrt{6n} = \sqrt{6n}.$$

The firm demands all units of labor on offer, and produces as much output as there is demand.

Thus, the market equilibrium determines a division of time into labor and leisure, and also certain quantities of consumption. Could people be better of if they used their time in a different way? Does the market a good job at determining the use of resources? This normative question will be discussed later.

There is one curious point about this example that I want to mention in conclusion. What happens if the number n of consumers grows very large? the above analysis shows: nobody changes their labor supply, and the amount of leisure that they enjoy. But all consumption quantities decrease as n increases. Thus, all consumers' utility decreases. Who benefits? the capitalist. His consumption, equal to $\sqrt{6n}$, grows without bounds as n increases.

Don't read any political statement into this example. It is interesting to think about the intuition, i.e. *why* we get these changes when n increases. I shall leave that to you as a task that you might take up on your own. In our next example, capitalists will fare much worse than in the first example.

Example 2

In this section we study an example with two different consumption goods, and two different firms that produce this good. Firm 1 produces good 1 from labor. It has production function:

$$f_1(L) = 2L,$$

where the index "1" denotes firm 1 (and good 1). Firm 2 produces good 2 from labor. Firm 2 has production function:

$$f_2(L) = L.$$

Note that the production of good 1 is more efficient than the production of good 2: with one unit of labor, one can make 2 units of good 1, but only 1 unit of good 2.

There are n consumers. They all have, for simplicity, the same utility function:

$$U(\ell, q_1, q_2) = (q_1)(q_2)^3,$$

where ℓ is leisure, q_1 is the consumed quantity of good 1, and q_2 is the consumed quantity of good 2. This shows that consumers don't really care about their leisure at all. We assume, though, that they have an initial endowment with leisure of 12 hours. Obviously, in equilibrium, they will sell all of that as labor. We'll come to that below.

Observe also in the utility function that the exponent for good 2 is much higher than the exponent for good 1. In a sense, this shows that consumers value good 2 a lot. But recall that the production process for good 2 is inefficient. What we are going to work out is how much labor, in general equilibrium, goes into the production of good 1, and how much goes into the production of good 2.

There is also one capitalist. She owns both firms. Her utility function is:

$$U(\ell, q_1, q_2) = (q_1)^2(q_2).$$

She has no endowment with leisure.

Now let's find a general equilibrium for this example. Let's again start with firms. What is firm 1's cost function? To produce q_1 units of good 1, firm 1 needs L units of labor, where L is determined by:

$$q_1 = 2L,$$

that is:

$$L = \frac{q_1}{2},$$

and therefore the cost function of firm 1 is:

$$C_1(q_1) = w\frac{q_1}{2}$$

where w is the wage. Therefore the profits of firm 1 are:

$$p_1 q_1 - w\frac{q_1}{2},$$

where p_1 is the price of good 1. There is something special about this profit function: it is a linear function of q_1. You can see this by writing profit like this:

$$\left(p_1 - \frac{w}{2}\right) q_1.$$

The large brackets indicate the profit per unit of good 1. The firm's profit is just that number, times q_1. Marginal cost are constant. Therefore, as q_1 increases, the profit per unit does not change.

So, how would firm 1 maximize profits? If $p_1 < \frac{w}{2}$, the term in big brackets is negative, and the firm should produce nothing. Supply is zero for such prices. On the other hand, if $p_1 > \frac{w}{2}$, then there is no limit to how much the firm wants to produce. Every extra unit makes more profit. So, we can say that supply is ∞. if price happens to satisfy: $p_1 = \frac{w}{2}$, then the firm makes zero profits for each unit, and is indifferent about how many units it produces. Every quantity is optimal, in some, disappointing, sense.

At this point, we can already conclude something about equilibrium prices. We don't actually have to calculate demand. The conclusion is that in equilibrium we must have that:

$$p_1 = \frac{w}{2}.$$

This is because it cannot be in equilibrium that the firm wants to produce ∞. Nobody is able to buy so much. It can also not be that the firm wants to produce zero. This is because the preferences are such that there is positive demand for good 1, whatever the prices are. We shall verify that later, but if you are familiar by now with demand for Cobb Douglas preferences, you will know that it is true.

We can go through the same steps for firm 2, and the situation is very similar. The cost function is: $C_2(q_2) = q_2$, the profit function is: $(p_2 - w)q_2$, and in equilibrium we must have:

$$p_2 = w.$$

Because in equilibrium the firms must make zero profits, we can ignore the capitalist. She has no income. Her life will be that of an ascetic: she does not work, and does not consume.

Let's turn to the n consumers. Because the consumers don't value leisure they will offer all their leisure time as labor, and therefore have income:

$$12w.$$

They have Cobb-Douglas preferences with coefficients 1 and 3, and therefore, they will spend 1/4 of their income on good 1, and 3/4 of their income on good 2. This implies that demand for good 1 is:

$$\frac{1}{4}\frac{12w}{p_1}.$$

But earlier we found out that $p_1 = \frac{w}{2}$. We can plug that in, and get demand for good 1:

$$\frac{1}{4}\frac{12w}{\frac{w}{2}},$$

and that expression is actually equal to 6. The total demand for good 1 is therefore:

$$6n.$$

Note that the price no longer appears in this. When prices are such that the firms make zero profits, the demand for good 1 is independent of any further details of prices.

The demand for good 2 can be found in the same way. It is:

$$\frac{3}{4}\frac{12w}{p_2}.$$

If we plug in what we found earlier, namely $p_2 = w$, then this is just equal to 9. Each consumer demands 9 units of good 2, and therefore, the total demand for good 2 is:

$$9n.$$

Again, the prices don't appear in this expression.

When prices are such that firms make zero profits, and are indifferent between all quantities, then they are, of course, also willing to supply the quantities that we just calculated, 6n units of good 1 and 9n units of good 2. Thus, we have that supply equals demand in the markets for both consumption goods.

How about the labor market? To produce 6n units of good 1, firm 1 needs $3n$ units of labor. To produce $9n$ units of good 2, firm 2 needs $9n$ units of labor. So, total demand for labor is $12n$. What is labor supply? Well, each consumer has 12 hours available every day, and sells all of these as labor. Therefore, supply in the labor market equals $12n$, and we find that supply equals demand in the labor market as well.

This is already the end of our analysis of example 2. In summary, the prices have to satisfy: $p_1 = \frac{w}{2}$, and $p_2 = w$, and then all markets are in equilibrium. Notice that we have again verified the two insights that we also mentioned in the previous example. Firstly, only the price ratios are determined in equilibrium. We can re-write what we found as:

$$\frac{w}{p_1} = 2 \text{ and } \frac{w}{p_2} = 1,$$

and this makes perhaps even clearer that we have only determined the price ratios. The second observation is that, after we had studied equilibrium in the markets for the two consumption goods, we actually did not gain any additional information from looking at the labor market. We just verified that what we had already found out about prices by looking at the consumption good markets automatically implied that in the labor market demand was also equal to supply. In other words: we only have to study equilibrium of supply and demand in all but one market, not in all markets.

How are resources used in equilibrium? The economy will use 3 times more units of labor in the production of good 2 than in the production of good 1. There are two reasons for this: first, the Cobb Douglas utility function of consumers has a higher exponent for good 2. Another reason is that the production of good 2 is less efficient than the production of good 1. For one unit of good 1 one needs 0.5 units of labor, but for one unit of good 2 one needs 1 unit of labor. one might expect that the economy puts less resources into the less efficient production process. But that is not how it ends up in equilibrium. Consumers value good 2, and therefore the economy puts a lot of resources into the production of good 2, to make sure that despite of the inefficiency enough of good 2 is produced.

Is it optimal to put so many resources into the production of good 2? We shall discuss this question later. It is a *normative* question. So far, we have stayed away from all normative questions.

Topic 17: General Equilibrium and Efficiency

Everything that we have done so far has lead up to the construction of the general equilibrium model that shows how the use of the economy's resources is guided by the "invisible hand" of the price system, by consumers' preference maximizing choices, and by firms' profit maximizing choices. We now turn to the all-important question: Is the resource allocation which the invisible hand directs us to "good"?

I have put "good" into quotation marks because, of course, a central concern here is that we need to define what we mean by "good." We are entering the area of value judgments here, and, of course, it is not easy to make value judgments. On the other hand, it is inevitable at this point. It is an often expressed view that letting the market decide about resource allocation is preferable to other systems, such as letting the government decide about resource allocation. We want to find out whether there is a scientific justification for the preference for markets over other systems. This makes some value judgment inevitable.

How to make value judgments, or even more basic, what do value judgments actually *mean*, is a question that is mostly addressed by philosophers. But there are parts of economics, often called "welfare economics," that study value judgment in the domain of economics. The research undertaken by economists in this area is closely related to philosophers' work on this topic.

The issue of normative judgments in economics will occupy us for a while, not just in this topic. In this topic we shall start by using the perhaps least controversial idea of what one might mean by "good," and apply it in our market setting.

Pareto Efficiency

The specific notion of "good" that we explore in this topic is the notion of "efficiency." We shall ask whether the market allocates resources efficiently, i.e. whether equilibria in the general equilibrium model are efficient. Now we have to say what we mean by "efficiency." We shall offer a very specific definition of efficiency, and when "efficiency" is defined in this way it is also called Pareto efficiency, after the Italian researcher Vilfredo Pareto (1848-1923). It is good to use the expression "Pareto efficiency," because, in every day life, we use the word "efficiency" in may different senses, and if we qualify the word and say "Pareto efficiency" we are reminded that the word is used in a very specific sense.

Here is the definition:

> A use of resources is called Pareto efficient if there is no other way of using the same resources that would increase *all* consumers' utility.

It seems pretty clear and uncontroversial that any use of resources that is not Pareto efficient is not a good use of resources. Note that there may be many different ways of using a given set of resources that are Pareto-efficient. For example, in many examples a Pareto efficient use of resources is to let everyone else work as hard as they can, but let me free-ride, and not do any work. Because everyone else works hard, firms produce a lot of consumption goods. The proposed Pareto efficient allocation passes all that output to me. Everyone else goes empty-handed. Why is this Pareto-efficient? We have to check whether there is some other use of resources that would give *everyone* higher utility. But, presumably, the use of resources that I just described gives me the highest utility I can ever have, for any use of resources. Any other allocation would leave me with less output. Even if everyone else is made better off, a different allocation would cause at least one person to be less well-off, me in this case. If that is so, this way of using resources is indeed Pareto-efficient.

Pareto-efficiency is a pretty weak definition of "good." Many uses of society's resources are all Pareto-efficient (although there are often even more uses of resources that are not Pareto-efficient). We shall ask whether markets bring about Pareto-efficient uses of resources, and, if the answer is positive, we will have shown that markets fulfill some minimum requirement. But we should not interpret this as a very strong statement about the advantages of markets as resource allocation mechanisms.

To make the concept of Pareto efficiency more concrete, let us go back to the first example from Topic 16. Recall that the resources available to society were 2n workers, each of whom was able to work at most 24 hours per day, and a production function that allowed us to transform L hours of labor into $\sqrt{L}$ units of output. Among the n consumers, the first n had utility function:

$$U(\ell, q) = \ell q^2$$

and the remaining n had the utility function:

$$U(\ell, q) = \ell^2 q.$$

There was also a capitalist who only consumed.

To make things very concrete lets set n equal to 5. The market then decided that the first 5 consumers worked 16 hours, the remaining 5 consumers worked 8 hours, so that the total number of hours worked is 120. Thus, we produce a quantity of the consumption good of $\sqrt{120} \approx 10.95$. If we use the calculations from the previous section, the first 5 consumers get each a quantity

$$4\sqrt{\frac{1}{6n}} = 4\sqrt{\frac{1}{30}} \approx 0.73$$

of the consumption good. The next 5 consumers get each:

$$2\sqrt{\frac{1}{6n}} = 2\sqrt{\frac{1}{30}} \approx 0.37.$$

The rest of the consumption good:

$$10.95 - 5 \cdot 0.73 - 5 \cdot 0.37 \approx 5.48$$

goes to the capitalist. Plugging this into the utility functions we have that the first 5 consumers have utility:

$$\ell q^2 \approx 4.26,$$

the next 5 consumers have utility:

$$\ell^2 q \approx 94.72,$$

and the capitalist has utility

$$q \approx 5.48.$$

Here, I have just assumed that the capitalist's utility function is: $U(\ell, q) = q$. We didn't specify that before. All that we needed to know up until now was that he only valued consumption.

Is this a good way to use the resources? This is what we shall investigate now. More precisely we shall check whether the use of resources is "Pareto efficient," that is, whether there is some other way of using resources that would give the first 5 consumers utility higher than 4.26, the next five consumers utility higher than 94.72, and the capitalist utility higher than 5.48.[1]

[1] Note that the outcomes seem much better for the second group of consumers than for anyone else in the economy. But recall that utility is an ordinal concept, and that utility comparisons across individuals are meaningless. Thus, the apparent inequality will not be a concern for us here.

Let us use naive "trial and error" to get a first sense for what the answer to our question should be. There are infinitely many alternative uses of resources. We'll pick one arbitrarily. How about we let everyone work a little less, and enjoy more leisure? So, we could let the first group of consumers only work for 12 hours, the second group of consumers only work for 4 hours, and then only produce $\sqrt{5 \cdot 12 + 5 \cdot 4} = \sqrt{80} \approx 8.94$ units of the consumption good. Of course, that is less than we had before, but maybe people enjoy leisure so much that in equilibrium everybody works too much? To check this, we have to figure out whether we can allocate 8.94 units of the consumption good among all the participants in the economy so that in the end they are better off than they were before. Let's try.

The first n consumers have in the general equilibrium a utility of:

$$\ell q^2 = 8 \cdot 0.73^2 \approx 4.26.$$

In the new arrangement, they have more leisure: $\ell = 12$. how much consumption do they need so that they are better off than before?

$$12q^2 > 4.26 \Leftrightarrow q > 0.56.$$

They need approximately 0.56 units of output, or a little more. Thus, we need to allocate at least 2.8 units of the output to these consumers. The next 5 consumers are better off if they get a q that satisfies:

$$20^2 q > 94.72 \Leftrightarrow q > 0.24.$$

Thus, we need to allocate at least 1.2 units of the output to these 5 consumers. That leaves a total 8.94-2.8-1.2=4.94 for the capitalist. But before he had 5.48. So, if we make all consumers better off, the capitalist will be worse off. You may not like capitalists, or you may be one. Whatever is the case, not *everyone* was made better off, and so our guess is not good. So far, we haven't found an allocation that makes everyone better of than in the general equilibrium.

Of course, we have only tried out one alternative to the equilibrium outcome. There are infinitely more that we could try. But I shall not bother you with further allocations. In fact, in this example, the general equilibrium outcome is indeed Pareto-efficient. Our search for a "better" use of resources, that is, a use of resources that makes everyone better off, is bound to fail. I won't explain to you for the moment why that is the case. You have to take my word for it.

Let's consider the second example from the previous section. In that example, unlike the first example, in the general equilibrium the capitalist had zero utility. So, essentially we can forget about the capitalist. Moreover, in the general equilibrium, all consumers worked for 12 hours, and acquired 6 units of good 1, and 9 units of good 2, giving them a fantastic utility[2] of:

$$U(q_1, q_2) = (q_1)(q_2)^3 = 6 \cdot 9^3 = 4374.$$

[2] Actually: remember that utility numbers don't mean anything by themselves.

To produce the output, setting as before $n = 5$ for concreteness, 15 units of labor went into the production of good 1, and 45 units of labor went into the production of good 2. Can we guess a way of using resources that makes everyone better off? We'll try again.

Remember from the example that, although consumers value good 2 more than good 1, good 2 is harder to produce: for making one unit of good 1 one needs only 0.5 unit of labor, whereas for making one unit of good 2, one needs one unit of labor. Maybe, everyone would be better off if more labor went into the production of good 1, and less labor went into the production of good 2?

Let's consider what would happen if we put 20 units of labor into the production of good 1, and 40 units of labor into the production of good 2. We could then produce 40 units of good 1 as well as 40 units of good 2. If we divided the output equally, this would mean that everyone would get 8 units of each of the two goods. Then their utility would be:

$$U(q_1, q_2) = (q_1)(q_2)^3 = 4096.$$

Comparing this to the utility in the general equilibrium, which was 4374, we can see that instead of making everyone better off, we are making everyone worse off by putting more labor into the production of good 1.

As for the first example, we have only tried out one guessed alternative use of resources, and concluded that it did not make everyone better off than they are in equilibrium. But the general results that we are going to discuss in the next section will imply that indeed, whatever we try, not everyone will be better off. There will be at least some people who are worse off. In this example, too, general equilibrium is Pareto efficient.

The First Welfare Theorem

The discussion in the previous section illustrates one of the most important results of economic theory:

> The First Welfare Theorem:
>
> Under some assumptions, the use of resources in general equilibrium is Pareto efficient.

This theorem describes one simple sense in which the "invisible hand" of the market guides the agents in the economy towards an efficient use of resources, even though each individual and firm only follows their own interests, and nobody plans for the economy as a whole.

We need to discuss two questions: (i) Why is this theorem true?, and (ii) What are the assumptions on which the theorem is based? Regarding the second question, note that in the statement of the theorem we mentioned that we were making some assumptions, but did not say what those assumptions are. So, our second task is to make those assumptions explicit.

I shall say a little bit about why the theorem is true, even though I shall not prove it. The arguments that I offer for the theorem are informal, not precise. The actual, precise proof of the first welfare theorem is actually astonishingly simple. But it does not give much intuition. Moreover, to write that proof down, I would need some definitions and notations which would only distract us at this point.

Let's think for the moment about a somewhat more narrow issue than the first welfare theorem addresses. Let's not question whether the economy produces the right quantities of each consumption good, and let's only check whether, given what has been produced, the assignment of these outputs to the different consumers is Pareto efficient, or whether there would be some other way of assigning the output to consumers that would make everyone better off.

It turns out that the marginal rates of substitution are a good indicator of whether we have assigned goods to individuals in a Pareto efficient way. I shall illustrate why this is so by means of an example. Suppose we consider any two consumers, and calculate their marginal rates of substitution between any two consumption goods, say good 1 and good 2. Call these marginal rates of substitution MRS_1 and MRS_2. Suppose they were not the same number, say:

$$\text{MRS}_1 > \text{MRS}_2.$$

Recall that MRS_1 tells us how much of good 2 consumer 1 is willing to give up in return for one extra unit of good 1. MRS_2 tells us the same thing for consumer 2. But we want to give MRS_2 a somewhat different interpretation. It is how much of good 2 we have to give to agent 2 if we want to take one unit of good 1 away from him. We thus reverse the direction of change: instead of giving consumer 2 one unit of good 1, we take one unit of good 1 away. Because we are talking about derivatives and slopes, that is, margins, the two directions don't differ. The MRS is the same, regardless of which direction we consider.

Now let's imagine the following change in allocations: We take one unit of good 1 away from consumer 2 and give it to consumer 1. We then take MRS_1 units of good 2 away from consumer 1, and give them to consumer 2. On balance, consumer 1 will be neutral towards this change. We take exactly as much away from him as he is willing to give up. Consumer 2, by contrast, will be better off: we have given him more than he needed to be compensated for the loss of one unit of good 1. Thus, we have almost achieved an allocation which both prefer: Indeed, if we take a tiny bit less of good 2 away from consumer 1 than he is willing to give away, then he too is better off, and consumer 2 will remain better off as long as this tiny amount is really tiny.

In other words, we have seen that, whenever

$$\text{MRS}_1 \neq \text{MRS}_2$$

the use of the given output of consumption goods is not Pareto efficient. Thus, it should better be true that in a general equilibrium this will never happen. Otherwise, the first welfare theorem could not be true. But now recall our familiar formula for optimal consumer choice. For consumer 1's choice to be optimal, we need to have:

$$\text{MRS}_1 = \frac{p_1}{p_2}.$$

(Let us ignore corner solutions for a moment.) The same has to be true for consumer 2:

$$\text{MRS}_1 = \frac{p_1}{p_2}.$$

But, if both MRSs are equal to the price ratio, then they are also equal to teach other. That is, indeed, in a general equilibrium, it will be true that:

$$\text{MRS}_1 = \text{MRS}_2.$$

This is an example of the "miracle of the market." People don't consciously arrange their consumption so that their marginal rates of substitution are equal to each other. Everyone only looks at the market prices, and then tries to maximize their own preferences given these market prices. But as a result, inadvertently, the marginal rates of substitution are equalized. This is how the "invisible hand" steers people towards efficiency using market prices as the strings on which it pulls.

Let us now look at the production side of the economy. Suppose there are two firms, making two different goods, goods 1 and 2, both using labor and capital as inputs. First, we shall argue that Pareto efficiency requires that these two firms have identical marginal rates of technical transformation. Suppose, indeed, these were not the same for both firms. Then one firm could give up more capital for one extra unit of labor, holding output constant, than the other firm would require in capital if we took one unit of labor away from it. So, we should take one unit of labor away from the latter firm, give it to the former firm, and compensate by transferring capital. Both firms would be better off, in the sense that they would produce larger output.

Thus, Pareto efficiency requires that for the two firms the marginal rates of technical substitution are the same:

$$\mathrm{MRTS}_1 = \mathrm{MRTS}_2.$$

In an equilibrium, both firms satisfy our familiar condition that the marginal rates of technical substitution must be equal to the factor price ratio:

$$\mathrm{MRTS}_1 = \frac{w}{r}, \text{ and } \mathrm{MRTS}_2 = \frac{w}{r}.$$

This means that the "invisible hand of the market" has guided the two firms to a situation where they have the same marginal rate of technical substitution, even though the two firms did not really think about this, but only sought to maximize their profits.

Informally, one says that this last paragraph shows how a market equilibrium ensures "efficiency in production." The argument before shows how a market equilibrium ensures "efficiency in exchange," where the word "exchange" alludes to the fact that we are considering a situation where the total quantity of goods is given and fixed, and all that we are considering is whether exchange among the two agents could make both agents better off.

Let us consider another aspect of efficiency: it is whether the "right" things are produced. We could imagine an economy that is very efficient at choosing how to allocate scare resources to different production processes, and that is also very efficient in allocating given output to the consumers, but that, unfortunately, produces goods that, in fact, nobody wants. Thus, we also want to consider the efficient coordination of the production and the consumption side.

Again, let us think about a very simple scenario. Suppose there was one factor only, labor, but it could be used to produce two different goods, good 1 and good 2. If we allocate output efficiently, then everyone will have the same marginal rate of substitution, say MRS. I claim that efficiency requires that:

$$MRS = \frac{\frac{\partial f_2}{\partial L}}{\frac{\partial f_1}{\partial L}}.$$

In words: the marginal rate of substitution must equal the ratio of the marginal products of labor in the two industries, the one producing good 1, and the one producing good 2.

Why do they have to be equal? We have to first think about what the ratio of the marginal products means. The ratio of the marginal products tells us how much less of good 2 we would have to produce if we took labor away from the production of good 2, and put it into the production of good 1, so that we can produce one extra unit of good 1. To see this more concretely, suppose that the marginal product of labor in the first industry, i.e. $\frac{\partial f_1}{\partial L}$, were 5, and the marginal product of labor in the second industry, i.e. $\frac{\partial f_2}{\partial L}$, were 2. To produce one more unit of good 1 we would then need $1/5$ units of labor. If we take those away from the production of good 2, we would lose $2 \cdot (1/5) = 2/5$ units of good 2. This is exactly the ratio of the marginal product of labor in industry 2 divided by the marginal product of labor in industry 1.

Now suppose the MRS were not equal to the ratio of the marginal products. For example, the MRS were larger than the ratio of the marginal products. Then we could produce one extra unit of good 1, and what we would lose in the production of good 2 would be less than consumers would be willing to give up. That is, all consumers would be better off, because they don't have to give up as much as they are willing to give up.[3]

[3] This assumes, implicitly, that we divide the amount of good 2 that needs to be given up equally among the consumers.

We have now convinced ourselves that Pareto efficiency requires:

$$MRS = \frac{\frac{\partial f_2}{\partial L}}{\frac{\partial f_1}{\partial L}}.$$

Now, because, in equilibrium, consumers choose optimally, the left hand side equals the price ratio p_1/p_2. How can we relate the ratio of the marginal products to the price ratios? Recall from our discussion of profit maximizing choices by firms, that a necessary condition for profit maximization in firm 1 is that:

$$\frac{\partial f_1}{\partial L} = \frac{w}{p_1},$$

and that:

$$\frac{\partial f_2}{\partial L} = \frac{w}{p_2}.$$

But, if we divide the second equation by the first equation, we find:

$$\frac{\frac{\partial f_2}{\partial L}}{\frac{\partial f_1}{\partial L}} = \frac{\frac{w}{p_2}}{\frac{w}{p_1}} = \frac{p_1}{p_2},$$

and thus both the MRS and the ratio of marginal products equals the price ratio, and thus they are equal to each other. Market equilibrium satisfies the necessary condition for efficient coordination between consumption and production.

The Conditions Under Which The First Welfare Theorem Is True

My statement of the first welfare theorem above was very vague: "under some assumptions ..." We now need to discuss what those assumptions are. I shall give you a list of three crucial assumptions.

