Methods and Techniques for Social and Economic Research: Syllabus

Paul Koster & Thomas de Graaff

2022-10-29

Contents

Preface **5**
What . . . 5
Why . . . 5
For Whom . . . 6

1 Introduction **7**
1.1 Theory, Models and Hypotheses . . . 7
1.2 Doing Research (in the Social Sciences) . . . 10
1.3 Statistical software . . . 13
1.4 Reading Guide . . . 13

2 Introduction to Economic Surplus **15**
2.1 Introduction . . . 15
2.2 Consumer choices and consumer value . . . 17
2.3 Poducer behaviour and surplus . . . 32
2.4 Analysis of economic surplus . . . 45
2.5 External costs and economic surplus . . . 49
2.6 Conclusion . . . 58

3 Behavioural Error and Economic Surplus **59**
3.1 Introduction . . . 59
3.2 Arguments against neo-classical valuation and responses . . . 62
3.3 Behavioural choice models with errors . . . 69
3.4 Behavioural errors and economic social surplus . . . 76
3.5 Behavioural error and policy recommendations . . . 85
3.6 Behavioral errors and pricing of consumption externalities . . . 89
3.7 Discussion and conclusion . . . 93

4 Moral Considerations and Economic Surplus **95**
4.1 Introduction . . . 95
4.2 Dealing with moral considerations at the valuation stage . . . 97
4.3 Redefining the economic pie . . . 103
4.4 Taxation and moral consideration . . . 111

4.5 Externality taxation and moral considerations 112
4.6 Conclusion and discussion 114

5 Regression Analysis in the Social Sciences 117
5.1 Introduction 117
5.2 So, what is the problem? 118
5.3 Univariate regression 125
5.4 Least squares assumptions for causal inference 139
5.5 Other least squares assumptions 141
5.6 Measures of fit 147
5.7 Conclusion and discussion 149

6 Modeling in the Social Sciences 151
6.1 Why more independent variables? 151
6.2 Multivariate regression analysis 156
6.3 Non-linear specifications 165
6.4 Using fixed effects in panel data 182
6.5 Conclusion and discussion 184

7 Specification and Assessment Issues 185
7.1 Specification of your model 185
7.2 Presentation of results 187
7.3 Potential sources of bias 190
7.4 Concluding remarks 195

8 In conclusion 197

A Derivation of the optimal cost function 201

B Taxation in the presence of behavioural error 205

C Background calculations for Chapter 4 207
C.1 Deriving the tax that optimizes adjusted social surplus 207
C.2 Pricing of a consumption externality 208

D Reviewing probability and statistics 209
D.1 Reviewing probability 209
D.2 Sampling in frequentist statistics 213

Preface

What

This syllabus provides the course material for the course **Methods and Techiques for Social and Economic Research**. With this course we would like to bridge the gap between on the one hand applied statistics and (micro-economic) modeling and on the other hand putting all this in practice when performing empirical research. As such this course can as well be seen as preparation for the bachelor thesis. But above all, the course aims to provide students with some tools that we see as very useful for research; not only in the socio-economic sciences but outside them as well.

As we only have a limited amount of time available for this course, the amount of topics we can deal with is by nature restricted. We decided to focus in the first three week on economic welfare from a behavioral perspective. Not only is the concept of economic welfare central to all most all economic theories, the behavioral perspective allows for more *reflection* on the neo-classical assumptions typically made in introductionary economic courses. The last three parts of the course move on to empirical analysis by focusing on the basics of applied econometrics and as such builds upon the foundations of the statistics course in the first period of the second year. But now we challenge the student to build more elaborate statistical models (based partly upon the theory of the first three weeks of this course) where specific attention is given to *presentation* and *interpretation* of the results.

Why

Although there are many and very good introductory textbooks on economic models and applied econometrics, the combination of the two is seldom seen. Apart from that there are two reasons why we wanted to write our own material. First, usually less time is spent on why certain, and a first sight very restrictive, *assumptions* are made. We want to bridge that gap and provide the student with more intuition on where models, evidence, and finally perhaps the "truth" (if there is such a thing) comes from. Second, how to *present* statistical evidence

and the *interpretation* of that evidence is very important but usually not given much attention.

For Whom

This syllabus assumes that the reader has a basic working knowledge of introductory economics, statistics and some calculus (typically those method courses students enjoy in their first year). The book can however be read as stand-alone, although that requires some more attentive reading and practising. Where we think it is necessary we provide (references to) background material. For the course **Methods and Techiques for Social and Economic Research** the syllabus should be read **in total** and it might be wise to read the relevant material *before* the respective lecture. The big advantage of course is that lectures and reading material now really go one-on-one.

Chapter 1

Introduction

In this first introductory chapter, we will lay out the relations between theory development and theory testing as they are the cornerstones of scientific progress. After all, a theory or idea can only be scientific if the theory can be tested and, if need be, refuted. If the theory cannot be tested then it is not science. We will also explain the basic workflow of scientific research and the tools needed with specific emphasis on research in the social sciences. We end with a reading guide where we discuss each chapter in this syllabus and the relations between the chapters.