The first is that firms and consumers are “price takers,” that is, all participants in the economy take prices as given, and don’t think that their own actions can in any way change the prices. Thus, for example, firms think that which price tag they have to put on their products is just dictated by market conditions. They don’t think that perhaps they can raise their prices, and thereby attract fewer, but not zero customers. Consumers don’t think that they can negotiate prices, in any market. Instead of saying: “we assume price-taking behavior by consumers and firms,” people also sometimes say: “we assume perfect competition.” These two phrases have exactly the same meaning.

Perfect competition is an unrealistic assumption. We also know situations in the real world where the assumption is patently not true. Some researchers maintain, though, that for certain purposes it is a good enough approximation of what we see in economic data. If we write a model in which the assumption is violated, which we shall indeed do later in this course, then the first welfare theorem is not true, and we can reach Pareto inefficient outcomes. This is because firms and consumers no longer respond to the market prices as they are, but instead they respond to their anticipation of how prices might change if they chose different quantities.

A second crucial assumption is that an equilibrium exists. Now, we have never encountered in this course, and will never encounter, a situation in which equilibrium does not exist. But this is because this is an introduction, not an advanced course in microeconomics. In an advanced course in microeconomics it would become clear that this, too, is an important, and somewhat restrictive assumption. When you study a system of millions and millions of markets, that all potentially interact with each other, the question whether one can find a price system such that in each and every market supply equals demand does not have an obvious answer. Whenever you try to adjust prices in one market, you potentially create a disequilibrium in some other market. Will this ever end? The answer is that we know some conditions under which equilibrium exists, but they are not without loss of generality. For example, these conditions include the assumption that consumers have decreasing marginal rates of substitution, which, as we discussed earlier, does not always seem plausible, at least on the basis of casual introspection.

The final assumption is perhaps even more hidden and implicit than the assumption that equilibrium exists. It is the assumption that for every good there is a market. But, if we let our imagination range freely, we can think of many goods for which there does not seem to be a market. For example, I would really like to buy a car that is painted in pink and blue stripes. But I just can't. The market for such cars does not exist. More seriously, perhaps you would like to buy an insurance that pays out some money to you if you get a bad grade in an exam, just to make you feel better. No market for such an insurance policy exists.

The lack of markets in the real world is even more subtle than this. To explain, let me introduce the phrase "externality." Many of our economic actions impose "externalities" on others. For example, when I listen to loud music in my office at the university, my office neighbor will be disturbed. The walls are thin in our part of the building. Or if I have a barbecue in my front yard, then I either benefit my neighbor, because she likes the smell, or perhaps I harm her, because she dislikes the smell. (Externalities can be "positive," or "negative.") Or, if I drive my car to work in the morning, I produce air pollution, which harms the environment that the rest of the world would like to enjoy. That is a negative externality that I impose on many other people, without ever asking them for permission.

This last phrase, "not asking them for permission," is the key. If there were a market for the right to have a barbecue in my own yard, then I could approach my neighbor, and ask her, before I have the barbecue, how much she is willing to pay, or how much I have to pay her, for my right to have the barbecue. And then the first welfare theorem would be true again. It is not the existence of an externality by itself that causes an inefficiency, it is the lack of a market in which this externality is traded, that causes the inefficiency.

Thus, when we say that there needs to be a market for every good, we really have in mind a lot of markets: markets for strange objects, insurance markets, and markets for externalities. Once some such markets don't exist, we have, in the language of microeconomic theory, "incomplete markets," and incompleteness of markets typically causes inefficiencies.

Completeness of markets is such an unrealistic assumption that we best treat the model of perfect competition in which the first welfare theorem holds as a benchmark. Some researchers, though, have also claimed that the data make the economy look as if there were complete markets. But that is not a claim that we can explain in detail, or assess, in this course.

The Second Welfare Theorem

Saying that every equilibrium of an economy with perfect competition will result in a Pareto efficient use of resources is not saying much. As we mentioned earlier, very unequal uses of resources, such as giving everything to me, is Pareto efficient. But still, society may not view this as a satisfactory use of resources. In particular, we may be concerned about distributional issues. In a vague sense, we might want people to lead "equally satisfactory" lives. Economics is sometimes misunderstood, or misrepresented, as a field that is not interested in inequality. But there is nothing in economics that would give reasons for saying that inequality may not be a legitimate concern in economic policy. We are discussing normative issues, and whatever research has been done by economists into the foundations of normative judgments does not at all imply that inequality or equality are irrelevant subjects.

Let's put ourselves into the shoes of a fictional policy maker, to whom we, as economists, have gone to advocate for market solution to the resource allocation problem, presenting as our argument that the market solution will be Pareto efficient. Suppose the policy maker forgets to ask us whether all the assumptions of the First Welfare Theorem are satisfied, but instead says that she is not satisfied with Pareto efficiency. She says that, yes, the market will result in a Pareto efficient allocation, but that this allocation will be very unequal, favoring some, and offering others almost nothing to live on, and that she thinks this is unacceptable. How might we respond to her?

There is another result in welfare economics that could come to our help. Not surprisingly in fact, given that the previous result was called the *first* welfare theorem, there is a second welfare theorem. It says the following:

> The Second Welfare Theorem:
>
> Under some assumptions, every efficient use of resources can be made into an equilibrium through appropriate redistribution of purchasing power.

How does this help us in addressing our fictional policy makers concerns? We could say to her: Pick whatever use of resources you deem equitable and desirable. For sure, you don't want to pick an allocation that is not Pareto efficient. For sure, you just want to pick a different allocation among the Pareto efficient allocations than the market would produce under current conditions. Now that you have picked it, you can still use the free market to achieve it. All that you have to do is redistribute purchasing power towards those who, in your opinion, would otherwise be too disadvantaged. But your concern for equality does not speak against the use of markets by itself, it only says that, before markets operate, you might want to redistribute purchasing power.

This claim is an important complement to the first welfare theorem, and it strengthens the case for market solutions to resource allocation problems very significantly. But like in the case of the first welfare theorem, there are a number of important questions that we need to address right away. The first is: what is an "appropriate redistribution of purchasing power?" What the theorem refers to is a system of taxes on some people, that are used to finance transfers to other people. most importantly, though, this system of taxes and transfers must affect in any way consumers' or firms' responses to prices. In other words: How much taxes someone has to pay, or how much transfers someone receives, must not depend at all on their economic choices. For example, we must not tax the consumption of a particular consumption good. This would mean that the price of that good relative to the price of other goods would increase for consumers, but go down for firms, as we saw in 15. But then the marginal rates of substitution and the marginal rates of technical substitution among different goods would not equal the ratio of the marginal products of, say, labor, in the production of those goods. The argument that we gave earlier for why this is true under perfect competition crucially assumed that consumers and firms respond to the *same* prices.

Taxes must also not be based on labor supply choices, for the same reason.

Thus, “appropriate” transfers in the language of the second welfare theorem do not include income taxes, one of the most important forms of tax most of us face. As you can see, our choices for “appropriate” taxes and transfers become more and more narrow the longer we think things through.

Taxes and transfers that are not related to any economic choice that agents make are called lump sum taxes and transfers. Redistribution involving such taxes and transfers is what the second welfare theorem means by “appropriate redistribution of purchasing power.” But lump sum taxes and transfers would be extraordinarily hard to implement in practice. We would have to say to someone: we are going to tax you not based on your income, or based on your consumption, but simply because we have investigated your DNA, and therefore we know you are able do very well in the market system, perhaps because you are very talented. But you have to pay the tax just because of your DNA, regardless of how you actually make use of your talents. Sometimes governments come close to introducing such taxes, for example when the British government imposes a tax called the “BBC license fee” on essentially all its citizens. But lump sum taxes and transfers are honestly very rare in practice.

So, now the second welfare theorem seems less persuasive. Let us dig further. There is also the vague phrase “under some assumptions.” This is not a course in which I could explain in detail what those assumptions are. But they are in fact very similar to the conditions under which we know a competitive equilibrium exists. They include, for example, the condition that marginal rates of substitution are decreasing. So, this further weakens the appeal of the second welfare theorem. They also include, implicitly, the assumption that markets for all goods exist. As with the first theorem, we should think of it as a benchmark, rather than a realistic theorem.

One way of phrasing the second welfare theorem is: In ideal conditions, we can separate the problem of efficiency and the problem of equality. We deal with equality through transfers. Then we let the market create efficiency. As you now know, this is only valid in a somewhat unrealistic benchmark situation. But maybe we are not too far from that benchmark in the real world, or maybe we are very far away. I don’t know the answer. The idea of separating efficiency and equality is, I believe, deeply ingrained in many economists’ minds.

Topic 18: A Difficulty in the Concept of Social Welfare

Welfare

In the previous topic we have seen how general equilibrium under perfect competition leads the economy to one of the Pareto efficient uses of general equilibrium. We have also seen that, if lump-sum redistribution of purchasing power is possible, we can move the equilibrium to another Pareto efficient allocation, perhaps because of inequality concerns. But how can we decide whether this redistribution is "good?" Some people will be better off, some other people will be worse off. Is it worthwhile?

Before discussing how one might answer this question let me raise another, similar question. Suppose lump-sum redistribution is impossible. We then have to use distortionary taxes such as the income tax to redistribute income. We know that this will take us typically away from Pareto-efficient allocations. But maybe some other purpose is achieved. Perhaps some more desirable state of equality is achieved. Is it worthwhile?

Often, economists discuss this question using a concept called "welfare." They ask whether a given policy enhances welfare, and they also discuss by how much welfare is increased, or decreased. But what do we mean by welfare?

If you follow the discussion in this topic carefully, you will emerge from it more skeptical towards what economists say on these subjects, including what they say on "welfare." We shall see that economic theory includes a result that should make us deeply skeptical of all concepts of welfare. In fact, this result should make us deeply skeptical whenever economists make value judgments, i.e. say what "should be done." The result that I shall introduce you to in this topic is very famous and well-known. Yet the difficulty that the result uncovers, is often ignored by economists. In the next topic we shall discuss one way in which, perhaps, one might bypass the difficulty, and provide foundations for economists' normative statements. But first I want to explain to you the result that I am talking about.

What might we mean by welfare? Suppose there is a given set of uses of society's resources. Let's use some new notation, and let's call this set simply

"A." Each element of this set is a description of one particular way of using people's time, including how much everyone is supposed to work, a specification of which of the many possible products are produced, and how much of each of them is produced, and also one of the many ways of distributing what has been produced among consumers. As we saw earlier, markets pick one of the elements of this set. To define what "welfare" is, we could seek to assign to each of these possible uses of society's resources a number called "welfare." What would this number mean? In this topic we shall treat it as if it was "society's utility function."

Like utility functions, the welfare numbers would then have no meaning, and really only the comparison of different elements of A would be meaningful. In other words: one possible use of resources has welfare 8, and another has welfare 16, does not mean that the second one has twice as much welfare as the former one, it only means that the second one is preferred over the first one, from society's point of view.

What we are really seeking to do is thus to create a society's preferences over the set A. Where should they come from? The most common answer within economics, and the answer that we shall accept here, is that society's preferences should, in some way, be based on consumers' preferences, and on those preferences only. Of course, how we derive society's preference from consumers' preferences is still to be determined. What I am saying here is that we should not take any other information into account except consumers' preferences. Society's preference should not reflect that society judges alcoholic drinks as "bad" even though consumers like them, because society thinks that clouding your mind is bad for you. Society values whatever consumers value. We also say: "society is not paternalistic," because society does not act like a father who thinks he knows better than his children what is good for his children.

There is an important principle embedded in the approach that we have just described. We mentioned it before. We are postulating that the only purpose of the economy is to satisfy your and my preferences. There is no other purpose. Policy objectives such as "growth," "low inflation," "balanced government budget," etc., are valid only in as far as they help individual consumers achieve what they want to achieve. If our preferences suddenly change, and we all value leisure much more than we did in the past, then the economy should shrink, rather than grow. Nothing would be wrong with that shrinkage. It would just reflect our new preference for a life of leisure.

Now suppose I ask everyone how they rank all possible elements of A, that is, about their preferences. Suppose then I sit down, remain silent for a while, perhaps think during that period, or perhaps not, and then announce what in my opinion society's preferences should be. You might wonder what went on in my mind during that period of silence. What did I think about?

What sort of answer might I give you? For example, I could say, I only looked at what my two friends said, and not what anybody else said. I then checked first which preferences they had in common, and decided to make those preferences part of society's preference. And for all other comparisons, I sometimes gave friend 1 priority, and sometimes friend 2. What I have then described to you is a *procedure* for defining welfare. I have not just explained to you my proposal for society's order in the circumstances that I found myself in, but I explained to you a general procedure that you could apply in other situations, that is, for other preferences of individuals as well.

I should add: Whenever I gave one friend priority, I also included all comparisons that were implied by transitivity. Only with this caveat leads my procedure to a transitive order for society.

After I gave you my answer, you might wish to test what my procedure would yield as society's preference in other circumstances, and check whether it is reasonable. That, in fact, seems to be the main way in which we discuss judgments over what is desirable for society. You rightly expect from me some "principled reasoning" that you can discuss, and it seems that "principled reasoning" must take the form of a procedure.

I have to admit, I actually don't have a procedure ready for you for constructing society's preference over A. But let's search for one together.

Let me try to describe symbolically what we are looking for:

consumer 1's preference consumer 2's preference ... consumer n's preference	The "procedure" $\longrightarrow$	society's preference

This looks a little bit like mathematics. And, indeed, it is possible to define mathematically what one means by a "procedure." It is a generalized sort of "function." The functions that you are familiar with map one number, or one list of numbers, into another number, or another list of numbers. Here, the function that we are looking for maps a list of consumers' preferences into society's preference. It is a function because it should be defined for all possible lists of preference, not just for one, just as the function $f(x) = x^2$ assigns not only to the number 3 the number 9, but it assigns to every arbitrary number another number.

To summarize, instead of saying that we are looking for a "procedure" that describes how we define "welfare," we can say that we are looking for a function that maps every list of individual consumers' preferences into society's preference. We won't be completely formal here. I want to put the idea of a function that describes how social preferences are constructed into your mind, but I shall not define it rigorously. This would take too much time here. Let me nonetheless use the word "social welfare function"

Most economic researchers mean by a "social welfare function" something slightly different from what I mean by it here. But calling the function that we have been talking about a "social welfare function" seems at this point of our discussion so natural that it seems excusable that I deviate from the conventional terminology in these notes.

Let's move on, and pretend we knew exactly what we mean by a procedure, or a social welfare function. How can we find a reasonable one? You and I might each propose one, and then we might compare them, and exchange arguments. But there's a more systematic way of looking for a procedure. We can write down some properties that we would like our procedure to have, and then we might ask which procedures have those properties. We might continue, writing down more properties, until we have figured out just one remaining procedure. Let us follow this approach.

The Condorcet Cycle

The first desirable property is this:

> Our procedure should assign to every list of individuals' preferences a complete and transitive preference for society.

To explain this property, and why it is desirable, let me give an example of a procedure that does *not* have this property, and that is thus ruled out by this first postulate. It is the procedure which determines society's preference through majority voting. More precisely, suppose, for every pair of alternatives, *a* and *b*, the social preference ranked *a* over *b* if and only if a majority of all individuals ranked *a* over *b*.

For what follows, it does not matter how ties are resolved in majority voting.

What can go wrong with this? Here is a simple example. Suppose the set *A* had only three elements. This just makes the example simple. Let the elements of *A* be denoted by *a*, *b*, and *c*. Suppose also there were only three people. And suppose these three people had the following preferences:

Person 1	Person 2	Person 3
a	c	b
b	a	c
c	b	a

This table means, for example, that person 2 ranks *c* above *a*, and *a* above *b*.

Now let's use majority voting to determine society's preference. How will we rank *a* and *b*? Well, persons 1 and 2 prefer *a* over *b*, whereas person 3 prefers *b* over *a*. Thus, a majority prefers *a* over *b*, and therefore majority voting leads to the conclusion that society ranks *a* over *b*. Similarly, you can verify that a majority prefers *b* over *c*.

We might now argue that if society ranks *a* over *b*, and *b* over *c*, then it must be that society ranks *a* over *c*, and therefore, we don't even have to check whether a majority of people prefers *a* to *c*, or whether it is the opposite. But let's check it anyway. What do we find? Person 1 does prefer *a* over *c*, but persons 2 and 3 actually prefer *c* over *a*. So, majority voting would say that *c* should be preferred over *a*, the opposite of what we expected to find!

But now the preference determined by majority voting is not transitive. We found that *a* is preferred over *b*, that *b* is preferred over *c*, and *c* is preferred over *a*. That makes no sense. How could society evaluate policies on this basis? If our current status quo is *a*, because society prefers *c* over *a*, we might advocate a policy that brings about *c*. Then, when we are in state *c*, we might advocate a policy that brings about *b*, because society prefers *b* over *c*. And then we could advocate for a policy that brings about *a*, because society prefers *a* over *b*. Continuing, we might endlessly cycle among *a*, *c*, and *b*, at each step thinking that we improved society's welfare. This is not sensible. Society should have a transitive preference.

Recall that we defined and discussed transitivity in Topic 2.

Thus, we have found that majority voting violates our first desirable property. The example that we have used to demonstrate this is called the Condorcet cycle.

Condorcet was an 18th century philosopher and mathematician.

It is very disappointing that we have to rule out the seemingly natural procedure of majority voting. But let's not be ruled by emotion. Let us simply continue our quest for a good procedure for constructing society's preferences. Perhaps we find a sensible alternative to majority voting?

Before continuing, one final point for this section: the first desirable property also requires society's preference not only to be transitive, but also to be complete. The purpose of this is to rule out cop-out such as the one saying that in the Condorcet cycle society simply does not rank a and c. We don't want to allow such a "cheap" resolution of the Condorcet cycle problem.

Arrow's Impossibility Theorem

Here is another cop-out that we could use to avoid the Condorcet cycle problem described in the previous paragraph. We could decide that society's preference is that a is preferred over b and that b is preferred over c, regardless of what people's preferences are. This might be based on the opinion that "objectively" a is better than b, and b is better than c, regardless of what anybody thinks. This would avoid the intransitivity problem. But this would clearly contradict our earlier argument that society's preference should be based on individuals' preferences, and not on paternalistic judgments. A simple requirement that captures this is as follows:

I seem to like the word "cop-out." But what is a "cop-out," and what is just simple, but sensible? I don't think I have a precise answer.

> Our procedure should respect unanimity: If everyone prefers a over b, then society's preference should also rank a over b.

Let's accept this as a desirable property. Here is "cop-out" one that leads to transitive preferences, *and* respects unanimity: society's preference is always that of person 1, regardless of what everyone else's preference is. This obviously creates a transitive preference for society, and it also respects unanimity. What's wrong with this? It is difficult to put one's finger on it, but it just does not seem fair. Let's not search for any deeper justification, but let's just write down directly that we don't like this procedure:

Society's ranking should be non-dictatorial, that is, the procedure should not just be: we take person X's preferences, and those are society's preferences.

Let's see whether we have written down all desirable properties for the procedure for defining society's preferences, or whether we have missed out on anything. Here is an example of a procedure that has all properties that we have described so far, and that we still might feel skeptical about. Suppose that there are four elements of A, and just three individuals. Suppose we ask everyone to rank the alternatives a, b, and c, and that we then add up for every alternative its rank in all three individuals' preferences. This allows us to construct a preference for society. Here is an example:

Person 1	Person 2	Person 3
a	d	d
b	a	a
c	b	b
d	c	c

Alternative a is ranked first by person 1, and second by persons two and three. So, the sum of its ranks is: 1+2+2=5. Similarly, we get for alternative b the sum of ranks: 8, for alternative c we get: 11, and for alternative d we get: 6. So, how would we rank the four alternatives? Obviously, a smaller sum is better. So, we would rank a over d over b over c. You can check that this procedure satisfies all the desirable properties that we have listed so far.

The procedure described here is known in the literature as the Borda count. Borda was another 18th century French mathematician.

Is anything wrong with it? Well, consider a different list of preferences:

Person 1	Person 2	Person 3
a	d	d
d	a	a
b	b	b
c	c	c

You can check yourself how society would now rank the alternatives: d over a over b over c. Thus, a and d have flipped rankings. Why? It is because d has moved up in person 1's ranking. But, person 1 might say I still prefer a over d. So, why do you change society's ranking of a and d, if my ranking has not changed? To answer this question, we might argue that person 1 previously ranked b and c between a and d, thus indicating that she really strongly disliked d, whereas in the second case it seems that maybe she no longer dislikes d as much. But if person 1 has taken an intermediate microeconomics class, she will respond: you are misinterpreting what I have told you. You are interpreting my preference as if it had some cardinal meaning, as if I expressed how strongly I prefer one alternative to another. But recall from intermediate microeconomics that all preferences are *ordinal*. You can't infer any cardinal preferences from them. It might have been that in the first scenario I was approximately indifferent between all alternatives, whereas in the second scenario I strongly prefer a over all other alternatives. Then your construction of social welfare makes no sense. And just from knowing my ordinal preferences, you just cannot tell whether this is the situation, or whether I really strongly dislike d only in the first but not in the second scenario.

We shall take the view here that person 1 just won the fictional debate. The appeal to intermediate microeconomic theory must win, given that these notes are about intermediate microeconomics. We shall therefore add a fourth desirable property to our list. This property rules out the procedure that we just described.

> Society's ranking of two alternative a and b should only depend on how every individual ranks a and b. It should not depend on how individuals' rank other alternatives.

This property is sometimes called Independence of Irrelevant Alternatives. Here is one way of thinking about it: suppose we have the task to decide whether we want to go to Northern Michigan, or to Chicago, for a brief weekend break. We ask everyone how they rank these two options. But then suddenly I ask you: how about Florida? How would you compare Florida to these destinations? You might say: "Don't change the topic, Tilman. We don't have the money to travel to Florida. You are asking me about an irrelevant alternative." If you argument is valid, then the principle of independence of irrelevant alternatives is a good property.

So, how much progress have we made? Have we narrowed down the set of procedures that we need to consider for constructing social welfare? It turns out, we have not just narrowed down the set quite a lot, we have weeded out *all* procedures!

The theorem assumes that A is finite. In our model of resource allocation, i.e. the general equilibrium model, there are infinitely many ways of allocating resources. Does Arrow's theorem apply? The answer is "yes," but some more mathematical apparatus would be needed to state the theorem for the case of an infinite set A. The theorem also assumes that A has at least two elements. If A had only two elements, then majority voting, for example, would be a perfectly fine procedure for defining social welfare. This is because, with only two alternatives, obviously no Condorcet cycle can arise. The Condorcet cycle needs at least three alternatives.

Theorem:
Suppose A is finite, and has at least three elements. There is no procedure that has all four desirable properties.

This is quite extraordinary. The four properties that we have listed really seem rather harmless desiderata. Yet, no procedure has them all.

Now I can explain to you how I chose the title for this topic: "A Difficulty In The Concept of Social Welfare." This is actually the title of a paper published in 1950 in the economics journal *Journal of Political Economy.* It is a quite an innocent title, but it contained a proof of the above result. The author was Kenneth Arrow, and this paper made him famous. The result lead to a huge amount of further research, all of which looks for ways how to proceed in the light of this result. The result did not just describe "a difficulty," it described *the* central difficulty in constructing a concept of social welfare. Arrow later won the Nobel Prize in economics, for this result, and for other achievements. He was arguably the most important economist of the 20th century. The theorem is known as Arrow's impossibility theorem.

I wrote that Arrow "proved" the above "Theorem." All of this suggests again that we are dealing with mathematics, and indeed, Arrow's proof is a piece of mathematics. As I explained earlier, the concept of a "procedure" for constructing a social preference is a mathematical concept, similar to the mathematical notion of a "function." Once this has been properly defined in our context, one can also define the three properties that we have listed formally, and then one can mathematically prove the theorem.

I should explain the proof to you. But this would go too far here. To understand the result better, you might try out various procedures for constructing social welfare that sound plausible to you, and check whether these procedures have all the properties that Arrow listed. I shall ask you to do some of this in the problems for this topic.

A central idea in Arrow's proof is this: The second to fourth desirable properties imply that the procedure for creating society's welfare ranking must be something that "looks a lot like" majority voting. But problems similar to the Condorcet cycle actually rule out not just majority voting, but all procedures that "look a lot like" majority voting.

Topic 19: Equivalent Variation, Compensating Variation and the Surpluses

We began our discussion in the previous topic by asking how economic policy measures that take us from one Pareto efficient outcome to another, or policy measures that redistribute income, but result in an outcome that is not Pareto efficient, should be evaluated. Arrow's impossibility theorem shows that coming up with a plausible notion of "welfare" that allows us to assess the welfare consequences of different policies is hard. But do we have to give up? In this section we shall describe some ways in which we can by-pass Arrow's impossibility theorem and construct some, perhaps reasonable, ways of evaluating economic policies.

Suppose we were considering a policy that makes *everyone* better off. Couldn't we agree that it is a good policy? Even if we cannot come up with *complete* welfare rankings, it should be uncontroversial that if one outcome is preferred by everyone to another outcome, the former outcome should have the higher "welfare." The unanimity property in Arrow's theorem expresses this idea, and the impossibility arises only if we try to combine this property with other desirable properties.