1.1 Theory, Models and Hypotheses

In 2021, Guido Imbens, Joshua Angrist and David Card received the Nobel price for economics (officially *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel*). The field they work in is applied econometrics with specific focus on finding **causal** relations. That means that with data they want to test whether phenomenon X has an effect on phenomenon Y. And that is what most applied econometric work nowadays is focused on. And we will focus on that as well in Chapter 5, 6 and 7. But how do you know what to test, or, in other words, where do phenomena X and Y come from? Those phenomena and possible relations which originate from scientific theories as you will have in all disciplines. And those theories are typically casts in models—usually in a very abstract manner. Models come in the form of computer simulations (such as with agent-based modeling), real physical models (as with displays), but often models are formulated in mathematical notation with the aim of being as precise, lucid and clear as possible. Chapter 2, 3 and 4 will indeed deal with a specific subset of economic models formulated with mathematical notation. But note that those models are not necessarily theory. Theory is the underlying set of relations and assumptions that can say something about the specific structure of models. But very often theory can lead to **multiple** types of models, each

perhaps highlighting different aspects of the underlying theory. An example of such a theory is the Law of Diminishing Marginal Utility: each additional unit of the same good is appreciated less by consumers. This theory can be expressed in many mathematical ways but the underlying concept as displayed in Figure 1.1 always remains the same. This type of function, increasing but slower and slower, belongs to the family of *concave* functions. A function with exponential growth (such as e^{gt} belongs to the family of *convex* functions).

Figure 1.1: Law of Diminishing Marginal Utility

So how to relate this with each other in scientific research? Well, when doing research you are interested in something that is not yet known (the research gap). Your aim is to (partly) fill this research gap by answering a research question. To answer this research question you need theory (a theoretical framework); what do you need to assume, what are the most important (moderator) variables, how do they relate with each other, and so on and so forth. From this theory you construct a model. Not necessarily a mathematical one. For example, you can also make a model in a Geographical Information System environment where you

visualize layers of information that you think are most relevant based upon theory (in this case often previous scientific literature). Or you make a simulation model examining risks of flooding by rivers. The final step is the stage where your model should provide you with some answers. Sometimes they are concerned with optimality (what is the best location of a new road in a GIS environment), prediction (where are river dikes most vulnerable), or with establishing a (causal) relation. And it is the latter that the second part of this course deals with. How can we know that there is a relation between phenomenon X and phenomenon Y and how do we know whether that relation is causal?

For that we use applied econometrics (which is a form of applied statistics but then in the social-economic sciences domain—the exact difference will be discussed in Chapter 5). And to establish a, hopefully causal, relation, we test our models with empirical data. Be aware, though, that the applied econometrics materials we teach in this course (and in all introductory courses of applied statistics and econometrics all over the world—the "101" courses) is based on so-called frequentist statistics (we will revisit its basic assumptions in Chapter 5). The exact definition is not important for now but know that it is intrinsically related with hypothesis testing.

And hypothesis testing is most often associated with that scientific philosopher—and perhaps the only one you know—Karl Popper (as displayed in 1.2).

Figure 1.2: Karl Popper

Popper was a so-called empiricist and claimed that theories in the empirical sciences (that includes most of the social sciences) can never be proven, only rejected. That is why you can reject a null-hypothesis (H_0), but **never** accept the alternative hypothesis (H_a). And this is very much related with the theoretical framework behind frequentist statistics. Loosely speaking, in frequentist statistics you construct a world where H_0 is true and try to reject that world with data (we will come back to this in Chapter 5)—but that world does not say something the validity of the alternative hypothesis. Now, Popper never claimed that rejecting one null-hypothesis will reject a whole theory. For that you need a larger body of evidence, including results from all sorts of studies—not only statistical ones, leading to a *general* consensus among the scientific community (Popper 2005).

In truth, although the scientific approach of working with null-hypotheses is a very valuable one (and remarkably practical), it does not lead to definitive answers. That is because contradicting models can lead to similar null-hypotheses. Moreover, often the direct connection between research question and null-hypothesis is not a direct one. Consider that you want to know the effect of X on Y, so your research question is: "What is the effect of X on Y?". But, in a frequentist world you only reject hypotheses—thus leading to results in the line of: "The effect of X on Y is *not* ...". This makes the evidence for your research question at least circumstantial and in the best cases indirect. The bottom-line here is that one needs to be careful in drawing conclusions based on null-hypotheses (and in a broader sense based on models in general). Scientific research typically advances very slowly—but hopefully in a robust and parsimonious way!

1.2 Doing Research (in the Social Sciences)

It is remarkable that, although in principle students are (should be) prepared for scientific research, they receive little guidance in *how* they should do scientific research. What are the tips & tricks of the trade and what—and, more importantly why—should you use with respect to specific (types of) applications and what is the relation between them. In our view most of this should belong in your first course upon entering the university (with the appropriate course title "Research Methods 101"). And some of it you indeed have learned in your first year, but in our experience students still lack "operational" knowledge. Therefore, we discuss below the four elements we think are among the most important—at least for this course. There are others, but for now this will do.