Obviously, almost no policy will make everyone better off. If this was our criterion, we would be stuck in the status quo. But here is another idea: When evaluating an economic policy, we could ask whether the policy would at least make it *possible* to make everyone better off? Specifically, if the policy were accompanied by a redistribution of purchasing power among the consumers in the economy, could we find a redistribution that makes everyone better off? This is a hypothetical question. If the redistribution takes actually place, then we are back to the unanimity criterion. But, perhaps, we could take the fact that a redistribution that makes everyone better off is *conceivable* as an indication that a policy is desirable, even if this redistribution does not take place?

Suppose there is a policy which benefits some people, and hurts some others. We could ask everyone who benefits the following question: How much money could I take away from you after the policy has been implemented, and you would be as well off as before the policy was implemented? This is a measure of how much, in Dollars, the policy is worth to the person. We could also ask everyone who is hurt by the policy the question: How much money would I have to give you after the policy has been implemented, and you would be as well off as before the policy was implemented? This is a measure of how much, in Dollars, the policy hurts someone.

If the total amount of benefits, in Dollars, is larger than the total amount of losses, in Dollars, then a re-distribution which benefits everyone is possible. I could take away from the beneficiaries money, slightly less than they have said their benefit is, so that they remain better off than before the policy. I could distribute this money among those who lose, and pay them slightly more than they lose in Dollars, and so they would be better off, too. Thus, my "cost-benefit" test measures whether it would, in theory, be possible to make everyone better off.

You may rightly ask what is the purpose of such a hypothetical test: it refers to a re-distribution that does not actually take place! Why should it be relevant? Well, to justify the test we could say that we use the amount of money that would be needed to compensate people and to make everyone better off as a measure of the "strength" of their preferences. So, if a policy hurts someone, we can take the amount of money that we would have to pay them to compensate for this loss as a measure of how strong their opposition to the policy is. And we could take the amount of money that we could take away from someone after the policy, so that they are as well off as before, as a measure of the strength of their preference in favor of the policy.

If these ideas are valid, why could we not use them in the previous section to find a procedure for constructing social preferences that satisfies all of Arrow's axioms? One obvious reason is that in the previous section we described the set of possible allocations quite abstractly, as arbitrary elements of some set A, whereas the ideas which we described above involve money. A second reason is that a social preference that is based on the hypothetical test that we are discussing violates Arrow's *independence of irrelevant alternatives.* The intuition if quite straightforward: how we rank alternatives a and b depends not just on individuals ordinal preferences over a and b, it also depends on their ranking of alternatives in which we give, or take away, money, so as to compensate them for the move from a to b. If you prefer a to b, but you prefer b to a if I give you 100 Dollars, then I have introduced a third alternative, $b + \$100$, and have used your ranking of this alternative to construct society's preferences over a and b. Thus, we are giving up independence of irrelevant alternatives now.

But earlier I had, perhaps, persuaded you that independence of irrelevant alternatives is desirable. The argument was based on the teaching of intermediate microeconomic theory that preferences are ordinal, not cardinal, and that "strength of preferences" is alien to ordinal utility theory. In this section we adopt a different view. We allow that the willingness to pay for some policy change is a measure of the strength of the preference for, or aversion to, this policy change by an individual. This idea appears plausible. But its philosophical basis is still unclear. Maybe, the difference between 300 Dollars and 3,000,000 Dollars is not that important to me? I wish.

We shall leave these questions for some other time, some other place. Here, we shall just go along with the approach that I described, without questioning it further.

Money-Metric Utility

Our objective is to describe utility loss and utility gain in Dollars (or in Euros, or in Renminbis, etc.) I have emphasized a lot in this course that preferences are an ordinal concept, that utility functions are just representations of preferences, and that utility numbers by themselves don't mean anything, and that the fact that one utility is twice as large as the other utility is not informative at all. Now, in a sense, I throw overboard my principles, and we are going to measure utility in Dollars, and we will give the Dollar numbers some practical meaning.

The idea is very simple: You will remember from topic 5 an object called the "expenditure function:"

$$E(p_1, p_2, \bar{U}).$$

It is the lowest expenditure level with which you can achieve utility level $\bar{U}$, if you choose your consumption judiciously, and if the prices of the two goods are p_1 and p_2 (and, of course, we consider a two goods world rather than a million goods world just for simplicity, not because anything substantial would change if we allowed a million goods).

Instead of saying that the expenditure function assigns a Dollar expenditure to a utility level $\bar{U}$, we could, of course, also say that the expenditure function assigns a Dollar expenditure to the indifference curve that represents utility level $\bar{U}$. Recall that we determine the expenditure function by finding the "cheapest" consumption bundle on the indifference curve corresponding to utility level $\bar{U}$.

Now consider this: suppose we start with a consumption bundle (q_1, q_2). We determine the corresponding utility level $U(q_1, q_2)$. And then we ask: "what would be the cheapest way of achieving utility level $U(q_1, q_2)$?" It may be that the cheapest way of achieving this utility level is not consuming (q_1, q_2) but to consume some consumption bundle that is on the same indifference curve as (q_1, q_2). In any case, the corresponding expenditure level is:

$$E(p_1, p_2, U(q_1, q_2)).$$

Notice that now we have found a way to assign to each consumption bundle (q_1, q_2) a money amount. Moreover, if two consumption bundles have the same utility level, they also get assigned the same amount. Finally, if one bundle has higher utility than another, then achieving the higher utility will also require higher expenditure, and therefore, it will get assigned a larger money amount. But all of this means that the function that we have just constructed is itself a utility function, that is, it represents the same preferences as $U(q_1, q_2)$. Or, put differently, the expenditure function is a monotone transformation of the utility function. Every monotone transformation of a utility function is itself a utility function.

The claim in the text that higher utility requires higher expenditure may seem obvious. But actually this claim assumes that the utility function is increasing in quantities. To see this, suppose there was just one good, and my utility were the negative of the amount of that good. That is, the less I have of that good, the better. Then maybe to achieve utility level 5 I have to buy 5 units of the good, but to achieve utility level 4, I only have to buy 4 units of the good. Thus, in this case, in which utility function is not monotonically increasing, the expenditure level required to achieve a utility level is lower for higher utility levels.

Thus, we now have a utility function that is expressed in dollars. It assigns to each consumption bundle (q_1, q_2) the Dollar value $E(p_1, p_2, U(q_1, q_2))$. It is called the money-metric utility function. We are going to use this function to measure how much, measured in Dollars, people are better off, or worse off, if their consumption changes.

Obviously, the money metric utility function depends on the prices p_1, p_2. If they change, we get a different money metric utility function. But let's for the moment just imagine we had picked those prices. Then we would have a Dollar valued utility function, as we wanted. The fact that prices enter as well will play a role momentarily, in the next sections.

Equivalent and Compensating Variations

Now let's consider some policy intervention. Suppose, for example, we force a consumer, who so far has consumed (q_1, q_2) to consume instead $(\hat{q}_1, \hat{q}_2)$. We just tell the consumer: you now have to eat more broccoli and less chocolate, say. The consumer will complain that she is now worse off, because she doesn't really like broccoli. But by how much is she really worse off? Let's imagine that this consumer's choice alone does not affect the market prices p_1 and p_2. The consumer is a small fish in a big pond. Then a natural way of measuring the loss of utility in Dollars uses the money-metric utility:

$$E(p_1, p_2, U(q_1, q_2)) - E(p_1, p_2, U(\hat{q}_1, \hat{q}_2)).$$

It is this sort of calculation that we will use to evaluate policy interventions. Let us therefore look at the above formula very carefully. If the consumer initially, before we intervened, chose optimally, then the first term in the difference will be the consumer's income level, by the duality argument that we discussed when considering consumer theory. Now the new consumption that we force on the consumer, more broccoli and less chocolate, may actually cost the consumer more than his income, and we might have to give him additional money. But when measuring his utility loss we ask a different question. We ask: if we did not force you to change your consumption, but instead just took away some of your income, which income loss is equivalent to the utility loss that you suffer eating more broccoli and less chocolate?

So far this seems plausible. But now let us look at a more complicated case. Suppose we are looking at a policy intervention that changes some of the prices, Perhaps some taxes are introduced, for example. I shall focus on the case that the price of good 1 changes. Let us denote the new price of good 1 by $\hat{p}_1$. Then we have two possible money metric utility functions, and correspondingly two possible measurements of the consumer's loss in Dollars:

$$E(p_1, p_2, U(q_1, q_2)) - E(p_1, p_2, U(\hat{q}_1, \hat{q}_2)),$$

and

$$E(\hat{p}_1, p_2, U(q_1, q_2)) - E(\hat{p}_1, p_2, U(\hat{q}_1, \hat{q}_2)).$$

These two measures have their own name. The first one is called the Equivalent Variation. The second one is called the Compensating Variation. The first one measures, at the old prices, which variation in the consumer's income would have been equivalent to the policy intervention that we are considering. The second one measures, at the new prices, what we would have to pay the consumer to bring her back to the old utility level, after we have introduced the policy intervention.

Let me offer another way of phrasing the intuitive interpretation of equivalent and compensation variation. Let us suppose that the policy intervention causes a utility loss for the consumer. The equivalent variation tells us how much money the consumer would be willing to give up from his income to avoid the policy intervention. The compensating variation tells us how much money we have to offer to the consumer to make him willing to agree to the policy intervention (if he has any say). These two numbers may be different. Which one should we use to measure the consumer's aversion to the policy? We shall return to this question later.

We can represent these two possible measures of the effect of a change on consumers in a graph. I show them in Figures 1 and 2. In both Figures, I represent an increase in the price of good 1 from p_1 to $\hat{p}_1$. The result is a rotation inwards of the budget line around its intersection point with the vertical axis. Thus, the original budget line is the straight line connecting the intersection point at Y/p_2 on the vertical axis to the intersection point Y/p_1 at the horizontal axis. The new budget line is the straight line connecting the intersection point at Y/p_2 on the vertical axis to the intersection point $Y/\hat{p}_1$ at the horizontal axis. The consumer's consumption originally was (q_1, q_2), and after the price increase it is: $(\hat{q}_1, \hat{q}_2)$.

By how much has the consumer's utility dropped, expressed in the money metric utility function, with prices being the original prices p_1 and p_2? We are first going to find the equivalent variation. We have to find the expenditure level $E(p_1, p_2, U(\hat{q}_1, \hat{q}_2))$. For this, we have to find the expenditure minimizing consumption bundle on the indifference curve through (q_1, q_2) and on the indifference curve through $(\hat{q}_1, \hat{q}_2)$. The former is obviously simply the income, Y. The latter is determined in Figure 57, where I have labeled the expenditure minimizing consumption bundle as $(\hat{e}_1, \hat{e}_2)$. Note that the corresponding "constant expenditure line" must have slope $-p_1/p_2$. So, it is parallel to the original budget line. Now, the expenditure level corresponding to that consumption bundle is $E(p_1, p_2, U(\hat{q}_1, \hat{q}_2))$. This expenditure level cannot straightforwardly be read off the diagram. We have to imagine it. But the drop of income from Y to the expenditure level corresponding to the expenditure minimizing consumption bundle pointed out in Figure 57 is the “Equivalent Variation.”

Figure 57: Equivalent Variation

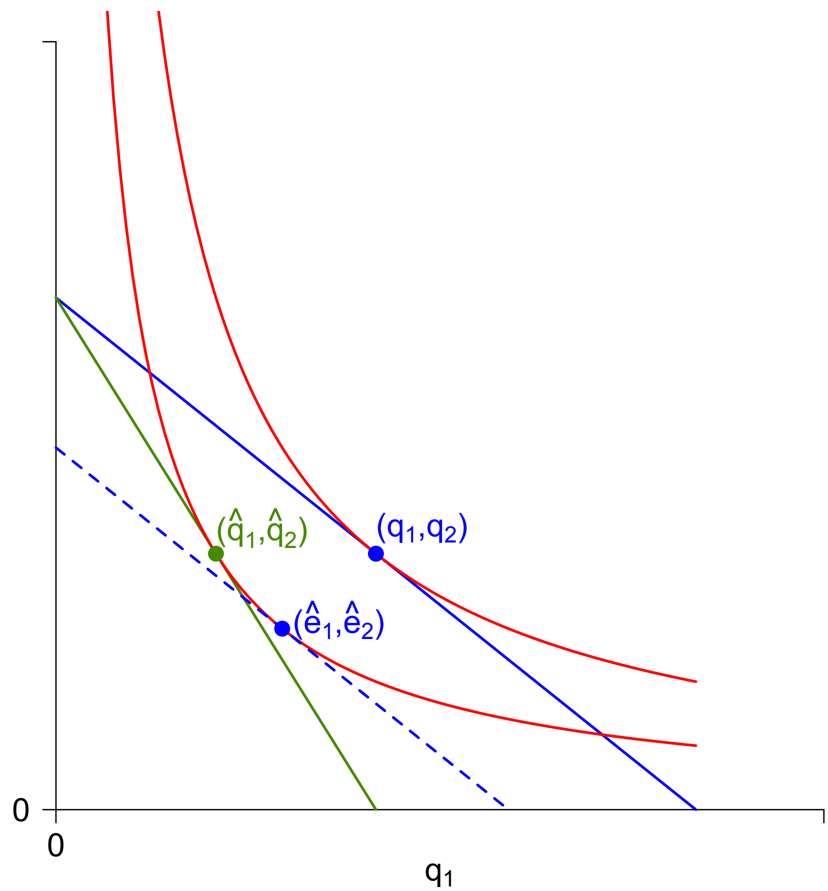


Figure 58 is very similar, but now we are representing the compensating variation. Now, it is clear that $E(\hat{p}_1, p_2, U(\hat{q}_1, \hat{q}_2))$ must equal Y. Thus, what we have to calculate is $E(\hat{p}_1, p_2, U(q_1, q_2))$. Similar to Figure 57, this is the expenditure level corresponding to the expenditure minimizing consumption bundle that I have labeled $(\bar{e}_1, \bar{e}_2)$ in Figure 58. The constant expenditure line is parallel to the new budget line, but tangential to the old indifference curve. The expenditure level corresponding to the expenditure minimizing consumption bundle in Figure 57, minus income Y, is the compensating variation.

We have obtained two monetary measures of the Dollar loss in utility caused

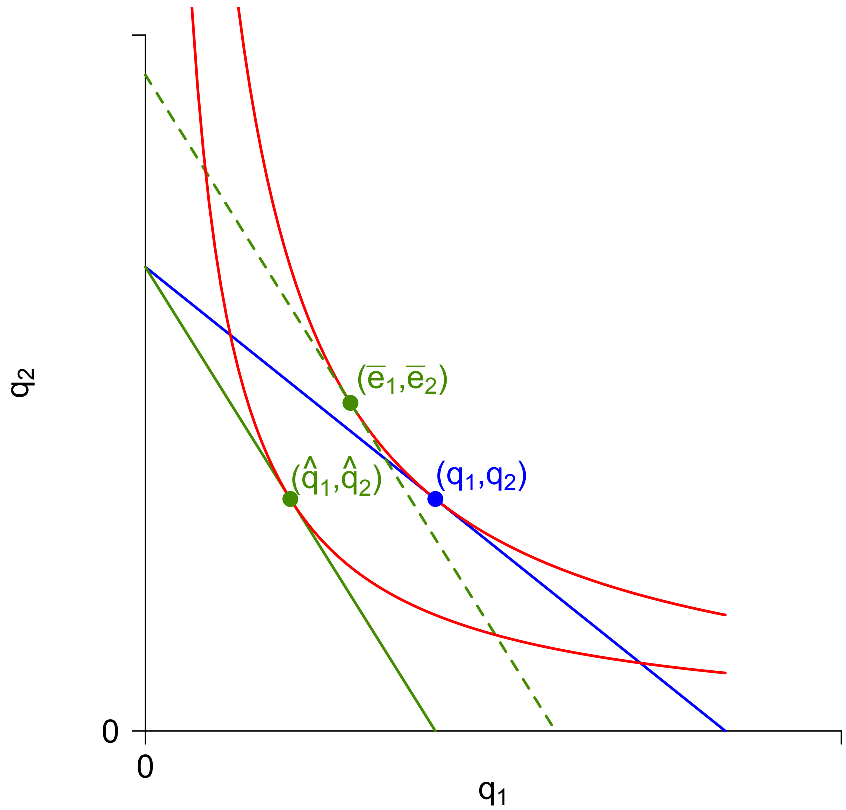


Figure 58: Compensating Variation

by an increase in the price of good 1. That is a little confusing. But the more natural measure to look at is the compensating variation. If the price of good 1 goes up, then the compensating variation tells us how much at least we have to give the consumer to compensate her for the increase in price. Maybe some other people gain from the increase in price. Then perhaps we can use the compensating variation to test whether a re-distribution of purchasing power would, in principle, be possible so that everyone is better off than before after the price increase.

I have shown the equivalent and the compensating variation for an increase in price. But the concepts are also defined for a decrease in price. Then we measure how much the consumer gains when the price goes down. Perhaps you can construct the graphs corresponding to the compensating and equivalent variation of a price decrease yourself, by appropriately modifying Figures 1 and 2?

A Numerical Example

Let us suppose that the utility function is the simplest Cobb Douglas utility function we can think of: $U(q_1, q_2) = q_1 q_2$. Income is: $Y = 20$. Prices are initially $p_1 = 1, p_2 = 1$, and the price of good 1 rises to $\hat{p}_1 = 2$. We are now familiar with how to calculate demand for Cobb Douglas utility functions. Before the price increase, the demand was $q_1 = 10, q_2 = 10$ with corresponding utility level $U(q_1, q_2) = 100$. After the price increase, the optimal demand is $\hat{q}_1 = 5, \hat{q}_2 = 10$ with corresponding utility level $U(\hat{q}_1, \hat{q}_2) = 50$. Thus, there is a utility loss of 50. But this number doesn't really mean anything. We want to calculate an equivalent Dollar amount.

To find the equivalent variation, we have to calculate how the new utility, $U = 50$ could be achieved in an expenditure minimizing way if the prices are the original prices: $p_1 = 1, p_2 = 1$. This is a standard expenditure minimization problem. Let me solve it explicitly. To minimize expenditure we have to set the MRS equal to the price ratio.

$$\frac{q_2}{q_1} = \frac{1}{1} \Leftrightarrow q_2 = q_1.$$

(Recall that we have calculated many times that the marginal rate of substitution for this utility function is q_2/q_1.) Our second condition is that we have to reach the target utility level:

$$q_1 q_2 = 50.$$

Plugging back in:

$$q_1 q_1 = 50 \Leftrightarrow q_1 = \sqrt{50}$$

and hence, because $q_1 = q_2$:

$$q_2 = \sqrt{50}$$

The consumption bundle that we have now determined is the consumption bundle that we have denoted in Figure 57 by $(\hat{e}_1, \hat{e}_2)$. Thus, in our example: $\hat{e}_1 = \hat{e}_2 = \sqrt{50}$. The corresponding expenditure level is simply $\hat{e}_1 + \hat{e}_2$, because both prices were originally 1:

$$E(p_1, p_2, U(\hat{q}_1, \hat{q}_2)) = 2\sqrt{50}.$$

The equivalent variation is:

$$E(p_1, p_2, U(q_1, q_2)) - E(p_1, p_2, U(\hat{q}_1, \hat{q}_2)) = 20 - 2\sqrt{50} \approx 5.86.$$

Thus, anticipating the price increase, the agent thinks: that is as if someone right now took 5.86 Dollars away from me.

Now we calculate the compensating variation. We have to calculate the expenditure minimizing way of achieving the old utility level 100 at the new prices, $\hat{p}_1 = 2, p_2 = 1$. The condition that marginal rate of substitution equals price ratio is now:

$$\frac{q_2}{q_1} = \frac{2}{1} \Leftrightarrow q_2 = 2q_1.$$

But we also need to reach the target utility level:

$$q_1 q_2 = 100.$$

Substituting:

$$q_1(2q_1) = 100 \Leftrightarrow q_1 = \sqrt{50},$$

and, substituting back:

$$q_2 = 2q + 2\sqrt{50}.$$

We have found the consumption bundle that we denoted in Figure 58 by $(\bar{e}_1, \bar{e}_2)$. Specifically: $\bar{e}_1 = \sqrt{50}$ and $\bar{e}_2 = 2\sqrt{50}$. The corresponding expenditure level is:

$$2\bar{e}_1 + 1\bar{e}_2 = 4\sqrt{50}.$$

And hence the compensating variation is:

$$E(\hat{p}_1, p_2, U(q_1, q_2)) - E(\hat{p}_1, p_2, U(\hat{q}_1, \hat{q}_2)) = 4\sqrt{50} - 20 \approx 8.28.$$

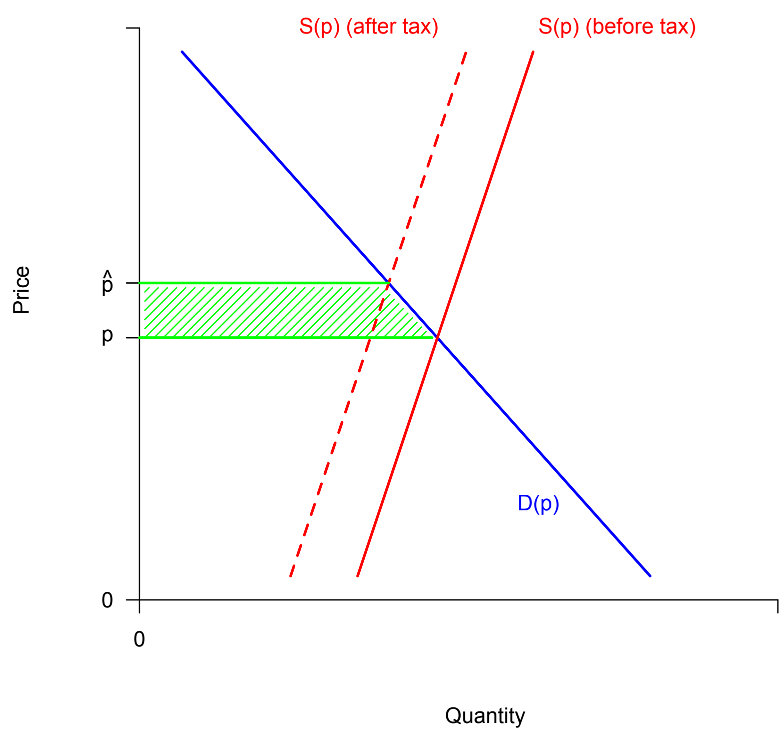
Thus, to bring the consumer back to his old utility level, after the price increase, we would have to pay her 8.28 Dollars.

The two numbers that we got, 5.86 Dollars and 8.28 Dollars, are quite far apart! In practice, though, they are typically much less far apart. It is a feature of our example that the difference is so big.

The Surpluses

Suppose we are in the situation discussed in the previous two sections: A tax is imposed on good 1 and the price of good 1 rises. We have already seen in Topic 14 how to study the effects of such a tax. In Figure 59 I reproduce that analysis, with some further information that we shall study in this section. I assume that the tax is paid by firms, so that the supply curve shifts. Consumers' price goes up from p to $\hat{p}$. Consumers are worse off. What is the corresponding loss in Dollars?

Figure 59: Loss in Consumer Surplus



I have shaded an area in green in Figure 59. It is the area under the demand curve, between the horizontal lines through prices p and $\hat{p}$. The size of this area is called the "loss in consumer surplus." And, as it happens, it is closely related to the equivalent and the compensating variation. In fact, if the expenditure on the good that we are looking at is only a small fraction of the consumer's income, then equivalent, compensating variation and loss in consumer surplus are approximately the same. I shall not prove these facts for you here. You have to take my word for them.

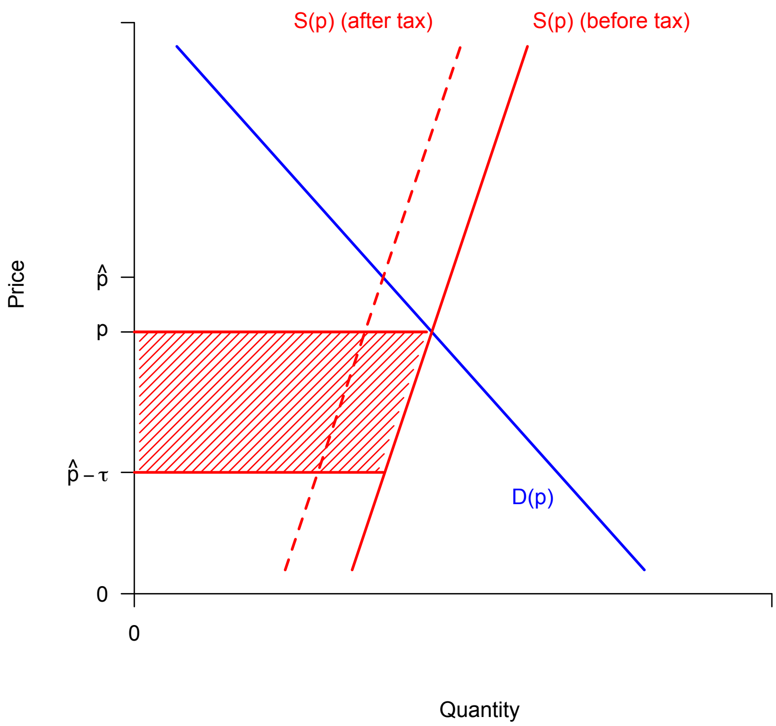
Why are these areas called "loss in consumer surplus"? The "consumer surplus" is defined as the area between the demand curve, and the horizontal line through the market price. In some sense, it is a measure of the contribution of this market to the consumer's money metric utility function. Note how vague the language is that I am using. But if you accept it for a moment, then you can see that the loss in consumer surplus is roughly the dollar measure that we are looking for.

The introduction of the tax also causes a change in firms' profits. I have indicated this in Figure 60. The profit loss is equal to the size of the red shaded area. The reason is that the area under the horizontal line through the price received by firms, and the supply curve, is the firm's profit. Why is it the profit? If you know a little calculus I can explain it: Recall that the supply curve is, in the simplest case, the marginal cost curve. Because the integral of the derivative is the function itself, by the Fundamental Theorem of Calculus, the integral of the marginal cost curve is the cost curve. Because the integral is the size of the area between the graph of a function and the x-axis, the size of the area under the marginal cost curve equals cost. The rectangle with one side being the horizontal line through the price, and the vertical line dropped from the price to the horizontal axis, is the revenue. The difference between these two areas, the profit, is the area between the horizontal line through the price, and the marginal cost = supply curve.

We are going to regard the loss in profits as another measure of the monetary effect of the tax. This is motivated by the fact that firms are owned by consumers. Thus, if they lose profit, the consumers lose income.

The tax also generates revenue for the government. The size of the revenue is the size of the rectangle with height τ, and the quantity that is produced in equilibrium. We shall regard this revenue as a "monetary benefit" of the tax.

Figure 60: Profit Loss



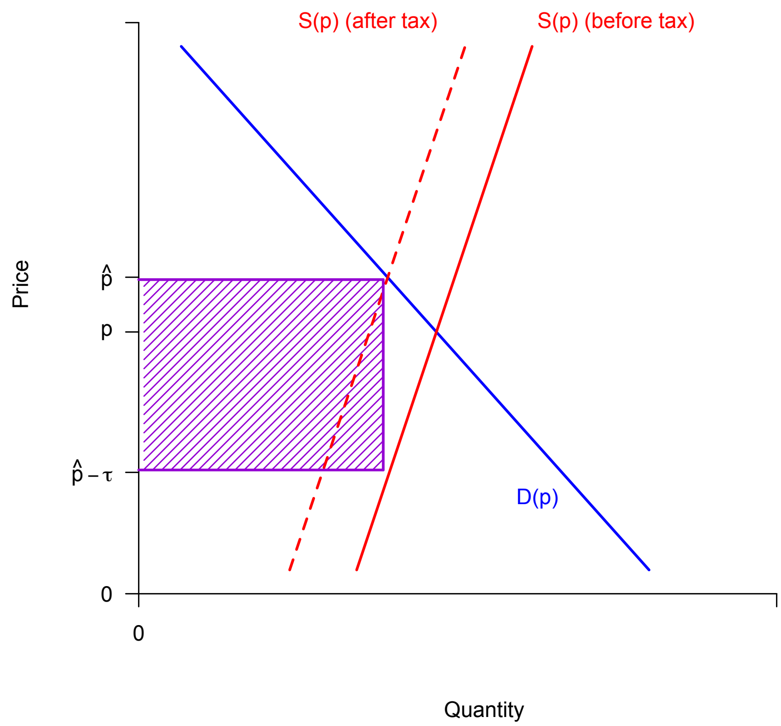
The government could give it back to consumers. I show the size of revenue in Figure 61.

Now we can make a rough calculation of the monetary effect of the tax: government revenue minus loss in consumer surplus minus loss in profits. If you go through Figures 3, 4 and 5, you can see that this difference is actually negative, and its size is the size of the triangle in Figure 62.

The size of this area is called the "deadweight loss." It indicates that, if we monetarize all the effects of a tax, we obtain that, de facto, it takes money away, and thus, everyone is made worse off.

Of course, this is a very rough calculation. The measure of equivalent or

Figure 61: Government Revenue



compensating variation that we use is very approximate. And equating government revenue with the monetary benefit of what the government does with the money is not well founded. Maybe the government spends the money poorly, or maybe it spends the money on some project that has huge benefits for consumers. This would have to change the calculation of deadweight loss.

I suspect that deadweight loss is a concept that you have encountered before. The purpose of this section was really to indicate a little bit how this concept is related to what we have done in this course. It is a concept that you should regard with caution. In the remainder of this course, we shall occasionally invoke this concept. But remember that there are many reasons to be skeptical.

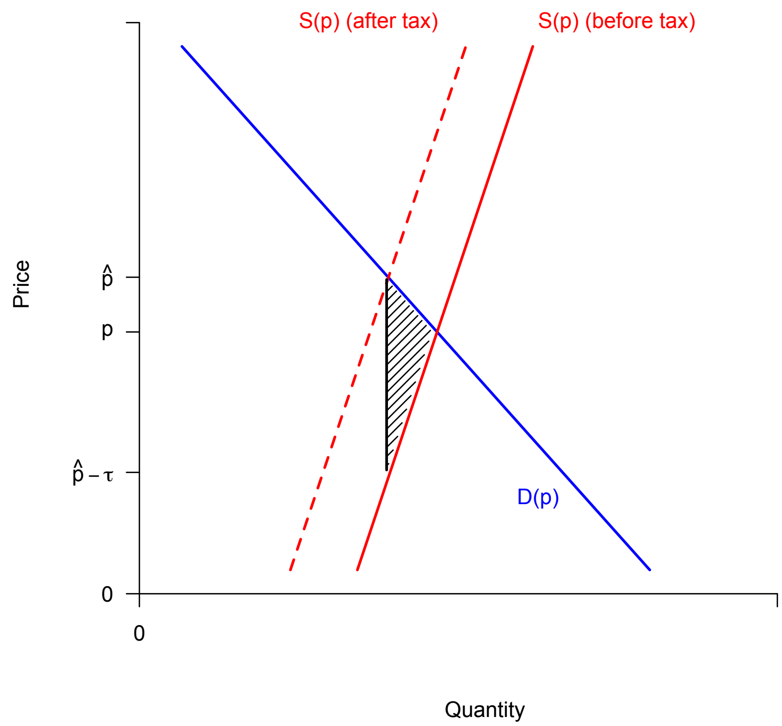


Figure 62: Deadweight Loss

Topic 20: Decisions Under Risk

Risky Choices as Probability Distributions over Outcomes

So far, we have only considered choices where the person who makes the choice, typically the consumer, knows with certainty what she is getting when she purchases quantities of goods. But often, when we purchase quantities of goods, we don't know what exactly we are getting. When you buy a previously used car, you may not know whether the car has been in an accident. When you choose from a menu in a restaurant, you often do not know what exactly the meal that you are ordering will taste like.

More generally, in every day life we have to make decisions under uncertainty all the time. For example, when we choose which courses to take towards our major at college, we don't yet know what the courses will really be like. Or when we decide whether to invest our money in stocks, and which stocks to buy, we don't exactly know how the price of those stocks will evolve over the next couple of years.

One way of thinking about decisions under risk that is popular among economists is that each choice that we may make leads to a probability distributions over "outcomes," and that choices differ only by which probability distributions over outcomes they lead to. For example, the four possible outcomes of the choice of a course at college might be that you learn very interesting things and the course is very difficult, that you learn not very interesting things but the course is still very difficult, that you learn very interesting things and the course is easy, or that you learn not so interesting things and the course is easy.

Of course, a detailed description of all the possible outcomes may have to include many more possible different outcomes than I listed for this example. But the distinction between choices and outcomes is useful.

The next, and also very important idea is that for each action we have a probability distribution over outcomes that describes for each outcome how likely that outcome is, if we choose the given action. For example, some course choice might lead to a probability of 1/2 that you learn very interesting things but it is very difficult, a probability of 1/12 that you learn not very interesting things but the course is very difficult, etc.

Where do these probabilities come from? Some choices that we can make have outcomes for which we can determine "objective" probabilities. For example, if you buy a lottery ticket that wins if one of 1 million numbers comes up with random probability, then you have a 1/1,000,000 chance of winning 1000 Dollars, and a 999,999/1,000,000 chance of winning nothing. But for many other choices, it is hard to say how likely it is that any particular outcome results. When you choose to invest in shares in Facebook, how likely is it that they go up by 10% over the next year? How should you arrive at an answer?

There is a lot of discussion of this question in the economics literature. We have to skip it here. The bottom line is that a rather persuasive argument supports that a rational decision maker should be able to come with some, admittedly subjective, probability assessment. In a nutshell, the decision maker should be able to compare a choice with a subjective risk to a choice with an objective risk, and then, by carefully questioning: "when I invest in Facebook shares, and get twice as much as I put in Facebook's shares doubles, and nothing otherwise" is this for me equivalent to putting up 1 Dollar, and getting back 2 Dollars if a tossed coin comes up "heads." and nothing otherwise?

In practice, of course, the questions that we ask ourselves need to be even more complicated, because there are many more possible outcomes of investing in Facebook than that either the money invested doubles, or that it loses all its values.

If we accept that every choice leads to some probability distribution over outcomes, it actually doesn't matter what the actions are called. We can simply identify each action with the distribution over outcomes that it leads to. For example, instead of naming the choice "take Intermediate Microeconomics," we may just name it: "with probability 1/2 have an interesting, but challenging class, and, also with probability 1/2, have a boring, and challenging class."

Let us introduce now some notation. We denote the set of possible outcomes of the choices that the decision maker has available by X. We shall assume that X is finite. This will make the mathematics simpler. The elements of X will be denoted by $x, y, \ldots, z$, etc. A probability distribution over X assigns to each element x of X a probability $p(x)$ that is between zero and one. Of course, the probabilities of all elements of X have to add up to one. We shall introduce a special notation for the set of all probability distributions over X. Following an old convention in economic theory we shall write for that set: ΔX, where you should read "Δ" as the capital letter "delta" in the Greek alphabet. Thus, any probability distribution p over X is an element of ΔX.

The small letter "delta" is written as "δ."

In summary, the perspective that we shall maintain in this Topic is that there is a decision maker who faces a choice among different alternatives, where each alternative leads to a different probability distribution over some set of outcomes X, that is, to some probability distribution p in ΔX.

I shall use some terminology that may seem strange to you: I shall call probability distributions over X also "lotteries." That is strange because in every day life the word "lottery" has a very specific meaning. Here, we use it in the sense of "lottery ticket," and, moreover, we use it even if the outcomes are not money payments, but simply elements of X. Hence, a course that is with probability 1/2 interesting but hard, and with probability 1/2 interesting but easy, will be called a "lottery." Indeed, every element p of ΔX is a "lottery."

It is useful to introduce the special case in which X has only three elements.

If this is the case, then we can represent the probability distributions over X in a simple, two-dimensional diagram. Suppose the elements of X are x, y, and z. Then, a probability distribution over X consists of three numbers: $p(x)$, $p(y)$, and $p(z)$. Because they have to add up to 1, once you know $p(x)$ and $p(y)$ you also know $p(z)$. It has to be: $p(z) = 1 - p(x) - p(y)$. Thus, to describe a lottery, we can focus on $p(x)$ and $p(y)$. Both numbers have to be between zero and one, and their sum must not be more than one.

We can show the possible probability distributions over x, y, and z using the triangle shown Figure 1. We display $p(x)$ on the horizontal axis, and $p(y)$ on the vertical axis. Both probabilities have to be between zero and one, and their sum must not be more than one. Therefore, all values in the shaded triangle are possible. Moreover, for any pair of values $p(x)$ and $p(y)$ in the shaded area, the probability of $p(z)$ is simply: $1 - p(x) - p(y)$.

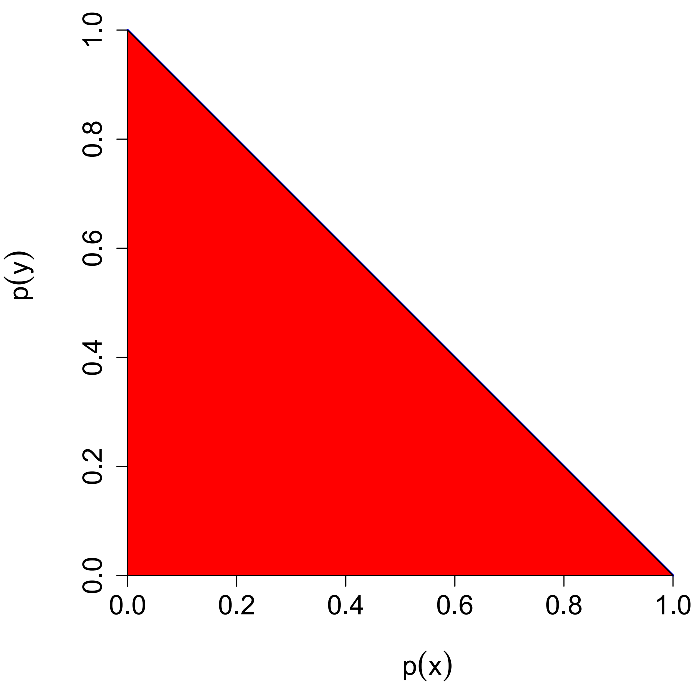


Figure 63: Probability Distributions over 3 outcomes: x, y and z.

Preferences Over Probability Distributions Over Outcomes

We want to think about rational choice when the outcomes of all, or some, of the choices are uncertain. in our new notation, this means that we want to think about a rational choice from the triangle depicted in Figure 63, or from a higher, or lower, dimensional triangle, if there are more, or less, than three choices.

If we adopt the ideas that we introduced in earlier topics, then we shall call a decision maker's choices rational if they look as if the decision maker had a transitive preference, and chose among all available actions the action that is best according to this preference. We could represent such a preference using indifference curves in the triangle shown in Figure 63. We show such indifference curves in Figure 64.

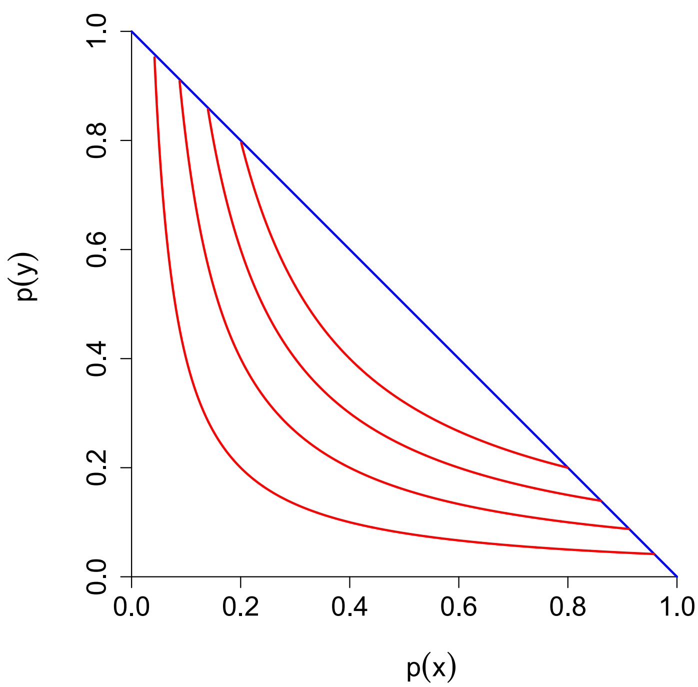


Figure 64: Indifference Curves Over Probability Distributions.

We might represent a preference over lotteries, i.e. probability distributions over X, by a utility function u. Thus, for any lottery p there would be a number $u(p)$, and the meaning of that number would be that a lottery p is preferred over a lottery q if $u(p) > u(q)$. Rational choice would mean that the decision maker has some utility function u, and chooses from the set of available lotteries p the one that maximizes u. This would be inline with the approach we have taken earlier in these notes. But, we shall modify this approach a little bit in the next section.

Expected Utility

Much of microeconomic theory has agreed to the point of view that a rational decision maker who makes decisions under risk doesn't just have some utility function u for lotteries, but that such a utility function must belong to a particular class of utility functions, namely those that can be written in the following form:

$$u(p) = p(x_1)v(x_1) + p(x_2)v(x_2) + \ldots + p(x_n)v(x_n),$$

where $x_1, x_2, \ldots, x_n$ is an enumeration of the elements of X, and where v is a function that assigns to every element of X some utility number.

When you look at the formula for $u(p)$, I hope you agree that we can interpret it as a "weighted average" of the numbers $v(x_1), v(x_2), \ldots v(x_n)$, where the weights are $p(x_1), p(x_2), \ldots, p(x_n)$. Such a weighted average, i.e. a weighted average where the weights are probabilities, is also called an "expected value," here the expected value of $v(x)$. We can think of $v(x)$ as the utility assigned to outcome x, if we knew in advance for sure that outcome x will occur. If there is uncertainty, then the utility of every outcome that is possible will be weighted using the probability that p assigns to that outcome. The decision maker chooses among all available lotteries the one that has the highest expected utility.

Pay attention to the distinction between "u" and "v." "$u(p)$" is the utility from a *lottery* p, and "$v(x)$" is the utility from an *outcome* x.

What we have now introduced is a very particular form for the utility function u, and, as yet without any explanation, we have claimed that every rational decision maker must have a utility function u that is of this form. This is very different from the case of, say, the theory of rational consumer choice that we considered earlier. There, we had no opinion about the utility function at all. Every utility function could be the utility function of a rational decision maker. When it comes to lotteries, by contrast, we shall argue that utility has to take this particular functional form.

Two points need to be argued, thus: what is the meaning of this particular function form, and why do some people believe that every rational decision maker has to have a utility function of this particular functional form.

Let's begin with the first question. It is useful to write down the equation for the indifference curves that are implied by the expected utility function. let us consider some fixed utility level, say $\bar{U}$ and let us consider which lotteries lead to this expected utility level if the utility from lotteries is of the expected utility form. We focus for simplicity on the case of only three outcomes. The equation of an indifference curve is:

$$\begin{aligned} u(p) &= \bar{U} \Leftrightarrow \\ p(x_1)v(x_1)+p(x_2)v(x_2)+p(x_3)v(x_3), &= \bar{U}. \end{aligned}$$

Let us replace $p(x_3)$ by $1-p(x_1)-p(x_2)$:

$$p(x_1)v(x_1)+p(x_2)v(x_2)+(1-p(x_1)-p(x_2))v(x_3) = \bar{U}.$$

We shall now solve for $p(x_2)$ in terms of $p(x_1)$ to obtain an equation for he indifference curve in the diagram in which $p(x_2)$ is on the vertical axis, and $p(x_1)$ is on the horizontal axis. We get:

$$\begin{aligned} p(x_1)v(x_1)+p(x_2)v(x_2)+(1-p(x_1)-p(x_2))v(x_3) &= \bar{U} \Leftrightarrow \\ p(x_1)(v(x_1)-v(x_3))+p(x_2)(v(x_2)-v(x_3)) &= \bar{U}-v(x_3). \Leftrightarrow \\ p(x_2) = \frac{\bar{U}-v(x_3)}{v(x_2)-v(x_3)} - \frac{v(x_1)-v(x_3)}{v(x_2)-v(x_3)}p(x_1) \end{aligned}$$

The main fact to notice here is that $p(x_2)$ is a linear function of $p(x_1)$. That is, for given v and $\bar{U}$, the equation shows that $p(x_2)$ is a linear function of $p(x_1)$ with intercept $\frac{\bar{U}-v(x_3)}{v(x_2)-v(x_3)}$ and with slope $-\frac{v(x_1)-v(x_3)}{v(x_2)-v(x_3)}$. These are linear indifference curves, since the slope does not depend on $p(x_1)$. Moreover, all indifference curves have the same slope, and are therefore parallel. They differ only by the intercept with the vertical axis. In Figure 65 I have drawn indifference curves into the triangle. Of course, whether these indifference curves are upward sloping or downward sloping, and at which angle, depends on the decision maker's taste, as expressed through the three numbers $v(x_1)$, $v(x_2)$, and $v(x_3)$. These are in a sense "parameters" of the expected utility function.

When there are four alternatives, the probability space is a tetrahedron, indifference curves are planes. In higher dimensions the geometrical objects get

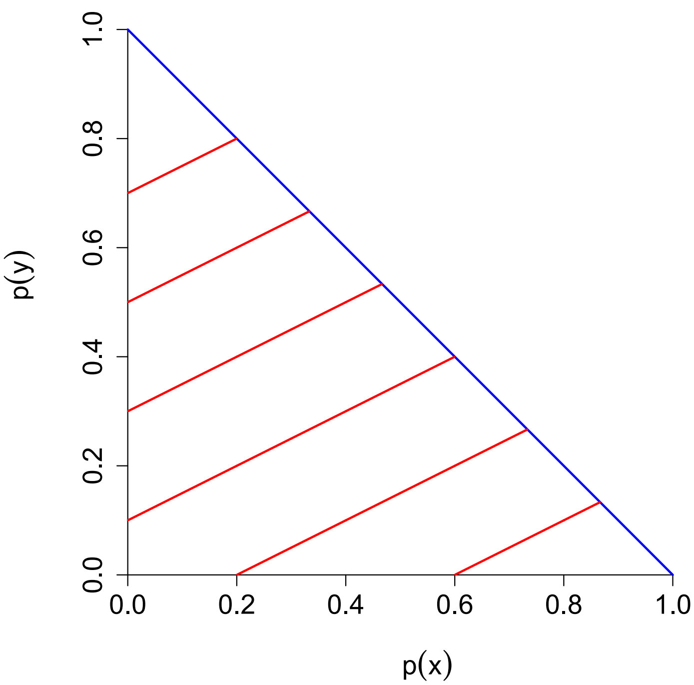


Figure 65: Linear and Parallel Indifference Curves Over Probability Distributions.

more complicated, but the idea is the same.

The claim of expected utility theory is thus that a rational decision maker should not only have transitive preferences, but also that a rational decision maker must have indifference curves of a certain form, or, as one can also put it, must have preferences that depend only on a small number of parameters. Note that this is quite different from the case of the preferences of a consumer over consumption bundles. In that case, we have really not made any further claim what a rational consumer's indifference curves should like. Why do we make this stronger claim when we consider choices among probability distributions over outcomes? We shall explain this in the next section.

Axioms for Expected Utility

The theory of expected utility maximization is due to John von Neumann and Oskar Morgenstern, the former a 20th century mathematician, and the latter a 20th century economist. John von Neumann and Oskar Morgenstern are most well-known among economists for their book, "Theory of Games and Economic Behavior," published in 1944. This book pioneered the use of game theory in economics. But, in an appendix, it discussed expected utility theory. We shall come to game theory later. For the moment, we shall just focus on expected utility theory.

Because the theory of expected utility theory is due to von Neumann and Morgenstern the function v, defined on outcomes, that we introduced in the previous section is often also called the "von Neumann Morgenstern" utility function. Because that is a lot of words, it is sometimes also abbreviated as the vNM-utility function.

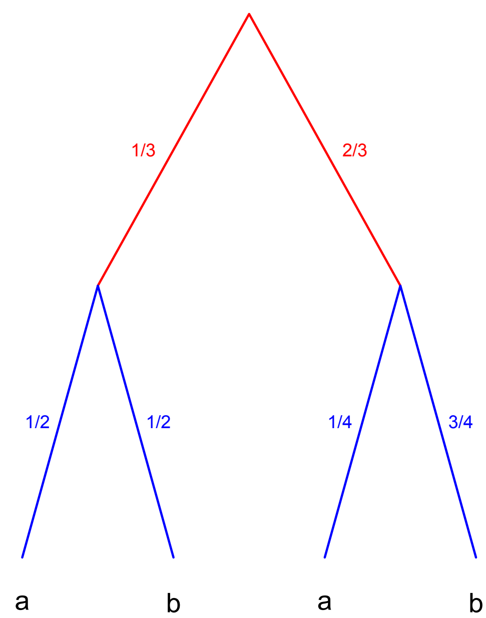
Why did vNM think that rational choice over lotteries means expected utility maximization? They argued that preferences over lotteries should satisfy an axiom that they called the Independence Axiom. I shall first write it down, but you may not understand it at first sight. I shall explain it.

Independence Axiom: If a rational decision maker prefers lottery p over lottery q, and r is some other lottery, then the decision maker should also prefer lottery $\alpha p + (1-\alpha)r$ over lottery $\alpha q + (1-\alpha)r$ for every α between 0 and 1.

What I need to explain here is what I mean by lotteries of the form "$\alpha p + (1-\alpha)r$ when α is between zero and 1. First, I shall give an intuitive explanation. Then, I shall give a mathematical explanation. For the intuitive explanation, I am going to interpret an expression like "$\alpha p + (1-\alpha)r$" as a "two stage lottery." To illustrate what I mean by a "two stage lottery," let me suppose for the moment there are two outcomes only, a and b. We set $p = (0.5, 0.5)$, that is, each outcome occurs with probability 1/2, we set $r = (0.25, 0.75)$, that is, outcome a occurs with probability 1/4 and outcome b occurs with probability 3/4, and we set $\alpha = 1/3$. Then $\alpha p + (1-\alpha)r$ stands for a lottery that can be visualized as shown in Figure 66.

Figure 66 shows a sequence of two random events that occur one after the other. First, we go left or right, left with probability 1/3, and right with probability 2/3. And then a or b is chosen. If, in the first stage, we went left, then a and b occur with probability 0.5 each, that is, lottery p. If, in the first

Figure 66: Two stage lotteries.



stage, we went right, then a and b occur with probability $1/4$, respectively $3/4$, i.e. lottery r. Thus: $\alpha p + (1-\alpha)r$ is here interpreted as a two stage lottery: first we decide between p and r, chosen with probability α and $1-\alpha$, and then, depending on the outcome in the first stage, p or r occurs.