1.2.1 Work tidy

Our first and most important tip is to work tidy. Try to make your work look **good**. And with work we mean everything you submit (such as tutorials, papers, examinations, and theses). And that is because lecturers are just like people and often think from primary instincts with their reptilian brain: if it doesn't look good, not much time is spent on arguing and thinking as well! Moreover, when your work is difficult to read, lecturers get annoyed. Making your work

look good and in the same time more lucid and transparent also serves a higher purpose as it is then easier to detect mistakes. Namely, everyone makes mistakes. The important thing is to detect them early, learn from them and remedy them. This advances science in general and is a very important feature of the scientific process. Chapter 7 will spent additional time on working tidy and making it looking good.

1.2.2 Know where your stuff is

A second very straightforward tip is to be organised and to know where your stuff is. Often, students come to us for help with all their files piled on a stack on their desktop and facing difficulties finding where their work is. It is always advisable to use a folder structure and have one folder for one project (or for one course). And to the use sub-folders for data, text, code, pdf's and so forth. A second tip for organisation is to think about versioning. As the well-known Figure 1.3 shows the number of versions of one file very quickly can get out of hand. Think at least about a consistent naming structure (perhaps with the date involved such as `paper_mt_20221215.doc`).

1.2.3 Make notes

One skill that in our opinion is given not too much attention is making *useful* notes. It has been proven that writing things down is beneficial; not only for remembering but also for understanding. And that seems to be best just by using a pen as this slows writing down and you have to think what to write down. Underlining or marking is useful less beneficial than writing accompanying notes. But when should you write notes? Well, when attending lectures of course but also when reading. To leverage your notes as much as possible it is important that you have a system where you can *retrieve* your notes and compare them with other notes. The latter is the hardest part, but is in the long run the most rewarding as new connections are created between lectures, courses, books, and years. You do not need any fancy tools for this (there is literally a ton of applications to be found on internet), Microsoft's Onenote or Evernote are more than good enough. Where the workflow typically is to first use pen and paper to *capture* notes and thereafter rewrite and organise your notes in a notes system.

1.2.4 Use a reference manager!

Perhaps the tool that has the quickest pay-off is a reference manager. For those of you who are not using one yet: **do** it. Why? Because you never have to think about your reference list again. All reference managers come with plugins for Word or other text-editors (or type-setters such as LaTeX) that enable you to *automatically* generate reference list based upon in-text citations which the reference manager can also provide. You only need less than an hour to set it up, but you very quickly become more efficient (and thus *save time* in future work). There are many reference managers out there, but we advise Zotero as it

Figure 1.3: Version confusion

is open source. There is both a cloud and desktop version and it comes with a handy tutorial. It also provides a plugin for your browser to automatically import the bibliographic details of the paper you are reading at the moment.

1.3 Statistical software

As quantitative research becomes more and more important in the social sciences you need software to **manage** your data and provides statistical and applied econometric **analyses**. Ideally, an open-source package is used (such as `R` or `Python`), but they have a steep-learning curve and do not work immediately out-of-the-box. The statistical software we use in this course is `STATA` and is more intuitive (compared to `R` or `Python` that is!) and, above all, all economists use it. So, the user base is large and that is important, because for each problem there is much material to be found on internet (including videos). Be aware though that there is one disadvantage in using `STATA` and that is that it is not open-source.

1.4 Reading Guide

This course will not concern itself with theory as such, but more with methods to describe that theory (the models) and how to test that theory (the applied econometrics). Chapters 2—4 deal with economic model. First, Chapter 2 revisits basic economic theory again with an emphasis on measuring (economic) welfare. Subsequently, Chapter 3 challenges this framework and especially the assumption that all actors make complete rational decisions (the *homo economicus* assumptions). What happens when you introduce behavioral errors—such at not enough time to consider all choices—in the model? Chapter 4 then introduces moral considerations—based on rather recent developments—to the model and starts to ask normative questions (how should welfare be distributed?).

Chapter 5 constitutes a break from theoretical modeling to empirical modeling and introduces the basic concepts of applied econometrics in the form of univariate regression. Chapter 6 extends this framework to a multivariate regression setting, but in the same deals as well with the translation of theoretical (socio-economic) models to empirical models that are testable. Chapter 7 discusses how to *specify* your model—which variables should you include and which variables not—and how to present your findings to a wider audience (that includes assessors). The final chapter summarizes and provides a general discussion.

Chapter 2

Introduction to Economic Surplus

Practical cost-benefit analysis is widely used in applied economic analysis and often employs a definition of economic surplus that adds up consumer and producer surplus and potential external costs. This chapter discusses the behavioral assumptions underlying this calculation and show how to operationalize the definition of economic surplus to derive economic surplus changes and externality taxes. It employs a mathematical approach suitable applied to economic partial equilibrium analysis.

2.1 Introduction

2.1.1 Background

This chapter introduces the classical economic perspective on economic surplus. Economic surplus is a quantitative estimate of the value that markets generate for society according to economists. Besides analysing the impacts of economic policies for societies (do policies have effects and do they work as intended?), one of the core tasks of economists is to give recommendations about what is supposed to be valuable for society (do policies add value?).

Economic analyses of the value that markets generate also are important in debates about sustainability. What economists contribute to this debate is that they seek to analyse behavioural choices of consumers and producers using statistical techniques and mathematical modelling and that they come up with quantitative estimates of value. The contribution of this chapter is that we move from conceptual analysis of market figures to mathematical descriptions and connections to statistical measurement. We do so for partial equilibrium analysis, meaning that we focus on the analysis of a market for a single good

(with multiple good, issues of substitution and complementarity comes into play).

Four conceptual pillars of economic analysis are important for applied economic analysis of surplus and this chapter aims to emphasize the structural relation between these pillars along the road:

i. **Optimization**: market actors choose their best (feasible) option.
ii. **Equilibrium**: market actors choose their best feasible option when interacting with others.
iii. **Empiricism**: behavioural models need to be tested with data.
iv. **Choice-based normativity**: choices express value of market actors and value of market actors is equal to value for society.

The first core principle states that consumers and producers optimize when they make their choices. For consumers this implies it is assumed that they maximize their utility. A mathematical description of utility is captured by the direct utility function which maps consumption quantities and remaining money to some cardinal number. We will discuss the mathematical behavioural model underlying consumer optimization in Section 2.2 and discuss the implications for consumer value in markets. We also show at some points how this model of consumer behaviour relates to pillar (iii).

For producers, pillar (i) implies that they minimize their costs given a production target when they are assumed to be price-takers. Section 2.3 discusses the basic cost-minimization model of producers and the implications for producer value in markets.

The second core principle states that equilibrium in markets results from the behavioral choices and *interaction* of consumers and producers. Section 2.4 of these lecture notes discuss partial equilibrium and implications for total economic value in markets. The discussion in these sections of the lecture notes is conceptually speaking rather basic as substitution to other markets is ignored. Some complexity is added by connecting equilibrium to the third pillar of empiricism. It is very important for you to understand the core concepts and assumptions and the calibration underlying a basic partial equilibrium model in order to understand the extensions that are provided in the weeks that follow.

The third pillar is discussed along the way. Empirical analysis can be used to estimate the structural parameters of observed demand functions and production functions using regression techniques. These techniques relate mathematical narratives of choice behaviour to empirical data. The mathematical structure that is imposed sometimes requires that parameters in regression have particular values. Hypothesis testing can be used to see which model fits the data better. Calibration can be used to connect estimated elasticities from the literature to market demand and supply functions. Examples are provided to show how you can obtain quantitative estimates of total economic value.

Section 2.5 applies the model of market value and introduces the idea of external costs. This is a key concept in environmental economics and analyses of sus-

tainability. From an economic perspective, external costs resulting from market trade lower the value of a market and can be a motivating factor for regulatory intervention of a government. For a classical economist a sustainable market is a market where external costs are correctly priced. Section 2.5 discusses the regulation of an environmental externality in a market and thereby provides the necessary bridge between the economic theory and your bachelor program. Analytical examples are provided to show how you can obtain quantitative estimates of total economic value.

We will come back to the four pillars in the chapters that follow where we zoom in on the impact of bounded rationality and other views on normativity.

2.2 Consumer choices and consumer value

2.2.1 Utility functions, inverse demand and demand

Why do consumers choose to consume particular meat products? What is the consumer value of newly developed vegan product? How does households' demand for solar panels respond to price decreases? To study these research questions, it is helpful to develop a model of consumer choice. For economists, models of choice are mathematical narratives about how consumers make their choices. The story of consumer choice by economists assumes that consumers optimize the choice value that they derive from the consumption of market goods given constraints they face. In this section the consumer story is partly simplified and we investigate the decision of the consumer to spend money on the market good versus spending money on other goods or services given a budget constraint. We therefore do not study market goods that are close complements or close substitutes.

Value is mathematically conceptualized using a direct utility function that depends on the consumption levels. The direct utility function of the consumer is defined by $U(Q, G)$, where Q is the demand for the market good and G is the demand for the outside good. The price of the market good is given by p and the price of the outside good is given by p_G. Both prices are given for the consumer and cannot be directly affected by their behaviour. This assumption rules out forms of bilateral bargaining about the market price between consumers and producers.

The price for the outside good can be interpreted as the costs or price index for a basket of goods for the consumers, where the market good is excluded from this basket. When prices of products in the basket increase over time, this will impact the price of the outside good.