Now let's consider the independence axiom. Its meaning is visualized in Figure 67. The Independence Axiom refers to two two-stage lotteries. In Figure 67 I am not showing the second stage lotteries explicitly. They are just named as "p, q, r." Now compare the two two-stage lotteries in Figure 67. They are almost identical, except that, after the left-hand branch that is chosen with probability α, in one of them p occurs, and in another one q occurs. The right hand branch is identical in both two-stage lotteries.

The independence axiom says that the decision maker's preferences over these two two-stage lotteries should only depend on their preference between p and q, because that is the only aspect in which the two two-stage lotteries differ. They are identical in all other aspects. That is, the comparison between

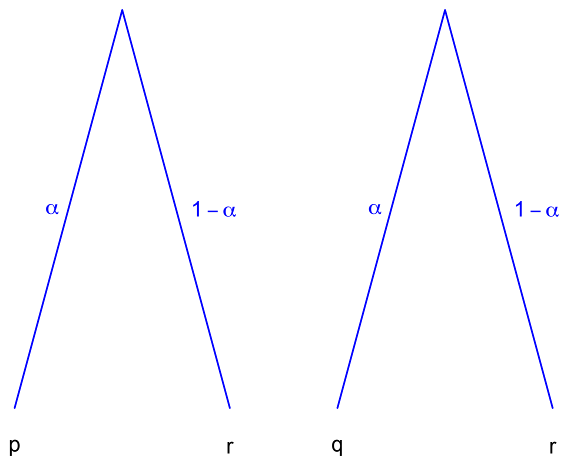


Figure 67: vNM's Independence Axiom.

the two two-stage lotteries should be independent of the first stage choice, as long as it has in both two-stage lotteries the same probability, and it should not depend on what follows the right hand branch, as long as that is also the same in both two-stage lotteries. Thus, the decision maker's preferences among the two-stage lotteries should only depend on what changes as a function of her choices, it should not depend on things she cannot change.

If described in this way, the Independence Axiom sounds to me eminently reasonable. Nonetheless, if we observe people's choices in practice, it is violated often. That is troubling, because we want our theory to be useful in describing what people do in practice. Some researchers think that, to a first approximation, the deviations from the Independence Axiom that we observe in practice don't matter. In any case, the expected utility theory is the benchmark theory about choices under uncertainty, and therefore, we need to explain it here first, and won't get to any alternative theories, because this is an introductory course.

Now let me briefly come to the mathematical explanation of the Independence axiom. Consider again Figure 66. In one in three cases we choose the left hand branch, and then, in one in two cases, we choose *a*. This means that in one in six cases, *a* occurs. But *a* can also occur on the right hand side. In 2 out of 3 cases the right hand branch is chosen, and in one in four cases, then a follows. This means that we have another 2 out of 12 cases that a is chosen. Altogether, the probability that a is chosen is therefore: $1/6 + 2/12$, which equals $1/3$. With the remaining probability, that is, $2/3$, b occurs. Thus, really the two-stage lottery in Figure 66 boils down to a one stage lottery where a is chosen with probability $1/3$, and b is chosen with lottery $2/3$. We can write this as follows:

$$\frac{1}{3}\left(\frac{1}{2},\frac{1}{2}\right)+\frac{2}{3}\left(\frac{1}{4},\frac{3}{4}\right)=\left(\frac{1}{3},\frac{2}{3}\right).$$

More generally:

$$\alpha\,(p(a),p(b))+(1-\alpha)(q(a),q(b))=(\alpha p(a)+(1-\alpha)q(a),\alpha p(b)+(1-\alpha)q(b)).$$

The right hand side of this equation is just a one stage lottery. It is this lottery when we write $\alpha p + (1-\alpha)q$. The two stage lotteries that I used to explain the independence axiom are really just a metaphor. They make the independence axiom plausible. The mathematical content refers to lotteries that are constructed as in the equation above. This equation generalizes in a simple, hopefully obvious way to the case of many outcomes.

Von Neumann and Morgenstern's main insight is that a transitive preference satisfies the independence axiom, and some other smaller conditions, if and only if it can be represented by an expected utility function, and therefore has linear and parallel indifference curves. Thus, this very firm restriction on the shape of indifference curves is justified by the independence axiom. In consumer theory, say, as we discussed it earlier, there is nothing similar to the independence axiom that would justify restriction to particular forms of indifference curves.

vNM's result is a mathematical theorem with a proof. We won't formalize things to this extent, and so I shall not offer a proof. But I do want to show you that expected utility preferences do satisfy the independence axiom. Let's consider the case of three outcomes,a, b, and c. Then expected utility is:

$$p(a)v(a)+p(b)v(b)+p(c)v(c).$$

Therefore, the expected utility from $\alpha p + (1-\alpha)r$ is:

$$(\alpha p(a)+(1-\alpha)r(a))v(a)+(\alpha p(b)+(1-\alpha)r(b))v(b)+(\alpha p(c)+(1-\alpha)r(c))v(c).$$

The expected utility from $\alpha q + (1-\alpha)r$ is:

$$(\alpha q(a)+(1-\alpha)r(a))v(a)+(\alpha q(b)+(1-\alpha)r(b))v(b)+(\alpha q(c)+(1-\alpha)r(c))v(c).$$

Which of these two lotteries does the decision maker prefer? We can subtract the last expression from the one before, and we obtain as the difference between these:

$$\alpha p(a)v(a)+\alpha p(b)v(b)+\alpha p(c)v(c)-(\alpha q(a)v(a)+\alpha q(b)v(b)+\alpha q(c)v(c)).$$

The decision maker prefers $\alpha p + (1-\alpha)r$ to $\alpha q + (1-\alpha)r$ if this difference is positive, and vice versa if it is negative. But the sign of this difference does not change if we divide the difference by α. Then, the difference is positive if:

$$p(a)v(a) + p(b)v(b) + p(c)v(c) > q(a)v(a) + q(b)v(b) + q(c)v(c).$$

But this just means that the decision maker prefers p to q. The independence axiom says that just this should matter when the decision maker chooses between $\alpha p + (1-\alpha)r$ and $\alpha q + (1-\alpha)r$, and our calculation shows that this is indeed the case.

Uniqueness of vNM Utility

Suppose a decision maker with vNM utility function v has calculated his utility from two lotteries p and q, and discovers that the expected utility from p is larger than the expected utility from q. Now, suppose that he revises his opinion on what his utility function v is, and he adds the constant 5 to all values of v. How will p and q compare now? Well, if expected utility from p was initially:

$$u(p) = p(x_1)v(x_1) + p(x_2)v(x_2) + \ldots + p(x_n)v(x_n),$$

then, after adding 5 to all vNM utilities, it is:

$$p(x_1)(v(x_1)+5) + p(x_2)(v(x_2)+5) + \ldots + p(x_n)(v(x_n)+5).$$

Rearranging terms, this is equal to:

$$p(x_1)v(x_1) + p(x_2)v(x_2) + \ldots + p(x_n)v(x_n) + (p(x_1) + p(x_2) + \ldots + p(x_n))5.$$

And this is equal to:

$$p(x_1)v(x_1) + p(x_2)v(x_2) + \ldots + p(x_n)v(x_n) + 5,$$

because the probabilities in the bracket in front of 5 add up to 1. This is the same as:

$$u(p) + 5.$$

In other words, we have reached the, perhaps obvious, conclusion that if we add 5 to the vNM utility function we just add 5 to the expected utility. Thus, if initially the decision maker preferred p over q:

$$u(p) > u(q),$$

then it will also be true afterwards:

$$u(p) + 5 > u(q) + 5.$$

Adding a constant to the vNM utility function does not change how lotteries are compared, and therefore if we add a constant, we still have a representation of exactly the same preferences over lotteries.

Similarly, when we multiply the vNM utility function by a constant c that is positive, then all expected utilities get multiplied by that constant, and the comparison among different lotteries does not change. Thus, if a decision maker's preferences can be represented using the vNM utility function $v(x)$, they can also be represented by the utility function $cv(x)$, and, adding a constant a, $cv(x) + a$, we still have the same preferences. In the jargon of microeconomic theory one says that vNM utility functions are unique up to increasing ($c > 0$) linear transformations.

How about non-linear transformations? For example, suppose we square the vNM utility function. Will the preferences remain the same? The answer is "no." Here is an example: a decision maker with utility function $v(a) = 1$, $v(b) = 4$, and $v(c) = 6$ prefers to get b for sure, i.e. $p(b) = 1$, over a and c both with probability 0.5, i.e. $q(a) = q(c) = 0.5$, which only gives expected utility: $0.5 \cdot 1 + 0.5 \cdot 6 = 3.5$, i.e. less than the utility 4 that the decision maker gets from b. But now let's square all utilities. Then b gives utility $4^2 = 16$, but the lottery gives utility $0.5 \cdot 1^2 + 0.5 \cdot 6^2 = 18.5$, and thus now the lottery is preferred.

Why do non-linear transformations, even if they are monotonically increasing, change the preferences? In the context of the theory of utility maximizing consumption choices we said earlier that any monotonically increasing transformation, not just linear such transformations, leave the preferences unchanged. The difference is that here, in the context of choice among lotteries, we take expected values of the utility function. We did not do that when considering consumption choices. When we take expected values, only linear increasing transformations of the utility function leave preferences unchanged.

vNM Utility Functions for Money

Now let us consider a special case, namely the case that the outcomes, that is, the elements of the set X, are just Dollar amounts. Dollar amounts are real numbers, and therefore, we can set:

$$X = \mathbb{R}.$$

So far, we have assumed that X is finite. Now we have a case where X is infinite. But this is just a technicality. The theory of expected utility maximization can be extended to this case. It is just harder to write down.

Expected utility maximization now means that whenever the decision maker chooses among lotteries that have Dollar amounts as outcomes, she calculates the anticipated average value of a vNM utility function defined for Dollars that assigns to every Dollar amount x a utility amount $v(x)$. It seems plausible that v is a monotonically increasing function.

It is very important that our theory says that a rational decision maker maximizes the expected value of the vNM utility function v, but that it does not say that a rational decision maker necessarily maximizes the expected Dollar amount. Beginning students often get confused about this point. Of course, a possible special case is: $v(x) = x$, and then maximizing expected vNM utility and expected Dollars is the same; but that is really just a special case.

The fact that an expected utility maximizer maker might not maximize expected money is interesting in its own right. Suppose, for example, that a decision maker can choose some investment that gives him 5 Dollars with probability 0.5, and 10 Dollars with probability 0.5, and another one that has no risk at all. It just yields 7.5 Dollars. If these investments were evaluated on the basis of their expected Dollar value, they would just be the same. Both have expected Dollar value 7.5 (admittedly, for the second investment, it is a little artificial to call this an "expected value"). But an expected utility maximizer might prefer either investment over the other. Indeed, if an expected utility maximizer prefers the risky investment over the safe investment, we shall say that this decision maker is "risk loving," whereas, if she prefers the safe investment over the risky, we shall say that the decision maker is "risk averse."

The precise definitions of risk aversion and risk loving are as follows:

An expected utility maximizer is risk-loving if, whenever he chooses between a risky investment that yields x Dollars with probability p and y Dollars with probability $1 - p$, and a safe investment, that yields for sure $px + (1 - p)y$ Dollars, he prefers the risky investment. He is risk-averse if he prefers the safe investment, and he is risk-neutral if he is indifferent.

An important result says that a decision maker is risk-loving if his vNM utility function v is convex, and is risk-averse if his vNM utility function v is concave. So, for example, if a decision maker has vNM utility function $v(x) = x^2$, she is risk-loving. How can we see that v is convex? Remember that a function is convex if its second derivative is positive. Here, the first and second derivatives are:

$$v'(x) = 2x, v''(x) = 2.$$

Because 2 is positive, this function is convex. If, on the other hand, the vNM utility function is, for example, $v(x) = \sqrt{x}$, then the decision maker is risk-averse, because this is a concave function. We check this by checking that the second derivative is negative:

$$v'(x) = 0.5x^{-0.5}, v''(x) = -0.25x^{-1.5}.$$

A decision maker is risk neutral if his vNM function is linear. For example, if $v(x) = 2x + 3$, then the decision maker is risk neutral.

An Application

Let us now consider an application of the theory of decision making under risk. We shall study a risk averse decision maker whose vNM utility function of money is: $u(x) = \sqrt{x}$. We showed in the last section that this is a concave function, and therefore this decision maker is risk averse.

Let us consider this decision maker's optimal purchase of home insurance. Specifically, suppose that the decision maker faces a 10% risk that his house burns down. This is the loss against which he wants to insure himself. For simplicity, we shall assume that the only aspect of the burning of his house to the decision maker is that he loses money. Specifically, if the house does not burn down, his total wealth, including the value of the house, is \$100. But if the house burns down, he loses \$80, and his total wealth is only \$20.

Now imagine that there is an insurance company that offers this contract:

- If your house burns down, you get $\$x$.
- If your house does not burn down, you get \$0.

You can choose the value of x, that is, the level of your insurance coverage. The price of the insurance is $\$0.12x$. The question that we want to study is: Which value of x, that is, how much insurance coverage, should the decision maker optimally buy?

We calculate the decision maker's expected utility if he buys coverage of $\$x$. If his house does not burn down, he ends up with a total wealth of $100 - 0.12x$ Dollars, that is, the original wealth, minus the cost of insurance. If he is unlucky, and the house does burn down, he has $20 + x - 0.12x = 20 + 0.88x$ Dollars. His expected utility, using the square-root vNM utility function, is therefore:

$$u(x) = 0.9\sqrt{100 - 0.12x} + 0.1\sqrt{20 + 0.88.x}.$$

Let's find the value of x for which this expected utility is maximized. We set the first derivative equal to zero:

$$u'(x) = 0.9\frac{1}{2\sqrt{100 - 0.12x}}(-0.12) + 0.1\frac{1}{2\sqrt{20 + 0.88x}}0.88 = 0.$$

Solving this equation requires some unpleasant algebra. But here is how to proceed. First, we re-arrange terms to get:

$$\frac{0.9 \cdot 0.12}{2\sqrt{100 - 0.2x}} = \frac{0.1 \cdot 0.88}{2\sqrt{20 + 0.8x}}.$$

Next, we square both sides:

$$\frac{(0.9 \cdot 0.12)^2}{4(100 - 0.2x)} = \frac{(0.1 \cdot 0.88)^2}{4(20 + 0.8x)}.$$

And next we multiply both sides by the product of the two denominators:

$$(0.9 \cdot 0.12)^2(20 + 0.8x) = (0.1 \cdot 0.88)^2(100 - 0.2x).$$

This is a linear equation in x, which we can solve easily:

$$x = \frac{(0.1 \cdot 0.88)^2 \cdot 100 - (0.9 \cdot 0.12)^2 \cdot 20}{(0.9 \cdot 0.12)^2 \cdot 0.8 + (0.1 \cdot 0.88)^2 \cdot 0.2}.$$

The solution is approximately:

$$x \approx 49.74.$$

We still have to check the sign of second derivative of the expected utility, to make sure that we really do have a maximum. For simplicity I shall skip this step here. It turns out that we have indeed found a maximum.

The conclusion is: The decision maker's optimal choice of x, that is, home insurance coverage, is: \$49.74. This does not cover all his losses when his house burns down, which are \$80, but it covers some of the loss. The decision maker pays for this a premium of $\$0.12x = \$5.97$.

Why does the decision maker only buy partial insurance, not full insurance, i.e. why is x strictly less than \$80? The answer is that the price of coverage, $0.12x$, is too high. The expected money payment to the decision maker is only $0.1x$. Because he is risk averse, he buys insurance even if the price is higher than this, but he will only buy partial, not full insurance. One can show that, had the premium been $0.1x$, then the decision maker would have bought full insurance.

Topic 21: Game Theory

Although economics is sometimes portrayed as the field in which we explain everything as an equilibrium of supply and demand, a different methodology has taken over large parts of economics research since the 1970s: the methodology of game theory. The introduction of game theory into economics was first proposed by the mathematician John von Neumann and the economist Oskar Morgenstern in their seminal book published in 1944 "Theory of Games and Economic Behavior." Game theory's breakthrough in economics came, however, much later, starting in the 1970s. Game theory became influential among economists only after some developments that followed the publication of von Neumann and Morgenstern's book: Nash's invention of the equilibrium notion named after him, Harsanyi's model of games with incomplete information, and Selten's contribution to the theory of dynamic games.

Today, game theory is arguably the dominant methodology used by researchers in microeconomic theory. The name "game theory" for a while invited the misperception that this was not really something serious, and that the primary interest of researchers in game theory was in games such as chess. When one describes more accurately what game theory is concerned with, it sounds a little less entertaining and exotic: game theory is about rational choice of strategies when each agent's utility depends not only on their own choices, but also on the choices of other agents. That is, of course, true in chess, where my own strategies, although typically very sophisticated, somehow always come undone by the choices of my opponents. But it is also true in many situations of every day life, including the part of life that we focus on in economics.

Whenever there is an externality, i.e. an effect of some agent's choices on another agent's utility, then there is, by definition, a strategic interaction of the sort game theory studies. If a firm competes with another firm, each firm's profits will depend on the choices of the other firm. When two employees work in a firm, what they achieve, and perhaps how they will be rewarded, will depend on both employees' effort choices. When countries choose the tariffs that they wish to impose on imports from other countries, then the effect of these tariffs depends on the tariffs chosen by other, competing importers. Potential applications for game theory are everywhere in economics.

In this section, I shall introduce the most basic game theoretic ideas. I first describe them briefly at the general, abstract level. But it is best to then quickly move on to examples. Examples are the best way to understand game theory. Therefore, the rest of this section consists of examples. For more on game theory, I have to refer you to textbooks on game theory. There is an abundance of excellent introductory books on game theory. In this topic, we shall cover only a tiny percentage of what is covered in those books.

Game Theory

The models that game theory considers are called games. To describe a game in game theory, one must specify:

1. the players;

   This is just another name for the agents involved in the interaction. Depending on the application, they can be firms, consumers, countries, employees, etc. In general, we will write for the players: $i = 1, 2, \ldots, N$.

2. the strategy sets;

   Strategies are the choices available to players. For example, if the players are firms, then maybe their strategies are the choice of product they want to produce, where they locate their production facilities, which prices they charge, etc. In general, we denote by S_i the set of available choices of player i, where i could be any of $i = 1, 2, \ldots., N$, and we denote specific strategies by the symbol s_i.

3. the utility functions;

   The defining feature of a game is that every player's utility depends on *all* players' choices. For example, a firm's profit will depend not only on her own choice of product, but on *all* (relevant) firms' choices of products. In general, we specify for each player i a function that indicates player i's utility as a function of *all* players' choices: $u_i(s_1, s_2, \ldots, s_N)$. The fact that we are dealing with a game is reflected by the fact that all the other players' strategy choices are in the brackets. Note that the utility functions are just representations of ordinal preferences, as they have been throughout this course.

Once we have specified a game that we hope is a good model of some interesting economic situation, we next try to make a prediction about players' behavior. The most common approach is to find the Nash equilibrium, or, if there are several, the Nash equilibria, of the game. Here is the definition:

Many people say "Nash equilibriums" when they refer to the plural of "Nash equilibrium." But, equilibrium is a word with Latin origin, and the Romans typically indicated the plural of a word that ends in "um" by replacing the "um" by "a." By saying "equilibria," you can show off the quality of your education.

> Strategies $(s_1^*, s_2^*, \ldots, s_N^*)$ form a Nash equilibrium if for every player i the strategy s_i^* maximizes player i's utility u_i provided that player i correctly anticipates all other players' choices s_j^* $(j \neq i)$ and regards them as given and fixed, independent of her own choice.

Another common way of describing equilibria is as follows. A Nash equilibrium is a list of strategies, one for each player, such that no player can gain by changing their own strategy if the strategies of the other players remain unchanged. Or, if you are more mathematically inclined, I can write the mathematical definition for you: strategies $(s_1^*, s_2^*, \ldots, s_N^*)$ form a Nash equilibrium if for every player i:

$$u_i(s_1^*, \ldots, s_{i-1}^*, s_i^*, s_{i+1}^*, \ldots, s_N^*) \geq u_i(s_1^*, \ldots, s_{i-1}^*, s_i, s_{i+1}^*, \ldots, s_N^*)$$

for all $s_i \in S_i$.

Here, we assume that player i is somewhere in the middle of the sequence $1, 2, \ldots, N$. If player i is player 1, then the definition would have to be modified like this: $u_1(s_1^*, \ldots, s_N^*) \geq u_i(s_1, \ldots, s_N^*)$ for all $s_1 \in S_1$, and it would have to be modified similarly if player i were player N.

As this is the concept most commonly used in game theory, it is important to understand *why* we use it. The key idea that underlies the concept of Nash equilibrium is that for each player, somehow, the other players' behavior has become predictable. This may have been through experience from repeated interactions, or through deep strategic analysis. The solution concept itself can be used without being specific about which of these reasons motivates us to use it.

Whether Nash equilibrium is a good concept for predicting players' behavior clearly depends on the game. For example, as much as I have tried to play chess, other players' behavior in chess has somehow never become predictable to me. Whenever I play a better player, that is, almost always, I am surprised by that player's choices. This is a sign that play is not in a Nash equilibrium because I have not anticipated correctly the other player's choices.

The situation is different, perhaps, when firms have to choose their prices. It may have become predictable for each firm which price the competitors normally choose, and also when they have sales, by how much they cut their prices during sales, etc.

One can speculate for a long time about circumstances in which Nash equilibrium seems a realistic concept, and circumstances in which it doesn't. As in other contexts in these notes, I suggest you do not form a firm opinion on this without first looking at data. Ultimately, the question is whether interpreting real world choices as Nash equilibrium choices yields interesting and useful insights into people's behavior, and we can decide this only with data.

Example 1: Group Homework

Consider two students: $i = 1, 2$. They are working together on a group homework project. We want to study how much effort each student puts into the homework if the students make rational decisions. Let us denote the time that student i spends on the homework by t_i, and let us assume that that the time spent on homework cannot be more than one hour: $0 \leq t_i \leq 1$.

The more time the students spend on the homework, the better their grades become. Let's combine the improvement in the quality, and the importance that each of the two students attaches to a better grade in just one expression that we shall call the "benefit from joint homework:" $\ln(t_1 + t_2)$. We assume that the benefit from joint homework for student 1 is given by this expression, and also for student 2. Here, we use the logarithm really just for convenience: it is an increasing and concave function, and that is all that matters for us.

Each student also derives benefit from leisure, that is, from the time spent not working. If student i works for t_i hours, then what remains as leisure is $1 - t_i$. We shall assume that the benefit from this leisure for student 1 is: $\ln(1 - t_1)$. To make things more interesting, let's assume that the benefit from leisure to student 2 is smaller than for student 1: $\frac{1}{2}\ln(1 - t_2)$, in other words: student 2's opportunity cost of working are lower than student 1's. Putting things together, we have that student 1's total utility is:

$$u_1(t_1, t_2) = \ln(t_1 + t_2) + \ln(1 - t_1),$$

and student 2's total utility is:

$$u_2(t_1, t_2) = \ln(t_1 + t_2) + \frac{1}{2}\ln(1 - t_2),$$

the benefit from better homework, plus the benefit from leisure.

We have now specified the three components of a game: the set of players (here the two students, $i = 1, 2$), the players' strategy sets (here just the set of all t_i between 0 and 1), and the utility functions (see above). Note that each student's utility depends on the student's own choice, but also on the other student's choice. This is reflected by the first of the two terms in the utility function: both students' efforts enter.

We are going to find the Nash equilibrium, or the Nash equilibria, of this game. First, let us imagine we were student 1, and we had to decide how much we want to work. Because student 1's utility not only depends on student 1's effort level but also on student 2's, also student 1's optimal choice may depend on student 2's choice. Let's check whether it does, and, if so, how it depends on student 2's choice.

Suppose student 1 maximized her utility, that is: $u_1(t_1, t_2) = \ln(t_1 + t_2) + \ln(1 - t_1)$, expecting t_2 to be a certain level. For now, we just leave open which level that is, and just write "t_2" for it. The first order condition for student 1's optimal choice is:

$$\frac{\partial u_1}{\partial t_1} = 0.$$

Finding the derivative, and then solving for t_1, we get:

$$\begin{aligned}
\frac{1}{t_1 + t_2} - \frac{1}{1 - t_1} &= 0 \Leftrightarrow \\
\frac{1}{t_1 + t_2} &= \frac{1}{1 - t_1} \Leftrightarrow \\
t_1 + t_2 &= 1 - t_1 \Leftrightarrow \\
2t_1 &= 1 - t_2 \Leftrightarrow \\
t_1 &= \frac{1}{2} - \frac{1}{2} t_2
\end{aligned}$$

To make sure we have a maximum, we also calculate the second derivative:

$$\frac{\partial^2 u_1}{(\partial t_1)^2} = -\frac{1}{(t_1 + t_2)^2} - \frac{1}{(1 - t_1)^2}.$$

The second derivative is clearly negative for all t_1, so that u_1 is concave in t_1, and therefore we can be sure that the solution to the first order condition is a maximum.