Net spendable income of the consumer is assumed to be exogenous and given by Y. This income is after tax income and can be used freely by the consumer. Total spending of the consumer is given by $pQ + p_G G$, because it is assumed that all net income is spent either on the market good or the outside good. The

budget constraint of the individual is then given by:

$$Y - pQ - p_G G = 0. \tag{2.1}$$

From this budget constraint we can derive the consumed quantity of the outside good G:

$$G = \frac{Y - pQ}{p_G} \tag{2.2}$$

The quantity consumed of the outside good increases in the income level Y, decreases in the price of the basket of goods p_G, and decreases in consumer spending pQ on the market good. Substituting this quantity in the direct utility gives us an expression that only has the quantity level of the market good as a choice variable:

$$U\left(Q, \frac{Y - pQ}{p_G}\right) \tag{2.3}$$

The optimal choice of the quantity of the market good is governed by the utility of consumption (first argument) and the disutility of spending money on the market good (second argument). Optimal consumption can be derived by setting the total derivative with respect to the market quantity to 0:

$$\frac{dU}{dQ} = U_Q - \frac{p}{p_G} U_G = 0 \tag{2.4}$$

Here, $U_Q = \frac{\partial U}{\partial Q}$ is the marginal utility of consuming the market good. Here we use the notation $\frac{dU}{dQ}$ for the total derivative (both arguments of the utility function), and the notation $\frac{\partial U}{\partial Q}$ for the partial derivative with respect to Q (only the first argument). Marginal utility of consumption is the slope of the direct utility function with respect to the first argument Q. Furthermore, $U_G = \frac{\partial U}{\partial G}$ is the marginal utility of consumption of the outside good. This is equal to the price of the outside good multiplied with the marginal utility of income $p_G \frac{\partial U}{\partial Y}$. The second part of the first-order condition therefore captures the disutility of paying a price p for the market good. It also accounts for the relative price of the market good compared to the price for the basket of goods. From the first-order condition an implicit expression for the inverse demand curve can be derived. This expression shows that the ratio of prices is equal to the ratio of marginal utilities.

$$\frac{p}{p_G} = \frac{U_Q}{U_G} \longleftrightarrow p = p_G \frac{U_Q}{U_G} \tag{2.5}$$

An implicit solution means that the price p potentially appears at the right-hand side of the equation. This inverse demand curve is the maximum marginal

willingness to pay (MWTP) for the good. It depends on the relative size of the marginal utilities and increases in the price of the basket of goods. For many practical purposes the price of the basket of goods is normalized to $p_G = 1$ resulting in inverse demand as a function of the ratio of marginal utilities:

$$p = \frac{U_Q}{U_G} \tag{2.6}$$

The MWTP can be employed as a measure of consumer benefits. Much of the analytical work of demand analysists focusses on particular specifications that result in explicit solutions for the inverse demand curve. These explicit solutions have only exogenous parameters and quantity Q at the right-hand side and can be analysed using empirical data.

2.2.2 Examples of demand and inverse demand functions

To illustrate the model, we provide an example of three particular utility functions. First, we discuss a direct utility function without income effects and exponential marginal utility of consumption. Second, we discuss the well-known Cobb-Douglas utility function. This utility function captures non-linear impacts of net spendable income on direct utility. Third, we discuss a specification of the direct utility function that captures minimum threshold levels for consumption. Such a specification is particularly relevant for particular basic goods that a consumer cannot do without.

2.2.2.1 The exponential-linear utility function

Suppose direct utility is specified as:

$$U(Q, G) = \frac{A}{\alpha}\left(1 - e^{-\alpha Q}\right) + BG \tag{2.7}$$

The marginal utility of consumption is given by the partial derivative with respect to the market good:

$$\frac{\partial U}{\partial Q} = Ae^{-\alpha Q} > 0. \tag{2.8}$$

This marginal utility is positive implying that direct utility increases when consumption of the market good increases. The marginal utility of the outside good is given by the partial derivative with respect to G:

$$\frac{\partial U}{\partial G} = B > 0. \tag{2.9}$$

It is also positive implying more of the outside good is considered to be better.

The direct utility is concave[1] in the consumption of the market good and linear in the consumption of the outside good. This can be proved by investigating the second-order partial derivatives which are negative for consumption (first argument) and 0 for the outside good (second argument):

$$\frac{\partial^2 U}{\partial Q^2} = -\alpha A e^{-\alpha Q} < 0, \qquad \frac{\partial^2 U}{\partial G^2} = 0. \tag{2.10}$$

The intuitive appeal of the assumption of concavity is that the additional utility gained from consuming an additional unit of the market good is decreasing in the consumption level. This implies that the *marginal utility* of consuming the market good is positive, but decreasing. For example, people might enjoy eating meat, but the additional enjoyment of meat consumption decreases in the consumption level. At some point meat saturation kicks in and marginal utility becomes (close to) 0.

In line with pillar (i), the consumer is assumed to maximize utility while accounting for the budget constraint. From the budget constraint we can obtain the consumption level of the outside good:

$$Y - pQ - G = 0 \longleftrightarrow G = Y - pQ. \tag{2.11}$$

Substituting gives:

$$U(Q, G) = \frac{A}{\alpha}(1 - e^{-\alpha Q}) + B(Y - pQ). \tag{2.12}$$

The first-order condition for utility maximisation is given by the total derivative of the direct utility function with respect to the market good:

$$\frac{dU}{dQ} = Ae^{\alpha Q} - pB. \tag{2.13}$$

The first positive part is the marginal utility of consumption which is positive and the second negative part the marginal utility of payments for the market good. From the first-order condition one can solve for the inverse demand curve which expresses price p as a function of quantity Q. Furthermore, one can also solve for the demand curve which expresses quantity as a function of the price. In order to have a unique solution we have to check the second-order total derivative:

$$\frac{dU^2}{dQ^2} = -\alpha A e^{\alpha Q} < 0. \tag{2.14}$$

[1] Remember, concave function have decreasing marginal returns such as in Figure 1.1

This second-order total derivative is negative, implying that the optimum is a unique maximum. We start with solving for the inverse demand curve $p(Q)$[2], which is given by:[3]

$$p(Q) = \frac{A}{B}e^{-\alpha Q}. \tag{2.15}$$

Now we have an explicit solution for the MWTP because the right-hand side of this equation only depends on Q and exogenous utility parameters. At $Q = 0$, the MWTP is highest and equal to $\frac{A}{B}$. For higher values of Q, the MWTP decreases which is in line with decreasing marginal benefits of consumption. The parameter α governs the steepness of this decrease. The higher the parameter α is, the steeper the decrease of the marginal benefits.

Using the first-order condition for utility maximisation (2.13), one can also derive the demand curve (again, by setting $\frac{dU}{dQ} = 0$ and then isolating Q):

$$Q(p) = \frac{1}{\alpha}\ln(A) - \frac{1}{\alpha}\ln(p) - \frac{1}{\alpha}\ln(B) \tag{2.16}$$

This expression shows that demand for the market good:

1. Decreases in the price p as $\frac{\partial Q}{\partial p} = -\frac{1}{\alpha}\frac{1}{p} < 0$;
2. Increases in the utility parameter A as $\frac{\partial Q}{\partial A} = \frac{1}{\alpha}\frac{1}{A} > 0$;
3. Decreases in the utility parameter B as $\frac{\partial Q}{\partial A} = -\frac{1}{\alpha}\frac{1}{B} < 0$;
4. Decreases in the utility parameter α as $\frac{\partial Q}{\partial \alpha} = -\frac{1}{\alpha^2}(\ln(A) - \ln(p) - \ln(B)) < 0$;

As this point it is useful to see how this analytical result can be connected to the third pillar: *empiricism*. The demand function can be empirically estimated using statistical techniques and data on quantities and prices. Suppose an empirical economist estimates the following linear-log regression:

$$Q = \beta_0 + \beta_1 \ln p + \epsilon. \tag{2.17}$$

From equation (2.17) one can derive some of the utility parameters. The parameter β_1 is expected to be negative as demand decreases in the price. The mathematical model suggests that this parameter should be equal to $\frac{1}{\alpha}$. We can therefore write $\alpha = -\frac{1}{\beta_1} > 0$. The constant parameter β_0 is—according to the behavioural model—equal to $\beta_0 = \frac{1}{\alpha}\ln(\frac{A}{B})$. We can therefore write $\ln\frac{A}{B} = \alpha\beta_0$ or $\frac{A}{B} = e^{\alpha\beta_0} = e^{-\frac{1}{\beta_1}\beta_0}$.

[2]Recall that an inverse demand curve is a function of Q that determines p

[3]Recall that for an optimal utility, which is what we aim for, the condition $\frac{dU}{dQ} = 0$ should hold, so from (2.13) we then know that $Ae^{\alpha Q} = pB$ leading to (2.15)

The estimated parameters resulting from the regression equation therefore have a one-to-one correspondence to the underlying utility maximisation problem. An assumption that needs to be made is that the error ϵ in the regression equation is irrelevant for the analysis and captures particular measurement errors at the side of the researcher. If this is not the case, errors capture another part of the marginal utility and things become more complicated.[4] Dealing with unknown regression noise therefore already leads to interesting epistemological questions for applied economic research that we will assume away for now.

Furthermore, the behavioural model analyses individual choice behaviour and therefore the measurement of prices and demand should ideally be done at the individual level to stay close to the behavioural mathematical narrative that is developed.