To summarize, we find that student 1's optimal choice, anticipating that student 2 will choose t_2 is:

$$t_1 = \frac{1}{2} - \frac{1}{2}t_2.$$

The expression on the right hand is also called player 1's best response function, and sometimes we also write:

$$\mathrm{BR}_1(t_2) = \frac{1}{2} - \frac{1}{2}t_2,$$

where the "BR" stands for "best response."

We see that, indeed, student 1's optimal choice depends on student 2's choice. It is decreasing in t_2, unsurprisingly. The more time the other student puts into the homework, the less time student 1 puts into the homework.

To figure out what his best choice is, student 1 still must form some expectation of what student 2 is going to choose. To do that, he might think about student 2's optimal choice. He can calculate it, using the same calculus approach that we used here for student 1, and he will find:

Figure 68 shows the graphs of the two best response functions.

$$\mathrm{BR}_2(t_1) = \frac{2}{3} - \frac{1}{3}t_1.$$

Thus, student 2's optimal choice depends on what student 2 expects student 1 to choose. But this raises a new difficulty for student 1: he has to predict how student 2 thinks about student 1's choice. Maybe student 2 puts himself into the shoes of student 1, and calculates that student 1's optimal choice is:

$$\mathrm{BR}_1(t_2) = \frac{1}{2} - \frac{1}{2}t_2.$$

Now, student 1 recognizes that he must figure out what student 2 expects that student 1 expects that student 2 does.

This is a moment at which, naturally, you might experience a little bit of mental fatigue. We might wish to take a rest, and then continue as before, but, wearily, we anticipate that this may well go on forever. Perhaps, though, we can come up with an approach that helps us cut through this Gordian knot. This is what the concept of Nash equilibrium does. Instead of modeling players who work through the long sequence of thoughts about other players' behavior that we started to indicate in the previous paragraph, we model players who somehow have learned to correctly predict the other player's choice.

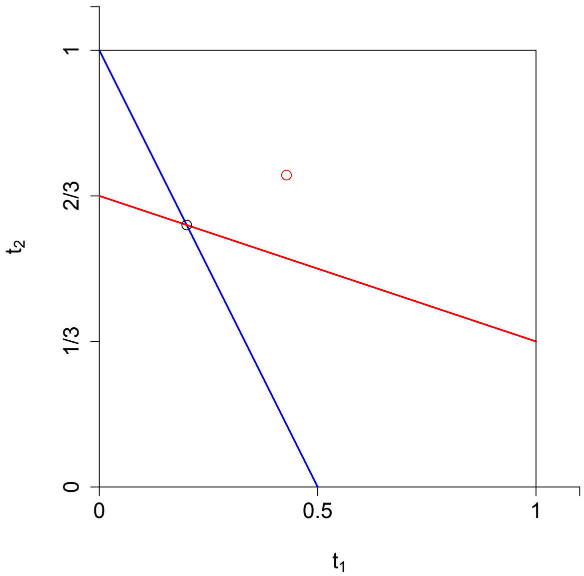


Figure 68: Best Response Functions for Example 1

The best response function of player 1 is indicated in blue. The best response function of player 2 is indicated in red. The black circle indicates the Nash equilibrium. The red circle indicates the effort levels that maximize the sum of utilities.

A Nash equilibrium must solve these two equations:

$$\begin{aligned} t_1 &= \frac{1}{2} - \frac{1}{2}t_2 \\ t_2 &= \frac{2}{3} - \frac{1}{3}t_1 \end{aligned}$$

These equations say that each choice must be a best response to the other player's choice. That is what the definition of a Nash equilibrium requires. If we find values t_1^* and t_2^* that solve these equations, then we get a pair of values such that, if student 1 anticipates that student 2 will choose t_2^*, then student 1's optimal choice is t_1^*; and if student 2 anticipates that student 1 will choose t_1^*, then student 2's optimal choice is t_2^*.

It is not hard to solve the two equations. We substitute the second equation into the first, and get:

$$t_1 = \frac{1}{2} - \frac{1}{2}\left(\frac{2}{3} - \frac{1}{3}t_1\right) = \frac{1}{6} + \frac{1}{6}t_1,$$

which we can solve as follows:

$$\begin{aligned} t_1 &= \frac{1}{6} + \frac{1}{6} t_1 \Leftrightarrow \\ \frac{5}{6} t_1 &= \frac{1}{6} \Leftrightarrow \\ t_1 &= \frac{1}{5}. \end{aligned}$$

Plugging back into the equation for t_2:

$$t_2 = \frac{2}{3} - \frac{1}{3} \cdot \frac{1}{5} = \frac{3}{5}.$$

In summary, our equilibrium choices are:

$$t_1^* = \frac{1}{5} \text{ and } t_2^* = \frac{3}{5}.$$

In the Nash equilibrium, student 1 spends only 1/5 of an hour, that is, 12 minutes, on the homework, and student 2 works for 36 minutes. That student 2 works longer than student 1 reflects that student 2 has half as much utility from leisure, that is, half as large opportunity cost of work, as student 1. Of course, we need to do the calculation above to figure out that this leads to student 2 doing exactly three times as much work as student 1.

The Nash equilibrium is marked in Figure 68 with a black circle.

Before we leave the example, let us do one more calculation. We might wish to ask whether there is some way for the students to raise both their utilities, by making different choices. How is this even logically possible? Observe that the defining characteristic of a Nash equilibrium is *not* that both players make choices that ensure that *both* players' utilities are maximized; it is that each player maximize *their own utility*, taking as given and fixed the other player's choice. No player ever considers the effect of their actions on the other player's utility.

As a heuristic to finding what is "best for both," we might try maximizing the sum of the two students' utilities. Recall that adding utilities isn't really very meaningful, given that utilities are only "ordinal." This is why I call the calculation that is about to follow a "heuristic." We are trying out something that looks simple and natural, but we have no completely solid reason why what we are doing makes sense.

After this apology, let us find the sum of the players' utilities:

$$\begin{aligned} & u_1(t_1, t_2) + u_2(t_1, t_2) \\ = \quad & \ln(t_1 + t_2) + \ln(1 - t_1) + \ln(t_1 + t_2) + \frac{1}{2}\ln(1 - t_2) \\ = \quad & 2\ln(t_1 + t_2) + \ln(1 - t_1) + \frac{1}{2}\ln(1 - t_2) \end{aligned}$$

Our objective is to find the values of t_1 and t_2 that maximize this sum. This is a two-dimensional maximization problem, because there are two choice variables that enter the expression we want to maximize. You might not be familiar with how to solve such maximization problems, and we shall not be completely rigorous about how to do this. But it might seem plausible that, analogous to one-dimensional optimization problems, we have to set some first derivative equal to zero. This would mean here that the first derivative of the sum with respect to t_1, given t_2 must be zero, and that the first derivative of the sum with respect to t_2, given t_1, must be zero. These two first derivatives are partial derivatives. Let's calculate them and set them equal to zero:

$$\frac{2}{t_1 + t_2} - \frac{1}{1 - t_1} = 0 \text{ and } \frac{2}{t_1 + t_2} - \frac{1}{2}\frac{1}{1 - t_2} = 0.$$

The first equation sets the partial derivative with respect to t_1 equal to zero, and the second equation sets the partial derivative with respect to t_2 equal to zero. We can re-write the two equalities as:

$$\frac{2}{t_1 + t_2} = \frac{1}{1 - t_1} \text{ and } \frac{2}{t_1 + t_2} = \frac{1}{2}\frac{1}{1 - t_2}.$$

Cross multiplying yields:

$$t_1 + t_2 = 2(1 - t_1) \text{ and } t_1 + t_2 = 4(1 - t_2).$$

We have two linear equations in two unknowns. Let me, for brevity, leave it to you to work out the steps needed to solve them. The solution is:

$$t_1 = \frac{3}{7} \text{ and } t_2 = \frac{5}{7}.$$

We have not been completely rigorous. But I can assure you that indeed these two choices maximize the sum of the two students' utilities.

The values of t_1 and t_2 that maximize the sum of players' utility are marked in Figure 68 with a red circle.

What have we shown? If both students sought to maximize the sum of utilities, they would both spend *more* time on homework than they do in the Nash equilibrium: 3/7 is more than 1/5, and 5/7 is more than 3/5. Do these choices make both agents better off, or do they just raise the *sum* of utilities? In the Nash equilibrium, student 1's utility is $\ln(1/5+3/5)+\ln(1-1/5) \approx -0.446$ (recall that utility being negative has no particular meaning). When students maximize the sum of utilities, then student 1's utility is: $\ln(3/7+5/7)-\ln(1-3/7) \approx -0.426$, and thus student 1's utility has increased. For student 2, utility in the equilibrium is: $\ln(1/5+3/5)+0.5\ln(1-3/5) \approx -0.681$ whereas when both students maximize the sum of utilities it is: $\ln(3/7+5/7)+0.5\ln(1-5/7) \approx -0.493$, which is also an increase. We find that indeed both students would be better off if they worked more. The Nash equilibrium is *not* Pareto efficient.

Example 2: A Picnic

To become further familiar with the concept of Nash equilibrium, let us consider another example. N people participate in a picnic. Each person i has to choose how much of a certain dish to bring: $0 \leq q_i \leq 1$. The cost of bringing q_i units are: $0.5 \cdot q_i$. Each person i has Cobb-Douglas benefit from the picnic: $q_1 \cdot q_2 \cdot ... \cdot q_N$. Therefore, the total utility is:

$$u_i(q_1, q_2, \ldots, q_N) = q_1 \cdot q_2 \cdot \ldots \cdot q_N - 0.5 \cdot q_i.$$

To find Nash equilibria of this game, we shall restrict our search to strategy combinations such that everyone chooses the same quantity $q_i = q^*$. Suppose i anticipates that everyone else will choose $q_j = q^*$. We have a Nash equilibrium if i's best choice is then to also choose $q_i = q^*$. For which values of q^* is this true?

Of course, as Example 1 shows, in general in Nash equilibria not all players make the same choice. Even if, as in Example 2, all players are ex ante identical, it need not be the case that in Nash equilibrium all players make the same choice. Therefore, the assumption that we are making: $q_i = q^*$ for all players i, is not without loss of generality.

What is i's optimal choice when everyone else chooses q^*? Player i's utility is then:

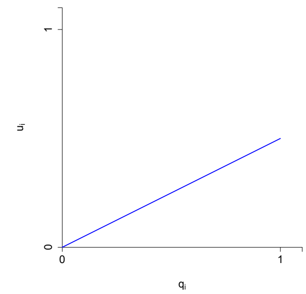
$$(q^*)^{N-1} \cdot q_i - 0.5q_i.$$

We can re-write this as:

$$\left((q^*)^{N-1} - 0.5\right) \cdot q_i.$$

Now observe that this is linear in q_i with slope $(q^*)^{N-1} - 0.5$. Thus, the graph of player i's utility function, with q_i on the horizontal axis and utility on the vertical axis, is a straight line. Recall that player i can only choose a value of q_i that is between zero and 1. Therefore, if the straight line is upwards sloping, i's optimal choice is $q_i = 1$. If it is downwards sloping, i's optimal choice is $q_1 = 0$. If it is horizontal, then any choice of q_i is optimal.

Here is a graph that shows player i's utility function, with her own choice q_i on the horizontal axis, in the case that $(q^*)^{N-1} - 0.5 > 0$.



In the case shown here, the best choice for player i is $q_i = 1$. This is the first of the three bullet points. The second is about the case that the line is downward sloping, going into negative utility, and the third is about the case that the line is horizontal, coinciding with the horizontal axis.

We have calculated the slope of the line that represents player i's utility above. It is: $(q^*)^{N-1} - 0.5$. Hence, for example, it is positive if: $(q^*)^{N-1} > 0.5$. We can therefore summarize player i's optimal choices now as follows:

- If $(q^*)^{N-1} < 0.5 \quad \Rightarrow \quad$ Optimal choice: $q_i = 0$.
- If $(q^*)^{N-1} > 0.5 \quad \Rightarrow \quad$ Optimal choice: $q_i = 1$.
- If $(q^*)^{N-1} = 0.5 \quad \Rightarrow \quad$ Any choice $0 \leq q_i \leq 1$ is optimal.

The graph of this best response function for player i is in Figure 69.

Now recall that looking for a Nash equilibrium means here that we are looking for a value q^* such that, if everyone other than i chooses q^*, then it is optimal for i to also choose q^*. Let's search for such values. Let's begin with $q^* = 0$. In this case, $(q^*)^{N-1} < 0.5$, and therefore player i's optimal choice is $q_i = 0$, and hence there is indeed a Nash equilibrium in which everyone chooses $q_i = q^* = 0$. In this Nash equilibrium, no picnic takes place.

Let's look at larger values of q^* next. If q^* is strictly positive, but $(q^*)^{N-1} < 0.5$, then i's best response is zero, and therefore we do not have a Nash equilibrium. This applies to all q^* such that $(q^*)^{N-1} < 0.5 \Leftrightarrow q^* < 0.5^{\frac{1}{N-1}}$. None of these is a Nash equilibrium.

What about the case that $q^* = 0.5^{\frac{1}{N-1}}$? Then the line that represents player i's utility is horizontal, and every choice is optimal. In particular, of course, then also $q_i = 0.5^{\frac{1}{N-1}}$ is an optimal choice for player i. Thus, is this a Nash equilibrium? It really is a matter of definition. But the precise definition of Nash equilibrium, as we gave it above, just requires that player i can not find a strategy that gives him strictly higher utility than the equilibrium strategy. It is OK if player i can find some other strategy that gives him exactly the same utility. Therefore, in this example, in the language of game theory one would say that $q^* = 0.5^{\frac{1}{N-1}}$ is indeed a Nash equilibrium.

Let us consider even higher values of q^*, namely values such that: $q^* > 0.5^{\frac{1}{N-1}}$. In this case, player i's best response is always $q_i = 1$. Therefore, the only value that corresponds to a Nash equilibrium in this range is: $q^* = 1$.

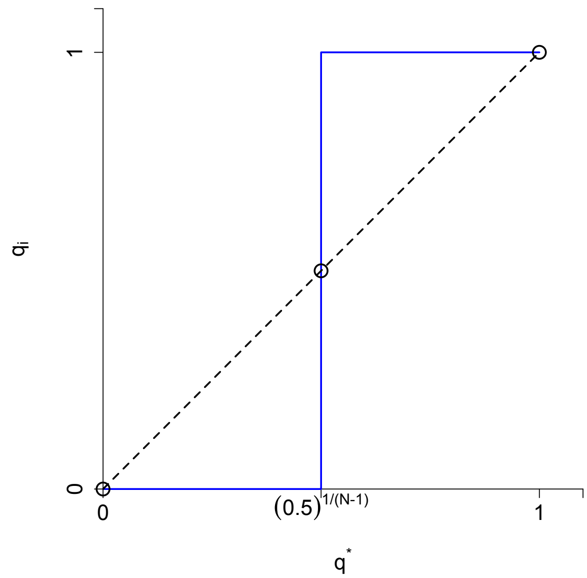


Figure 69: Best Response Function for Example 2

The best response function of player i is indicated in blue. q_i is on the vertical axis. On the horizontal axis is q^*. The figure shows the best response of player i if *all* other players choose q^*. The dashed line is the 45° line. The three Nash equilibria are circled in Figure 69.

The bottom line of this discussion is that there are three Nash equilibria:

- $q^* = 0$, which means that there is no picnic, and everyone's utility is zero.
- $q^* = 0.5^{\frac{1}{N-1}}$, in which case there is a picnic, and everyone's benefits from the picnic equal their cost of contributing to the picnic, so that everyone's utility is zero.
- $q^* = 1$, in which case there is a picnic, and everyone benefits more from the picnic than they contribute, so that everyone has positive utility (which happens to be 0.5).

It happens frequently that games have more than one Nash equilibrium. There will be more examples below. What does it mean that there are several Nash equilibria? If we interpret Nash equilibria as the outcome of a process by which players have come to be able to predict the other players' choices, then we have multiple equilibria if this process can arrive at multiple steady states. It may then, for example, depend on the history which equilibrium has been reached. In our example, if everyone has become used to expect that nobody else contributes to the picnic, then this will be self-enforcing. But if everyone brings $q^* = 1$ to the picnic, then this can also become self-enforcing. The same is true for $q^* = 0.5^{\frac{1}{N-1}}$. Intuitively speaking, this example shows that, whether a party is a success, may depend on whether everybody has come to expect the party to be a success.

Example: Pick a Number

There are N players. Every player i picks a number between 0 and 100. The winner is the player who picked the number that is closest to 2/3 * (the average of everyone else's number). The winner gets $1. If there are several winners, the Dollar is equally shared among them.

It is easy to find best response functions. Player i should simply choose:

$$x_i = \frac{2}{3} \cdot \frac{x_1 + x_2 + \ldots + x_{i-1} + x_{i+1} + \ldots + x_N}{N-1}.$$

The numerator on the right hand side is just the sum of everyone else's numbers. To find Nash equilibria we have to find combinations of numbers, one for each player, such that this condition holds for all players.

Let's use a little trick to find all such combinations of numbers. Let's write the best response conditions for players 1 and 2:

$$\begin{aligned} x_1 &= \frac{2}{3} \cdot \frac{x_2 + x_3 + \ldots + x_N}{N-1} \\ x_2 &= \frac{2}{3} \cdot \frac{x_1 + x_3 + \ldots + x_N}{N-1} \end{aligned}$$

Next, we subtract these two equations from each other:

$$x_1 - x_2 = \frac{2}{3} \frac{x_2 - x_1}{N-1}$$

which is equivalent to:

$$\begin{aligned} x_1 - x_2 + \frac{2}{3(N-1)}(x_1 - x_2) &= 0 \Leftrightarrow \\ \frac{3(N-1)+2}{3(N-1)}(x_1 - x_2) &= 0 \end{aligned}$$

One of the two factors on the left hand side must be zero. Because the fraction is not zero, it must be that the term in brackets is zero, which means, it must be that:

$$x_1 = x_2.$$

Whereas in Example 2, for simplicity, I restricted attention to Nash equilibria in which all players make the same choice, for Example 3 I prove that there are only Nash equilibria in which all players make the same choice. This is because my objective in Example 3 is to show that there really is just one Nash equilibrium, and that it is somewhat surprising which choices players make in this Nash equilibrium.

We could have derived the same result for any arbitrary pair of players. Therefore, we can conclude that all players must choose the same number:

$$x_i = x^* \text{ for all players } i.$$

Now suppose that all players other than player i choose the number x^*. Which number should player i choose? We can go back to the first order condition derived earlier and plug in x^* for the choices of all players other than i. We get:

$$x_i = \frac{2}{3} \cdot \frac{(N-1)x^*}{N-1} = \frac{2}{3}x^*.$$

I show this best response function in Figure 70.

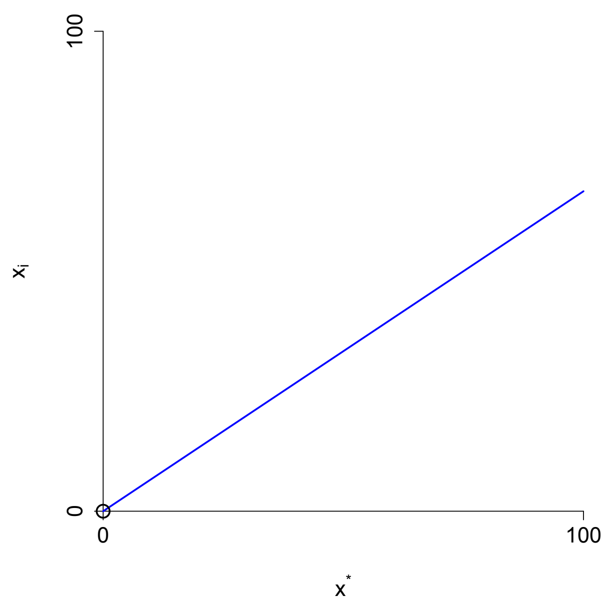


Figure 70: Best Response Function for Example 3

The best response function of player i is indicated in blue. x_i is on the vertical axis. On the horizontal axis is x^*, the number that all players other than i choose. The only Nash equilibrium of the game is circled in Figure 70.

Now, because all players have to choose the same number x^*, player i must find it optimal to choose $x_i = x^*$ when players play a Nash equilibrium. Plugging this in we find:

$$x^* = \frac{2}{3}x^*.$$

This is equivalent to:

$$\frac{1}{3}x^* = 0.$$

Therefore, the only Nash equilibrium is:

$$x^* = 0.$$

That is, all players choose the number 0. All players “win.” All players get exactly $1/N$ of the prize.

Interestingly, many experiments with this game have been conducted, including experiments where the game was described in large national newspapers, and readers were invited to submit by email their strategy choices. I have to admit that typically people do not choose the Nash equilibrium. What has not been tested, though, is whether, if one plays the game repeatedly, and lets people adjust their choice as they see what others do, in the long run play will converge towards the Nash equilibrium.

More Examples

I now go with you through a couple of examples of games in which each player only has a finite number of possible strategies. In all examples so far, each player had infinitely many strategies. Also, in the examples in this section there are only two players. I show you both a common way of representing such games in a table, and how to find Nash equilibria of the game in the table.

The first game is the most famous example in game theory: the Prisoners' Dilemma. The players in this game are two prisoners. Both are in custody. The authorities have limited evidence: they can prove for each one separately that they have committed a minor crime, which will put them each for 2 months in prison. Maybe each one of them can easily be convicted of not having returned the coffee mug when having coffee in a coffee shop. But the authorities suspect that the two prisoners have committed another, more significant crime, together. This crime might have been hacking together my computer, and putting errors into my lecture notes (just an example). Suppose this major crime could be punished with an additional 8 months in prison.

The prosecutor has a problem: she has no evidence she could use to convict the prisoners' of the major crime yet. She only suspects that they are guilty. The only way they can be convicted of the major crime is by confession. There is no evidence. But one confession is enough to convict both. Suppose, now, that the prosecutor sets up the following scheme: Each prisoner will be interviewed separately, and will be invited to confess to the major crime. If one confesses, and the other one does not confess, the one who has confessed will go away free, whereas the one who has not confessed will be punished for both, the minor, and the major crime. If both confess, then the punishment for the minor crime is waived, but they have to serve prison time for the major crime. In Figure 71 I show a table that represents this game.

	Don't Confess	Confess
Don't Confess	-2,-2	-10,0
Confess	0,-10	-8,-8

Figure 71: The Prisoners' Dilemma

The players in the game are the two prisoners. Let's call them "player 1" and "player 2." Each has two choices. Player 1's choices correspond to rows in Figure 1, and player 2's choices correspond to columns in Figure 71. Once we have picked one choice for each player, we can find the cell where the row and the column intersect. In each cell we display two numbers. The first one is the utility of player 1. The second one is the utility of player 2. Thus, for example, if player 1 confesses, and player 2 does not confess, player 1's utility is 0, and player 2's utility is -10. Note that I have chosen the utilities to be equal to months in prison, and I have put a minus sign in front of the prison months because, probably, utility decreases the more time one has to spend in prison.

Let us try and find Nash equilibria of this game by determining best responses. For each player, we only have to find two best responses, one for each of the two strategies of the other player. For example, let's determine the best response for player 2 if player 1 does not confess. If player 2 does not confess, then her utility is -2, whereas, if she confesses, it is 0. Therefore, her best response is to confess.

In a similar way, for each player, for each strategy of the other player, we can find some best response. In Figure 72 I indicate in red for the Prisoners' Dilemma the maximum utility that player 1 can get if she chooses a best response to player 2's strategy, and I indicate in blue the maximum utility that player 2 can get if she chooses a best response to player 1.

	Don't Confess	Confess
Don't Confess	-2,-2	-10,0
Confess	0,-10	-8,-8

Figure 72: Best Responses in The Prisoners' Dilemma

We have a Nash equilibrium when, in a cell, both utility numbers are colored. In Figure 72, you can see that the Prisoners' Dilemma has just one equilibrium, namely the strategy pair (Confess, Confess). Note that in this Nash equilibrium both players get to spend 8 months in prison. If they didn't confess, they would spend only 2 months in prison. That would be better for both of them. The game is called a "dilemma" because if players act in their own interest, the outcome "don't confess" will not occur. It is not stable. Each player, when assuming that the other player does not confess, can reduce their prison time by confessing.

Our next example is called the game of "Chicken," and the name refers to the verb to "chicken out," meaning, to avoid a confrontation because of fear. The story that is often told with this game is this: two car drivers drive towards each other in the middle of the road. They are testing each other's courage. Each can either stay on course, or swerve to the left. If both stay on course, the outcome is disastrous for both. The one who serves to the left first, loses out to the one who stays on course, although at least she survives. If both swerve, both survive, but neither has proven their courage. In Figure 73 I display the game with utilities that represent what may be players' incentives in this game.

One can see in the suggestion that chicken run away from dangerous situations either a slander about chicken, or a compliment for chicken.

irrational (?)

	Stay On Course	Swerve to the Left
Stay On Course	-2,-2	2,-1
Swerve to the Left	-1,2	0,0

Figure 73: The Game of Chicken

For both players, the lowest utilities, -2, are realized if they both stay on course, and therefore crash into each other. if both swerve to the left, they both survive, but receive somewhat unimpressive utilities of 0, because neither has been able to demonstrate their "courage." Each player has the highest utility, 2, if they stay on course whereas the other swerves to the left. The player who is then revealed to be less courageous than the other gets utility -1, which is bad, but still better than getting killed in a crash.