2.2.2.2 The Cobb-Douglas utility function

One of the disadvantages of the introduced direct utility function of the previous section, is that the marginal utility of consumption of the outside good is assumed to be constant for each consumption level of the market good. The behavioural narrative therefore does not account for the plausible intuition that consumers with a lower income derive more utility from an income increase compared to consumers with a lower income (see for example Layard, Mayraz, and Nickell (2008)). In order to account for this, we propose a Cobb-Douglas type utility function that captures decreasing marginal utility for the market good and the outside good and directly substitute the expression for the outside good:[5]

$$U(Q,G) = A\ln(Q) + B\ln(G) = A\ln(Q) + B\ln(Y - pQ). \tag{2.18}$$

The first-order condition is now given by the following total derivative:

$$\frac{dU}{dQ} = \frac{A}{Q} - \frac{Bp}{Y - pQ} = 0. \tag{2.19}$$

Solving this first-order condition for the quantity level gives the following demand function:[6]

$$Q(p) = \frac{A}{A+B}\frac{Y}{p}. \tag{2.20}$$

[4] An interesting question that gives a hint for solving this issue is: could you propose a direct utility function that deals with ϵ as a parameter of the direct utility function?

[5] It is also possible to specify this utility function as: $U(Q,G) = ZQ^A G^B$. Log-linearizing gives the utility function used in the text and has no impact on the first-order condition. The constant Z is irrelevant for the analysis of demand.

[6] This solution is unique. You can check this by proving that the total second-order derivative is negative.

The demand curve depends on the exogenous utility parameters A and B, increases in the income level Y and decreases in the price p. Higher income levels therefore result in higher demand for the market good. The consumer spends $pQ(p) = \frac{A}{A+B}Y$ on the market good. From the budget constraint we can work back the spending on the outside good which are equal to $Y - pQ = \frac{B}{A+B}Y$. This is an attractive aspect of the Cobb-Douglas specification: expenditure shares on the market good are directly related to relative utility parameters $\frac{A}{A+B}$.

For empirical purposes, the demand curve can be estimated using econometric techniques. Log-linearizing the demand curve gives the following regression equation:

$$\ln Q = \ln(\frac{A}{A+B}) + \ln Y - \ln p + \epsilon. \tag{2.21}$$

The Cobb-Douglas specification of utility therefore *requires* that the parameters for income and price in the demand function are 1 and -1 respectively. This can be tested using statistical hypothesis testing. The inverse demand curve for this particular utility function is given by:

$$p(Q) = \frac{A}{A+B}\frac{Y}{Q}. \tag{2.22}$$

A disadvantage of this specification from a behavioural perspective is that the MWTP for the first marginal unit of the good is infinite. This is because when $Q \longrightarrow 0$ then $p(Q) \longrightarrow \infty$. An advantage of this specification is that the MWTP for the market good is increasing in the net spendable income Y. For environmentally sustainable goods it is expected that this will be a plausible assumption.

2.2.2.3 A utility function that allows for minimum consumption tresholds

We continue the discussion on demand models with a third example. It might be that individuals need to consume a minimum amount of particular basic goods such as food or energy. The Stone-Geary utility function is able to deal with this. Again, we use a log-linearized version of the utility function:

$$U(Q, G) = A\ln(Q - \underline{Q}) + B\ln(G - \underline{G}), \tag{2.23}$$

In this equation $\underline{Q} < Y$ is the minimum amount needed of the market good and $\underline{G} < Y$ the minimum amount needed for the outside good. This minimum amount of the outside good can be interpreted as consumer spending needed for basic goods which are not the market good. Utility goes to minus infinity when consumption is lower or equal to these minimum amounts. Substituting $G = Y - pQ$ from the budget constraint gives:

$$U(Q, G) = A \ln(Q - \underline{Q}) + B \ln(Y - pQ - \underline{G}), \tag{2.24}$$

The first-order condition is given by the following total derivative:

$$\frac{dU}{dQ} = \frac{A}{Q - \underline{Q}} - \frac{Bp}{Y - pQ - \underline{G}} = 0, \tag{2.25}$$

which can be written as:

$$bpQ - bp\underline{Q} = AY - ApQ - A\underline{G}. \tag{2.26}$$

Rearranging gives:

$$Qp(B + A) = AY = Bp\underline{Q} - A\underline{G}. \tag{2.27}$$

Solving for the demand gives:

$$Q(p) = \frac{A}{A+B}\frac{Y}{p} + \frac{B}{A+B}\underline{Q} - \frac{A}{A+B}\frac{1}{p}\underline{G}. \tag{2.28}$$

The first term of the demand function is equal to the standard Cobb-Douglas demand derived in the previous section. The second term accounts for the threshold consumption of the market good. Market demand increases in the threshold demand $\underline{Q}$. The third term shows that market demand decreases in the threshold demand $\underline{G}$ as more money is needed to consume basic goods in the basket of outside goods. When the market good is not necessary and $\underline{Q} = 0$, market demand can be written as:

$$Q(p) = \frac{A}{A+B}\frac{Y}{p} - \frac{A}{A+B}\frac{1}{p}\underline{G} = \frac{A}{A+B}\frac{1}{p}\left(Y - \underline{G}\right). \tag{2.29}$$

Log-linearizing this demand function gives the following regression equation:

$$\ln Q = \ln \frac{A}{A+B} + \ln(Y - \underline{G}) - \ln p + \epsilon. \tag{2.30}$$

Demand is not linear in the log of the net spendable income anymore as first the necessary spending on the outside good should be accounted for. The price coefficient and the coefficient for $\ln(Y - \underline{G})$ still enter with a coefficient equal to 1. For applied work this means that if the Stone-Geary model of behaviour is correct, one first should subtract spending on basic goods $\underline{G}$ from the net spendable income, and introduce the logged variable as an explanatory variable in the regression.