What are Nash equilibria of this game? I have shown in Figure 73 the best response functions, using the same color coding as in Figure 72. One can easily see that there are two Nash equilibria: one where player 1 stays on course and player 2 swerves to the left, and another one where player 2 stays on course and player 1 swerves to the left. Thus, game theory does not predict which of the two players gives up first. It does predict that both of them trying to prove their courage, or both of them giving in, would not be a Nash equilibrium.

One may see each of the little games that we analyze in this section as a metaphor of something much larger. This is also true of the game of Chicken. We encounter games of Chicken frequently in real life: two sides with opposing wishes wait for one or the other to give in. The fact that there are two Nash equilibria may indicate that, who gives in, is determined ultimately by exterior characteristics of the players that have nothing material to do with the game, such as players gender. Maybe, women always give in, and men always get their way, or the other way round. Or maybe those who "power-dress" always get their way, and those who "under-dress" always lose. Once such an arbitrary convention is established, it becomes hard to break it. Nash equilibria may just be the reflection of arbitrary conventions.

The story that is typically told with our next example is that of a children's game. Two children each have a coin, and must secretly put it down with one side up, either heads or tails. If both of them turn out to have put down the coins with the same side up, then player 1 must pay a Dollar to player 2. If they have put down the coins with different sides up, then player 2 must pay a Dollar to player 1. Figure 74 shows this game.

	Head	Tails
Head	-1,+1	+1,-1
Tails	+1,-1	-1,+1

Figure 74: Matching Pennies

There is no Nash equilibrium in this game. I have marked best responses in Figure 74 with the same color coding as before, and you will see that there is no box in which both utility numbers are colored. The intuition is simple.

In this game, player 2 wants to do exactly the the opposite of what player 2 does, whereas player 2 wants to make the same choice as player 1 does. There cannot be an outcome of the game that is "in equilibrium," that is, an outcome that both players foresee, and that neither of them wants to change.

Intuitively, one can say that one player chases the other in this game. But the other player can run away, and thus there is no equilibrium. In game theory, this game is often used to motivate the concept of a "mixed strategy," that is, a strategy where a player chooses a strategy randomly. Suppose both players toss their coins into the air, and then leave them just the way they fall down. Then each side comes up for each player with probability 1/2. No player could do better by deliberately choosing one side or the other. We say in game theory that this game has a "mixed strategy" equilibrium in which both players choose each of the two sides with probability 1/2. Mixed strategies are strategies where players leave the choice of their strategy to some randomization device, and only choose the probabilities with which of the strategies is chosen. The "real strategies" are then called "pure strategies." Matching Pennies has no Nash equilibrium in "pure strategies," but it does have a Nash equilibrium in "mixed strategies." In this course, we will not study mixed strategies.

We conclude this section with an example that is not one of the famous examples of game theory, but that helps us to see how the method for studying Nash equilibria shown in the previous examples can be generalized to study Nash equilibria of an arbitrary game with two players in which each player's strategy set is finite. Consider the game in Figure 13.

	2	3	4	5
2	10,10	8,12	6,12	4,10
3	12,8	9,9	6,8	3,5
4	12,6	8,6	4,4	0,0
5	10,4	5,3	0,0	-5,-5

Figure 75: A Cournot Game

I have labeled the game as a "Cournot game," because it is a version of the Cournot game that we study in the next Topic, except that I have made the strategy sets finite, whereas in the next topic they will be infinite. I shall postpone the story that goes with this type of game until the next Topic. I have marked best responses in Figure 13 using the same color coding as before. Note that player 1 has sometimes several best responses, and sometimes just one, but that player 1's best responses decrease in size as player 2's strategy increases. The same is true for player 2's best responses, because the game is symmetric.

Strategies are numbers in Figure 13.

The game has three Nash equilibria: (4,2), (3,3), and (2,4). Note how these equilibria give the players different utilities. Each player prefers equilibria in which she chooses a higher number over equilibria in which she chooses a lower number. The equilibrium (3,3) seems like the most equitable outcome. Yet, it co-exists with equilibria that give one, or the other player an advantage.

These phrases make only sense if we compare different players' utilities with each other. In the Cournot model in the next section, the players will be firms, and the utilities will be profits. Perhaps it does make sense to compare profit numbers.

Topic 22: Imperfect Competition

We now want to use the methods that we introduced in the two previous topics and develop models of “imperfect competition.” Recall that by “perfect competition” we mean a market form in which no market participant thinks that they have any influence on the market price. By “imperfect competition” we mean any market form in which pricers are not taken as given by all market participants but in which there are at least some market participants who have some impact on prices, namely, some market participants choose prices. Clearly, in reality, prices must be determined by people who have some influence on these prices, and models in which we describe the process by which people determine prices have, perhaps, a chance of capturing important aspects of the world that the models considered so far have left out.

One situation that we often encounter in life is that we enter a store, choose something that we want to buy, and find that the seller has put a price tag on what we want to buy: the seller has chosen the price; as buyers, we have no impact on the price. We either take it or leave it. This is a very one-sided situation. As buyers, our role is very limited. Occasionally, it is the other way around. A buyer might announce that he is willing to pay a certain price for every collector’s item that he is offered, say a somewhat rare stamp. Then the buyer chooses the price, and the sellers can either sell, or not sell at this price, but cannot change the price.

In real life, we also encounter somewhat more balanced situations. Sometimes, there are negotiations between buyers and sellers over a price. For example, house prices are typically negotiated. Then buyers *and* sellers can influence the price.

Which prices are chosen, or how negotiations work out, will depend on the number of sellers offering similar or identical goods. It will also depend on the number of buyers trying to buy some of these goods. Prices will presumably also depend on *how* similar the goods offered by the sellers are, and how much buyers differ in their preferences.

There are microeconomic models of all these different ways in which buyers and sellers can agree on a price. A course like this is too short to consider all such models. We shall focus here on a subset of such models, namely those in which the sellers have the power to dictate the price. When roles are reversed,

And, perhaps, not just this course but even life itself is too short to consider all models that have been proposed by now.

and the buyers have the power to dictate the price, the theory is the same as the one which we present here, just with roles reversed. Models of markets in which prices are negotiated by both sides of the market are very interesting, but require more complicated, game-theoretic arguments than we can consider in this course.

Monopoly

We begin with a very simple and stylized model: there is just one seller. In the language of economics we call a seller who is the only one selling a particular good a "monopolist." So, we start with a model of monopoly.

No, by "monopoly" we don't mean the board game. Sorry.

Buyers may be utility maximizing, and may choose among many different goods. But assuming that all other goods' prices are given and fixed, we can describe the demand for the good that the monopolist has to offer by just one market demand function: $D(p)$. Although, as you know, demand need not always be downward sloping, we shall assume in this section that it is, to make things simpler. The downward sloping demand function summarizes buyers' behavior in our model. Beyond this, buyers will make no further appearance in the model.

The seller might have a given quantity of the good that she wants to sell, or she might not yet have produced the good, and only produce "on demand," once she knows how much her customers want to buy. These are two slightly different situations, and, for simplicity, we shall initially focus on the situation where production takes place after the monopolist knows how many people want to buy. Because the seller is then a producer, we shall regard her as a firm.

Earlier, we discussed how to determine a firm's cost function. Here, we take for granted that the firm has found for every quantity q what is the cost minimizing way of producing q, and we write for the corresponding cost $C(q)$.

Suppose the seller picks a price p, announces this price to consumers, and then waits to see how much consumers want to buy. Suppose also that the seller's goal is to maximize her profits. What will her profits be? They will be equal to revenue, $pD(p)$, minus cost, $C(q)$, where the quantity q is what buyers want to buy, i.e. $q = D(p)$, so that, putting things together, we obtain for profits:

$$pD(p) - C(D(p)).$$

We shall now study how to choose the optimal p. We shall assume that the firm has exact knowledge of the demand function (p). It turns out that this is then easiest if we re-write our problem a little bit. Instead of focusing on the choice of p, we can focus on the choice of q. This is because a downward sloping demand function establishes a one-to-one relationship between prices and quantities. Whether the monopolist thinks about which price he wants to set, or the quantity q that she wants to sell amounts to the same thing.

Let us take then the quantity sold as the firm's choice variable. We shall denote the price at which any quantity q can be sold by $P(q)$. The function P is called the "inverse demand function." It is indeed - mathematically speaking - the inverse function of the demand function.

Now we can write the firm's profit as:

$$\pi(q) = qP(q) - C(q).$$

We use here the Greek letter "π" to indicate profits. To find the optimal quantity, we set the first derivative of the profit function equal to zero:

$$\pi'(q) = 0.$$

To calculate the derivative, we have to apply the product rule to the expression for revenues. Differentiating the cost term is trivial. Thus, the first derivative is zero if:

$$qP'(q) + P(q) - C'(q) = 0.$$

Re-writing this a tiny bit, we get:

$$P(q) + qP'(q) = C'(q).$$

Here, the expression on the left hand side is the derivative of revenue, obtained using the product rule. We also call this derivative the marginal revenue. That is, "marginal revenue" is the entire expression on the left hand side of the above equation. The right hand side is marginal cost. The first order condition is often expressed as:

$$\text{marginal revenue} = \text{marginal cost}$$

We'll get back to this condition, but for a moment let us look at the second order condition. The second derivative of profit is:

We should also check for the possibility of "corner solutions." But let's not bother.

$$\pi''(q) = qP''(q) + P'(q) + P'(q) - C''(q).$$

When looking for a maximum, we want this expression to be negative. Suppose marginal cost are increasing, i.e. $C''(q) > 0$. This an assumption that we have used earlier already. Then the last expression is negative, because $C''(q)$ gets subtracted in the expression for the second derivative of profit. If demand is downward sloping, then also $P'(q)$ will be negative, so that the two terms in the middle of the second derivative of profit are negative. What remains is the first term. It may be positive, and still the whole second derivative is positive. But if we want to be safe, we can assume that $P''(q) \leq 0$, that is, that inverse demand is concave. With that assumption, the condition "marginal revenue = marginal cost" definitely gives us a maximum rather than a minimum.

Inverse demand is concave if demand is concave. So, this just means that demand is concave.

Now let's focus on the condition "marginal revenue = marginal cost." Notice its similarity with the condition "price = marginal cost," which we had identified as a necessary condition for profit maximization for a firm that takes prices as given and has increasing marginal cost. How does marginal revenue differ from price? Marginal revenue is: $P(q) + qP'(q)$, so that the difference with price is: $qP'(q)$. Observe that this is negative, because $P'(q)$ is negative. Marginal revenue is lower than the price.

Why is this? When the firm takes prices as given, and not affected by its own choices, and increases the quantity produced by 1 unit, then the extra revenue is just the price. But if the firm understands that the price will drop if it increases its quantity, there are two effects of an increase in revenue. The first is the effect that also a price-taker experiences: it will gain revenue, because it has produced a larger quantity. But the second effect is specific to the firm which is not a price-taker: it will lose some revenue, because the price has to be lowered. This price drop is $P'(q)$, and it applies to all the quantities that the firm sells, so that the total loss is: $qP'(q)$.

Suppose the firm considers how far to expand its output: as the output gets bigger, marginal cost increase: each unit has a higher marginal cost. The firm will want to stop expanding output once the extra revenue equals the extra marginal cost. Because marginal revenue is lower than price, it will stop expanding output earlier than a firm that just takes the price as given, and assumes that its own output decisions will not affect the price. Therefore, the monopolist will produce *less* than a firm that takes the price as given, i.e. a competitive firm. This is the distortion caused by monopoly: it produces less than a competitive firm. Because it produces less, the price will be higher than under perfect competition.

Auctions

A monopolist can often do better than just setting a price. When setting the price, the monopolist might not know what the buyers' willingness to pay is. He then has to guess how much he can get when he sets the price. But he might guess wrong. In such situations, auctions can do better. They can involve the buyers in a competition, and thereby drive the price up. Houses are often sold through auctions, as are works of art. Some government-issued licenses are sold through auction, for example licenses that allow the license holder to produce a certain amount of pollution. Licenses to use the radio spectrum for cell phone services are auctioned. Perhaps these are all contexts in which the seller's uncertainty about the buyers' willingness to pay is particularly large.

There is a very well-developed economic theory of auctions. It is an application of game theory. One question that researchers have studied particularly carefully is that of the optimal auction rules. It may not be obvious, but there are many different rules by which one can conduct an auction. There is for example, the silent auction, in which bidders submit bids, the highest bid wins, and the winner pays what she bid. This auction is in the language of auction theory called a "sealed-bid first price auction." But there are many alternatives. For example, you may be familiar with open outcry auctions, where bidders stand together in one room, and make and revise their bids openly, overbidding each other. Auction theory has tried to determine which of these auction rules gives the seller the highest revenue.

We shall focus on one particular auction rule that is not often observed in practice in exactly the form that we shall describe, but that is similar to an auction format that is often used in the real world. This auction format is the "sealed-bid second price auction." All bidders submit bids, the highest bidder wins, and - here is a surprise - does not pay her own bid, but pays what the second highest bidder bid. Thus, if the bids are \$5, \$10, \$12, and \$15, then the person who bid \$15 wins, but only has to pay \$12.

This looks strange: if you bid \$15, why doesn't the seller make you pay \$15? In fact, the sealed-bid, second price auction is a simplified version of another auction that is more familiar: suppose, in an open outcry auction bidders overbid each other, until, in the end, only one bidder remains, who then pays her bid. Suppose that, in fact, each bidder is not present in person, but sends a representative who has instructions to continue bidding until the price reaches some value $\$x$. Then the representative for whom the upper ceiling $\$x$ is highest will win, but, of course, he does not actually have to bid up to this level. Bidding ends when the person with the second highest bid limit drops out. At that point only one bidder is left. Thus, the bidder with the highest bid limit wins, but only has to pay the second highest bid limit. You can thus think of the bids in the sealed-bid, second price auction as the bid limits that bidders assign to their representatives in an open outcry auction.

This happens often in practice.

Suppose you are a bidder in a sealed-bid, second price auction. You are contemplating which bid to place. Suppose your value for the object is v. If you bid b and b turns out to be the highest bid, then your net utility is: $v - b_{(2)}$, where $b_{(2)}$ is the second highest bid, the one just below yours. If you lose, let's normalize your utility to be zero. How much should you bid? Obviously, the outcome will not only depend on your own bid, but also on the other bidders' bids. Thus, you are involved in a game.

Following the methods we developed in the previous topic, we can determine the bidders' best response functions, and then try to figure out a Nash equilibrium. Let us begin with best response functions. Take as given everyone else's bids. Actually, only the highest of these bids matters. Let's call it $\bar{b}$. What is your best response to $\bar{b}$?

We need to distinguish two cases.

- If your own value v is below $\bar{b}$, then you don't want to win. You could win only if you were to bid above $\bar{b}$, and then you would have to pay $\bar{b}$, which is above v. Therefore, your best response is any bid b below $\bar{b}$. Any such bid makes sure that you lose. Note that one such bid is v itself. If you set your bid equal to your value, you'll lose in this case, and that is what you want.

- If your own value v is above $\bar{b}$, you want to win. For this, you have to make a bid above $\bar{b}$. Moreover, it does not actually matter how much you bid, as long as you bid above $\bar{b}$. That is because you don't pay your own bid, but you pay the second highest bid, which will be $\bar{b}$. Thus, any bid above $\bar{b}$ is optimal. Note that, among the optimal bids, is, in particular, your own value v in this case. If you set your bid equal to your own value, you'll win, and that is what you want.

Looking back over the two cases above, notice one interesting point: In both cases it turned out that your own value v was one of your optimal bids. In other words, you have a bid, namely $b = v$, that is optimal, independent of what the other bidders do. In a sense, you have a "constant" best response function: always bid v.

The conclusion is, of course, that it is a Nash equilibrium if everyone just bids their value. But it is more than that: bidding v is optimal for everyone, regardless of what the others do. We say: bidding one's true value is a "dominant" strategy in the second price auction.

If everyone bids their value, the bidder with the highest true value wins, and pays the second highest value. That is, if there are three bidders, with values \$1, \$2, and \$3, then the bidder with value \$3 will win the second price auction, and will pay \$2.

Let's go back to the question we asked earlier: why don't we make that bidder pay \$3 instead of \$2? She is willing to pay \$3. Her bid shows that. Wouldn't the seller be better off charging the winner \$3? The answer is that the seller would not be better off, and the reasoning that leads to this conclusion is simple but important. If the seller announced that the highest bidder wins, and that this bidder has to pay what she bid, then it would no longer be optimal for bidders to set their bids equal to their true value. If they did, they would always have utility zero, regardless of whether they won or lost. Bidders would lower their bids. We don't quite know how far, of course. But, just as in the second price auction, what the seller gets from the winner is not that bidders' true value. It is in most circumstances inevitable that the seller's revenue is less than the highest willingness to pay of the buyers. In fact, auction theory has identified assumptions under which the maximum that the seller can get is, on average, exactly the second highest value, and therefore the second price auction is indeed optimal for the seller.

Auction theory is a large field of research, with many more results. One can regard it as the "more sophisticated" version of the theory of the monopolist. We shall not pursue it here, but instead we shall move on to the case of two or more sellers. Moreover, we shall shift our focus away again from auctions, and towards more simple selling mechanisms, such as fixing a price.

Bertrand's Model of Duopoly

The general expression for market in which there are two, or more, sellers, each of whom can influence the market price, is oligopoly. Here, we focus on the case that there are just two sellers. A market in which there are only two sellers is called a duopoly. A "duopoly" is thus a special case of an "oligopoly."

Let's suppose that they are trying to sell products that actually, at least from consumers' point of view, are perfect substitutes of each other, i.e. exactly identical. We can then think, as in the previous section, that consumers' behavior is described by a demand function, $D(p)$, and that the price p that we need to plug into the brackets is simply the lower of the prices charged by the two firms.

For what follows the exact form of the cost functions of the two firms is important. Let us assume that both firms have constant marginal cost, and exactly the same marginal cost. Let us denote those costs by c. Thus, if a firm sets the lower of the two prices, its profits are:

$$pD(p) - cD(p),$$

which we can also write as:

$$(p - c)D(p).$$

If it sets the higher of the two prices, its profits are zero. What if the two firms set the same price? We shall assume that then half of consumers go to one firm, and half of the consumers go to the other firm, so that each firm has profit:

$$(p - c)\frac{D(p)}{2}.$$

Note that firms are involved in a game: their strategies are their prices, and the "utilities" are the profits. We want to find a Nash equilibrium of this game. If we proceed as we have done with earlier games, we would now look for the best response functions. But there is an easier way to reason. Consider the following two points.

- If the smaller of the two prices were below c, then the firm charging the lower price would make losses, and it could increase its profit by raising its price above the other firm's price, and then selling nothing.
- If the smaller of the two prices were above c, then the firm charging the higher price would make zero profits, and could increase its profit by lowering its price, moving the price between the lower price, and c.

These two points together imply that, in a Nash equilibrium, the lower of the two prices must equal the marginal cost c. What about the higher price? If the higher price is strictly higher than c, then this cannot be a Nash equilibrium, because then the firm that sets the lower price makes zero profits, but can raise its profit by charging a price above marginal cost, but still below the other firm's price.

The inevitable conclusion seems to be that the higher price is actually not higher, but equal to the lower price. Both firms set their price exactly equal to marginal cost. Denoting the two prices by p_1 and p_2, we thus conclude:

$$p_1 = p_2 = c.$$

But is this indeed a Nash equilibrium? Maybe the Bertrand model is one of the games that does not have any Nash equilibrium. We need to check. Can any firm raise its profit by changing its price? If a firm lowers its price below c, it will make losses. If it sets a price above c, it will lose its customers. It will make zero profits, just as it does if it charges the price c. Thus, it does not *gain* by raising its price, and therefore, both firms setting price equal to c, is indeed the only Nash equilibrium here.

There are two paradoxes here: The first is that, even though a firm does not lose if it raises its price, in the Nash equilibrium it must set its price exactly equal to c. If it were to raise its price, then the other firm could follow, and also raise its price, and we would again be in a non-equilibrium situation. The Nash equilibrium itself is in a sense unstable: nobody makes any profits. Nobody can gain by raising their price, but nobody can lose by raising their price either.

The other paradox is that the outcome is the same outcome that would also prevail under perfect competition. Even though we have considered only two firms, and these firms definitely have influence on the price, in equilibrium they make zero profits. The model can be extended to the case that there are more than two firms. In that case, too, in a Nash equilibrium, customers will end up paying a price exactly equal to marginal cost, and firms will make zero profits.

Cournot's Model of Duopoly

Now suppose that the two duopolists of the previous section decided in advance how much to produce, and only after they were done with production decided about prices. The game in which firms have given produced quantities, and choose prices at which they are willing to sell these quantities, is hard to analyze. It often has no Nash equilibria.

Instead of analyzing Nash equilibria of the price setting game, we shall take a short-cut. The short c-ut is that we shall assume that, in the price setting game, the two firms simply choose the price at which they can sell everything that they have produced. That is, if firms have produced quantities q_1 and q_2, then the price that they will choose is the price at which $q_1 + q_2 = D(p)$. Using the inverse demand function P that we introduced when discussing the monopoly problem, we can also say that:

$$p = P(q_1 + q_2).$$

Both firms choosing this price need not be a Nash equilibrium of the price setting game, and it typically isn't. The short-cut nonetheless has some justification: it is known that even if we conducted an analysis of Nash equilibria of the game, the analysis of the decision how much to produce would lead to the same result as we get here.

Now let's consider the game in which firms choose the quantities that they produce. In this firm, each firm's strategy is a quantity $q_i \geq 0$. The firms seek to maximize their profits which are given by:

$$\pi_i(q_1, q_2) = q_i P(q_1 + q_2) - C_i(q_i).$$

The first term on the right hand side of the above equation represents firm i's revenues, and the second term represents firm i's cost. Profit is revenue minus cost.The expression for profit reflects the short-cut: the price will be $P(q_1 + q_2)$.

We shall now consider Nash equilibria of this game. These Nash equilibria are also called "Cournot equilibria," and the game that we are considering is also called the "Cournot model."

To find Nash equilibria, we again consider best response functions. To make notation simple, let us focus on firm 1. Firm 1's best response to a given choice q_2 of firm 2 satisfies the following first order condition:

For simplicity, we ignore boundary solutions and second order conditions.

$$\frac{\partial \pi_1}{\partial q_1} = 0.$$

To calculate the partial derivative of profit with respect to q_1, we have to use the product rule. We get:

$$q_1 P'(q_1 + q_2) + P(q_1 + q_2) - C_1'(q_1) = 0.$$

As in the monopoly model, we can re-write this condition, to compare the value of marginal cost to the value of the price:

$$P(q_1 + q_2) + q_1 P'(q_1 + q_2) = C_1'(q_1).$$

We see that firm 1 expands production not to the point at which marginal cost equals price, but to the point at which marginal cost equals price adjusted by a term that reflects that firm 1 anticipates a price and revenue decrease as it expands production. As in the monopoly case, this means that the quantity produced is smaller than the quantity at which price equals marginal cost.

How is firm i's best response going to change as q_2 increases? Well, in the first order condition the first term is just the price, and the price goes down as q_2 increases. However, we do not know how the second term in the first order condition changes, because it involves the derivative P', and we have not made any assumption that tells us how the derivative P' depends on q_2. Let us make such an assumption, namely suppose that P is concave. Then we know that the derivative of P decreases as its argument increases, and hence the second term in the second order condition also decreases as q_2 increases. In this case, we can say unambiguously that the optimal choice of q_1 decreases as q_2 increases.

In a Nash equilibrium, both firms choose best responses simultaneously. We can find a Nash equilibrium in the Cournot model in the same way as we did in the "joint homework" example of Topic 20. We draw the two firms' best response functions in Figure 76.

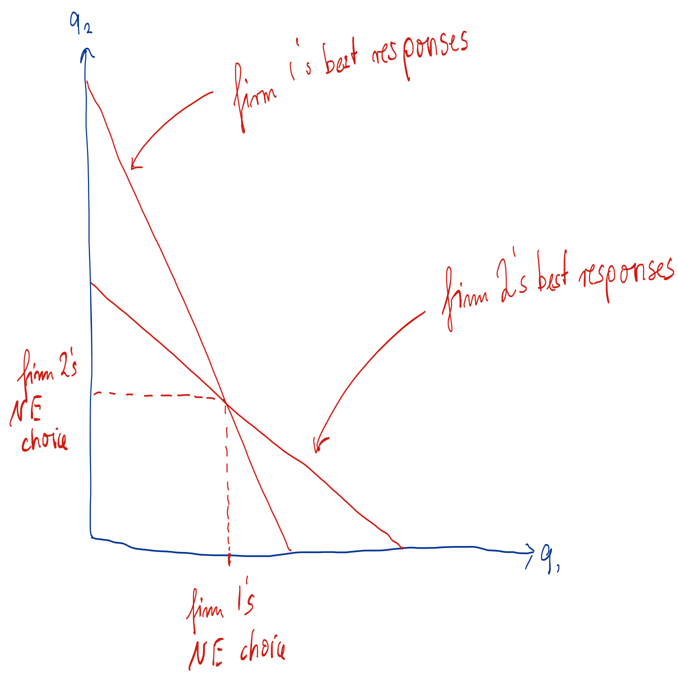


Figure 76: Best responses in the Cournot model.

What is the intuition that explains that the optimal q_2 creases as q_1 goes up? One might think it is this: the more firm 2 produces, the less "room" there is for firm 1 in the market. But the mathematical argument that we developed in the text shows that things are more complicated than this. The assumption that P is concave means that the larger q_2 is the larger is the price reduction that any increase in q_1 causes. Therefore, firm 1 chooses smaller quantities as q_1 increases.

If the best response functions happen to have the shape shown in Figure 76, then, as you can see in the figure, there is a unique Nash equilibrium. This equilibrium is also called the "Cournot equilibrium." But it is nothing else than the Nash equilibrium of the game that is called the "Cournot model."

But it is by no means guaranteed that there will be a unique Cournot equilibrium. Even if best responses are decreasing in the other firm's quantity, there may be multiple Cournot equilibria. I show an example in Figure 77 in which there are three Nash equilibria.

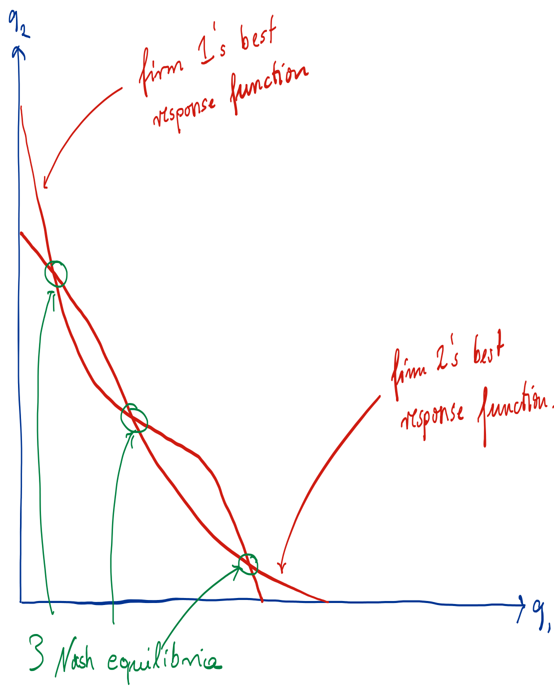


Figure 77: Best responses in the Cournot model.

A general analysis of Cournot equilibria is possible, but is perhaps a little too involved here. Instead, I shall consider in the next section a numerical example for which I can calculate a unique Cournot equilibrium. In fact, in that example, the best response functions are as shown in Figure 76. I calculate for that example also the monopoly solution, and the Bertrand equilibrium.

A Numerical Example

Suppose demand is given by: $D(p) = 12 - p$, and cost functions are given for all firms by: $C(q) = 2q$. Let us consider the monopoly model, and the Bertrand and Cournot models of oligopoly.

To study the monopoly model, let us first find inverse demand:

$$q = 12 - p \Leftrightarrow p = 12 - q,$$

so that the inverse demand function is:

$$P(q) = 12 - q.$$

Therefore, the firm's profit is:

$$\pi(q) = (12 - q)q - 2q,$$

The first order condition for profit maximization is:

$$\pi'(q) = 12 - 2q - 2 = 0 \Leftrightarrow 10 - 2q = 0 \Leftrightarrow q = 5.$$

The second derivative of profit is:

$$\pi''(q) = -2,$$

which is negative, and therefore we have found a maximum. To maximize profits a monopolist will choose to produce $q = 5$ units. The price will be: $P(5) = 12 - 5 = 7$. Note that the price is far above the marginal cost. Marginal cost are just: 2.

Now consider instead the Bertrand model, that is, assume that there are two firms with the same cost function $C(q) = 2q$, and that the two firms produce only *after* they have chosen their prices. Then the firms will both choose to set their prices equal to the marginal cost, that is, equal to 2. Consumers will only pay 2, and the total demand at this price is 10. Each firm will produce 5 units, but make profit zero.

Finally, let us consider the Cournot model. In this case, firm 1's best response function is obtained by maximizing:

$$\pi_1(q_1, q_2) = (12 - (q_1 + q_2))\, q_1 - 2q_1.$$

The first order condition is:

$$\frac{\partial \pi}{\partial q_1} = 12 - 2q_1 - q_2 - 2 = 0 \Leftrightarrow q_1 = 5 - \frac{1}{2}q_2.$$

The second derivative is -2, and hence we are dealing with a maximum. The same condition has to hold for firm 2:

$$q_2 = 5 - \frac{1}{2}q_1.$$

If we substitute this second equation into the first one, we get:

$$q_1 = 5 - \frac{1}{2}\left(5 - \frac{1}{2}q_1\right) = \frac{5}{2} + \frac{1}{4}q_1.$$

This is equivalent to:

$$\frac{3}{4}q_1 = \frac{5}{2} \Leftrightarrow q_1 = \frac{10}{3}.$$

Substituting back:

$$q_2 = 5 - \frac{1}{2} \cdot \frac{10}{3} = \frac{10}{3}.$$

Therefore, in the Cournot equilibrium both firms produce the quantity $\frac{10}{3}$, and then the price is

$$P\left(\frac{20}{3}\right) = 12 - \frac{20}{3} = \frac{16}{3}.$$

This price is lower than the price in the monopoly model, which was 7, but it is higher than the marginal cost, which are just 2.

It is interesting to ask how the Cournot equilibrium will change if there are more than two firms. Let us denote by n the number of firms. Focus again on best responses of firm 1. Firm 1's profit is:

$$\pi_1(q_1, q_2, \ldots, q_n) = (12 - (q_1 + q_2 + \ldots q_n))\, q_1 - 2q_1.$$

The first order condition for a profit maximizing choice of q_1 is:

$$\frac{\partial \pi_1}{\partial q_1} = 12 - 2q_1 - q_2 - \ldots - q_n - 2 = 0 \Leftrightarrow q_1 = \frac{1}{2}\left(10 - q_2 - \ldots - q_n\right).$$

This has to hold for all firms, for example for firm 2:

$$q_2 = \frac{1}{2}\left(10 - q_1 - q_3 - \ldots - q_n\right).$$

Suppose we subtract the first order condition for q_2 from the first order condition for q_1. Then we get:

$$q_1 - q_2 = \frac{1}{2}(q_1 - q_2),$$

which is equivalent to:

$$\frac{1}{2}(q_1 - q_2) = 0,$$

and this in turn can only be true if:

$$q_1 = q_2.$$

Thus, firms 1 and 2 have to produce the same quantity. With the same argument one can show that all firms have to produce the same quantity. Let us denote that quantity by q^*. The we can write each individual firm's first order condition as:

$$q^* = \frac{1}{2}(10 - q^* - \ldots - q^*),$$

where there are $(n-1)$ q^*s in the bracket.

The first order condition is therefore:

$$q^* = \frac{1}{2}(10 - (n-1)q^*).$$

Multiplying both sides by 2 we get:

$$\begin{aligned} 2q^* &= 10 - (n-1)q^* \Leftrightarrow \\ (n+1)q^* &= 10 \Leftrightarrow \\ q^* &= \frac{10}{n+1}. \end{aligned}$$

In the n-firm Cournot equilibrium therefore each firm produces the quantity $\frac{10}{n+1}$. The market price is then:

$$p = 12 - nq^* \quad \Leftrightarrow \quad p = 12 - n\frac{10}{n+1} \quad \Leftrightarrow \quad p = 12 - \frac{n}{n+1}10.$$

Notice that, when we plug in $n = 1$, we get in fact the monopoly price: $p =$

$12 - \frac{1}{2}10 = 7$, when we plug in $n = 2$, we get the price for two firms that we calculated earlier: $p = 12 - \frac{2}{3}10 = \frac{16}{3}$, and if we consider what happens as n, the number of firms, gets very large, the ratio $n/(n+1)$ converges to 1, and the price converges to $p = 12 - 10 = 2$, which means that with many firms the price is almost equal to marginal cost. That is, of course, also the price that the Bertrand model predicts.

Topic 23: Adverse Selection

In the previous topic we considered imperfect competition as a reason why markets may differ from the simple “supply=demand model.” In this topic we consider another reason: asymmetric information. What we mean by this is that some people have information that other people don’t. One says in this case that some people have “private information,” i.e. it is private to them - others don’t know it. We shall show how asymmetric information affects how the market works, and that the simple supply=demand model may not be quite sufficient.

What sort of information is it that some people know, and others not? As an example, suppose that we trade apples, and that I am a buyer of apples. It may be that my I know how much I value certain types of apples, and therefore how much I am willing to pay for them, and that sellers don’t. That is one type of asymmetric information. Another possibility is that I know that the harvest of apples this year was particularly bad, and that they are all infested by worms. This is different from the first case because the worm infestation is something that does not only affect me but that also affects others, maybe other buyers, or maybe even the sellers, if they don’t know either. This is a second type of asymmetric information. We shall call the first type of asymmetric information “asymmetric private value information,” and the second type of asymmetric information “asymmetric common value information.”

Putting things more generally: “private value information” is information that a market participant holds and that only affects that market participant’s valuation of good traded in the market. “Common value information” is information that a market participant has, and that affects other market particiapnt’s valuation of the good traded in the market as well. Putting this last point differently: if you have “common value information,” and you told me what you know, then my valuation for the good traded in the market would change. “Private value” and “common value” information are terms that are used in the literature on microeconomic theory with a more technical meaning than I have introduced here. Thus, if you read more advanced literature, you may have to get used to a somewhat more narrow definition of these terms. The way I use the terms here is, however, “in the spirit” of the more advanced literature.

If buyers and sellers in a market have “asymmetric private value information,” then this might not make a difference to the working of the market. Indeed, if our model of supply and demand were only valid if everyone knew everyone else's preferences, then it would be based on extremely narrow assumptions. The model is not meant that way. It is supposed to describe "decentralized" process of making decisions about the use of economic resources, and, in particular, it is meant to allow for the possibility that everyne knows only their onw utility function or their own production function, but not that of others. Thus, we shall not modify our supply=demand model to capture asymmetric private value information, but hope that the model also applies in this case.

Why do I say "hope" rather than asserting it applies? Because ultimately this is a question about the real world, and here, I don't give any evidence about the real world.

The second type of asymmetric information, "asymmetric common value infomration," creates problems for our supply=demand model. These problems are our subject in this topic.

Akerlof's Lemons Model

We shall only consider one model of asymmetric common value information, namely the model that is known as the "lemons model." It is due to George Akerlof, who described the model in a paper published in 1970. Akerlof asked in that paper a specific, and very practical question: "Why are 1 year old cars so much cheaper than new cars?" The answer that he suggested was this. Buyers who buy a new car know reasonably well what they are getting (except for the possibility of factory flaws). By contrast, buyers who buy a car after 1 year don't know whether the seller is selling the car because it has some problem, or because she doesn't need it anymore for some other reason.

I don't know whether that is actually true today. I am not even sure that it was true in 1970.

This problem of lack of information is then exacerbated by the following effect: Owners with negative information are more likely to sell than owners with positive information. Those owners who discover their car is good, rarely sell it. Those owners who discover their car is bad, frequently want to sell the car. Therefore, 1 year old cars *that are sold* are more frequently bad than new cars. Buyers anticipate that. That's why you can't sell a one year old car at a high price.

If you buy a used car, and it turns out to be poor quality, one also says you have bought a "lemon." That is why markets that are like the used car market that Akerlof described, are also called "markets for lemons."

Note that the form of asymmetric information that Akerlof described is, in our language, "asymmetric common value information." The knowledge that sellers have about their car is not just relevant to them, namely to their willingness to sell the car, it is also relevant to the buyer. Buyers want to avoid lemons. In other words: the quality of a car is not just a matter of private taste, but it is a matter that affects everyone's preferences.

Let us write a model that reflects Akerlof's ideas. We'll use very specific, and very simple numbers. Suppose there are 4000 buyers. Each buyer wants to buy one car. All buyer's willingness to pay for a good car is \$2000. But even if the car is bad, a buyer is still interested in buying it, but the buyer's willingness to pay for a lemon is only \$1000. All buyers are the same, and all buyers are risk neutral.

Sometimes students ask me: where do those numbers come from? The answer is: I made them up. I chose them so as to make my calculations easy.

There are 1000 car owners, i.e. potential sellers, who know that they have a lemon. They discovered it during the time that they owned their car. They want to sell as long as they get at least \$750 for their cars. There are also 1000 owners of good cars. They are more reluctant to sell than owners of lemons. They are willing to sell their cars only if they get at least \$1750.

Note how simple the model is. In particular, all buyers are identical, all sellers of bad cars are identical, and all sellers of good cars are identical.

We will be interested in the case of asymmetric information, when buyers don't know which cars are good, and which are lemons. But first, as a benchmark, let's ask what would happen if there were complete information. There will then be two markets, one for lemons, and one for good cars. In Figure 78 I have drawn supply and demand in both of these markets. I'll explain in the next three paragraphs how I have constructed the supply and demand curves.

In both markets, it is easy to understand where the supply curve comes from. Consider first the market for lemons. Supply of lemons is zero if the price is below the price that the sellers want, i.e. \$750. Supply is equal to the total number of owners of lemons, i.e. 1000, once the price rises above \$750. At price \$750, the owners of lemons are indifferent between selling and not selling. That is why for that price I have drawn a horizontal supply curve. In the market for good cars, the supply curve is constructed in the same way.

Now consider demand. In the market for lemons, if the price is above \$1000, no buyer is interested. If the price is equal to \$1000, buyers are indifferent between buying and not buying. Therefore, the demand curve is drawn at this

Market for Lemons:

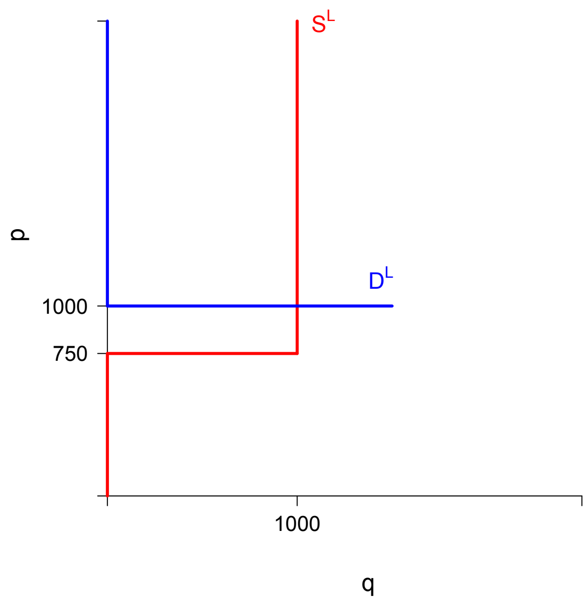
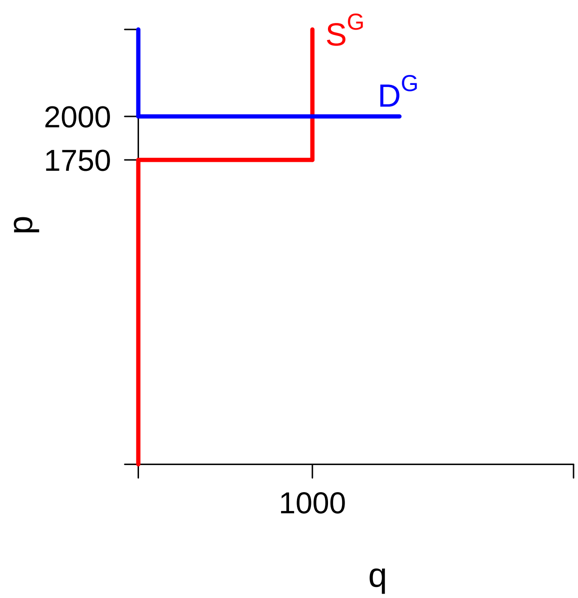


Figure 78: Supply and Demand with Complete Information

Market for Good Cars:



price as a horizontal line. For prices below $1000, buyers are potentially interested in buying, but whether or not they actually do buy depends on what the price in the other market is, i.e. which market offers the better buying opportunity. That is why I have not drawn the bottom part of the demand curve for prices below $1000. The demand curve in the market for good cars is constructed in an analogous way, and again I have left out the bottom part of the demand curve.

Why have I left out the bottom parts of the demand curves? It is because the two markets are connected. Whereas a seller either sells a lemon or a good car, a buyer considers buying in both markets. Therefore, I can draw the demand in one market only conditional on the equilibrium price in the other market. But we are about to determine the equilibrium prices. Therefore, I cannot yet draw the bottom parts of the demand curves. I can draw the top parts because they are in each market independent of what happens in the other market.

Even though I have drawn only incomplete demand curves for both markets, we can see in Figure 78 what the equilibrium prices will be. Lemons will be sold at the price $1000, and good cars will be sold at price $2000. Thus, the prices equal exactly the buyers willingness to pay. Buyers don't get a surplus in either market. They are indifferent between buying good cars, or lemons. Sellers, of course, make a good profit. In both markets 1000 cars will be sold. There are therefore 2000 buyers who don't get a car. But they are not dissatisfied, because the only price at which they could have got a car is exactly equal to their willingness to pay. So, either way, they have utility zero.

I have not gone in detail through the argument for this being the only equilibrium of supply and demand. But you can think of this case like we thought earlier about markets for perfect substitutes. In this model, lemons, offered at price p, are a perfect substitute for good cars, offered at price p+1000. In equilibrium, buyers must be indifferent between the two markets, otherwise demand would be zero and supply would be positive in that market. Moreover, because there are many more buyers than sellers, in equilibrium the buyers must be indifferent between buying and not buying. Some buyers will not get anything. This can happen in equilibrium only if they don't strictly prefer to get a car. These indifference conditions pin down the equilibrium prices in the case of complete information.

Now let us consider the same market with asymmetric information. Buyers don't know whether a car that they are offered is of good or bad quality. Thus, there is asymmetric information of the "common" value type: sellers have information that would affect buyers' willingness to pay if they knew it. There can then no longer be two separate markets, one for good cars, and one for lemons. Buyers don't know which car is which. There is only one market, a market for "cars."

Let us first construct a supply function. This is a good place to start because the supply side, the sellers, don't have incomplete information. They know the quality of their cars. I show in Figure 79 the supply function for cars.

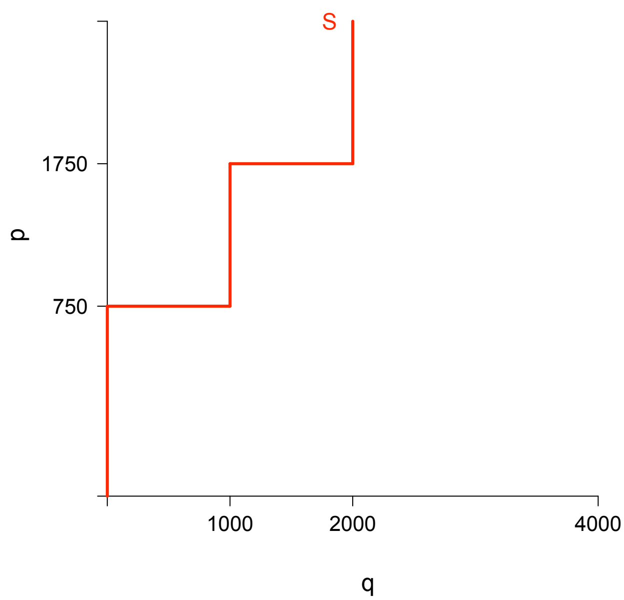


Figure 79: Supply With Asymmetric Information

The function has two steps. At price 750, supply increases from zero to 1000. That is because, at that and higher prices, the 1000 owners of lemons will be willing to sell their cars. When the price rises to 1750, or above, then also the owners of good cars will be willing to sell their cars. Therefore, at those prices, supply is 2000, the total number of cars in the market.

What becomes interesting now is the construction of the demand function. Recall that buyers don't know the quality of any particular car. They have to determine the “average” quality of a car offered in the market. But buyers will learn. In the long run, at any price p, buyers anticipate the “average" quality of cars offered at that price. The comparison between the value of this average quality and the price will determine whether buyers are willing to buy.

Based on the supply curve, we have four cases to distinguish:

- $p \geq 1750$: the owners of good and bad cars are willing to sell. Because there are equally many good and bad cars, buyers will anticipate that half of the cars are good and half are bad. Their their expected utility will therefore be:

$$0.5 \cdot 1000 + 0.5 \cdot 2000 = 1500.$$

  But, because this value is below the price, which is at least 1750, demand will be zero.

- $1000 < p < 1750$: only the owners of bad cars are willing to sell at this price. Buyers will anticipate that all cars that are traded are bad. The buyers’ willingness to pay is therefore just \$1000. But because the price is above \$1000, demand is still zero.

- $p = 1000$: also at this price, only the owners of bad cars are willing to sell. The buyers’ willingness to pay is \$1000. Therefore, the buyers are indifferent between buying and not buying. The sellers of the bad cars do want to sell, however. The supply is therefore 1000. All these cars will be sold. 3000 buyers will go empty handed. But that is OK. Buyers are indifferent between buying and not buying. The market price is exactly their willingness to pay.

- $750 \leq p < 1000$: Again, only the owners of bad cars are willing to sell. But now the price is strictly below buyers’ willingness to pay, and therefore all 4000 buyers want to buy. Demand at this price is 4000.

- $p < 750$: No car owners are willing to sell at this price. This makes it unclear what buyers are supposed to think about the quality of a car, if it nonetheless is offered at this price. But, in any case, buyers will be eager to buy. Demand is 4000.

This sequence of steps leads to the demand curve that is shown in Figure 80.

Let’s look, as always, for a point of intersection of supply and demand. The two curves intersect at the price $p = 1000$. This is the only equilibrium price.

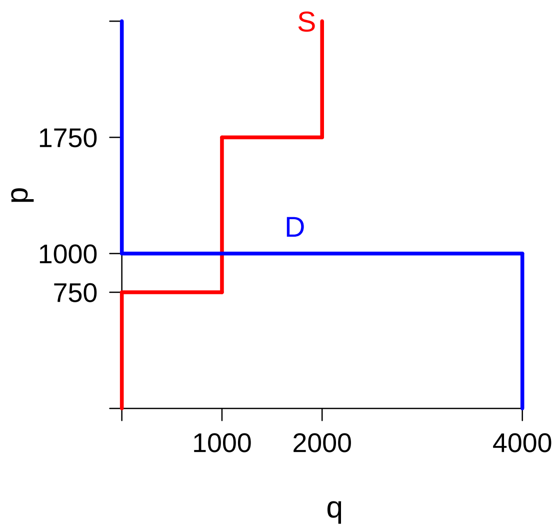


Figure 80: Supply and Demand with Asymmetric Information

At this price, only bad cars are traded. Buyers are indifferent between buying, and not buying, a car. They understand that only lemons are in the market. Thus, supply is 1000, and demand is anywhere between 0 and 4000. But in equilibrium, 1000 buyers buy a car, and the remaining 3000 buyers don't buy a car, but they don't mind, because they know they would get a bad car, and the price is exactly equal to their willingness to pay.

This is the bottom line: only bad cars can be sold. Good cars are not sold. This is inefficient, because buyers are willing to pay for good cars more than the owners of good cars require. But there is no way in this market for owners of good cars to distinguish themselves from owners of bad cars. Therefore, they will not be able to achieve a price at which they want to sell.

In practice, of course, owners of good cars can prove that they have good cars. They can pay a mechanic of the buyer's choice to inspect the car. Or maybe they offer a very good insurance. The fact that we see these practices in the real world can be interpreted as a response to the problem of asymmetric information. So, maybe it is best to think of the lemons model as the benchmark model that describes what would happen if there was no way of proving the quality of what you have. Moreover, in some markets it may be harder than in others to prove the quality of what you sell. Whenever I have bought a house, for example, there has always been a doubt in the back of my mind that the house had some serious problem that neither I, nor the house inspector, had found.

Now compare the case of asymmetric information to the case of complete information: In the case of complete information, not only the lemons, but also the good cars, were traded in the market. This was because, for either car, the buyers' willingness to pay was above the seller's reservation prices. Trade, therefore, was also efficient. Thus, we have indeed found an inefficiency in the market that is caused by asymmetric information.

One way of thinking at this market is as follows: If the price is very high, then there is large supply. As the price drops, one would expect demand and supply to meet. But there is another effect: as the price falls, the composition of the supply side gets worse: the good cars drop out of the market, and only the bad cars are sold. This is called *adverse selection.* Its effect is that, as the price falls, because buyers understand that the average quality of cars gets worse, also demand falls. This can be a downward spiral. In some models, ultimately, no cars are traded at all. Everyone expects that a car that enters into the market is terrible. This is possible even in cases in which, in principle, if buyers knew the quality of the cars, for every car there would be a buyer. Thus, adverse selection may cause extreme forms of market inefficiency.

Adverse selection may also occur on the demand side. For example, it may be that health insurance buyers, but not health insurance companies, know whether the buyer is healthy, or likely to get ill soon. Then, when the price of insurance is low, there is a lot of demand. As the price goes up, however, people who are healthy decide that it is no longer worthwhile to pay the price. Only people who are likely to get ill, and therefore people who are "lemons" from the health insurance companies point of view will buy expensive insurance. Charging a high price may not always mean higher returns per customer for the insurance company.

Sorry for the discriminatory word.

Adverse selection is, in fact, pervasive in markets. There are many other interesting models that deal with adverse selection, beyond the one explained in this topic. You will learn about those models at the latest in a future life.

Index