2.2.2.4 Demand modelling: conclusion

The take away of these three examples is that particular specifications of the direct utility function result in equations for individual demand curves that can be empirically operationalized and estimated making additional assumptions. Some specifications of the direct utility function are behaviourally more plausible whereas other specifications might be easier to operationalize using empirical data. This is the reason that economists prefer to have multiple modelling options in their toolbox.

2.2.3 Consumer benefits and surplus in markets

2.2.3.1 Market demand and elasticities

The behavioural models of the previous section can be viewed as models of individual choice behaviour. The demand and inverse demand functions are therefore in principal specified at the individual level. For the analysis of markets, these functions are aggregated into a market (inverse) demand function. This implies that when we analyse market quantities and prices in regression analysis, additional assumptions should be made to relate the results of these regressions to the underlying distribution of utility parameters in the population. Developing models that account for a correct mathematical aggregation to market demand curves often requires stringent assumptions on preferences.

Consider for example the Cobb-Douglas demand that is developed in section 2.2.2.2 and denote $Q_n(p) = \frac{A_n}{A_n+B_n} Y_n \frac{1}{p}$ as the individual demand. Assume we have a population of N individuals. The market demand can be given as the sum of the individual demands:

$$Q^D(p) = \sum_{n=1}^{N} \frac{A_n}{A_n + B_n} Y_n \frac{1}{p} = \frac{1}{p} \sum_{n=1}^{N} \frac{A_n}{A_n + B_n} Y_n. \tag{2.31}$$

Although aggregation is possible, it requires additional analytical assumption to be neatly done.

Because of complexity we avoid this aggregation discussion by neglecting the behavioural model underlying the decision of the population of consumers. We continue by directly interpreting the market inverse demand curve that is estimated in regressions as the total MWTP for the market good and assume the regression constant adjusts for aggregation issues. From now on we therefore use $p(Q)$ to denote the market inverse demand curve. This market inverse demand curve measures what the price of the market good would have to be for Q units of it to be demanded (Varian (2003), 2006, p.268). Consumers can buy the good multiple times and are sorted *along* the inverse demand curve on the basis of their MWTP where the intuition for the derivation of the MWTP stems from the behavioural model that is developed.

The MWTPs estimated from regressions of market demand can then be interpreted as measures of consumer benefits. The total consumer benefits for the market good are in turn given by the sum of the benefits of all consumers who participate in the market. Besides looking at market data, one could also operationalise the intuition of MWTP using survey research. In a survey we could for example ask people to express their WTP for an additional unit of consumption and sort the WTPs to obtain the inverse demand function. For example, we could ask the following question:

> What would be your maximum willingness to pay for an additional unit of these vegan chicken nuggets?

The answers to these questions can serve as an input to derive the inverse demand curve and the demand curve.

For applied work we are often interested in the price elasticity of market demand. This is defined by:[7]

$$\epsilon = \frac{\partial Q^D}{\partial p} \frac{p}{Q^D(p)}. \tag{2.32}$$

The definition of an elasticity is as follows: for a 1% increase in the market price, the market demand will change with ϵ%. For many market goods, this elasticity will be negative as higher prices result in lower demand. When the elasticity is lower than -1, the market good has elastic demand. When the elasticity is in between 0 and -1, it has inelastic demand.

We illustrate the concept of price elasticity of demand using an example. Brons et al. (2002) give an overview of price elasticities of demand for air travel. Figure 2.1 below is from their paper.

For short distance air travel the median value of the price elasticity is -1.5. This implies that for short distance travel a 1% increase in the price will reduce air travel for these trips with 1.5%. For medium-distance trips the median price elasticity of market demand is given by -1.4 implying that a 1% increase in the price reduces demand with 1.4%. For long-distance flights the median is -0.8, implying that a 1% increase in the price reduces demand with 0.8%. For the assumed numbers above, long-distance flight demand is therefore inelastic whereas short distance and medium distance flights have elastic demands. are therefore less sensitive to the price compared to travellers.

Another example of estimated price elasticities comes from a meta-analysis of Andreyeva, Long, and Brownell (2010). They give a summary of estimates of absolute price elasticities of demand for food categories in the US (displayed in Figure 2.2). All food categories have inelastic demand as elasticities are in between 0 and -1. The estimated price elasticities are between -0.27 (eggs) and -0.81 (food away from home). A 1% increase in the price of eggs therefore results

[7] Note that the ϵ here is different from that what we normally use in regression equations.

Fig. 4. (a) Distribution of price elasticity estimates from short-distance flights. (b) Distribution of price elasticity estimates from medium-distance flights. (c) Distribution of price elasticity estimates from long-distance flights.

Figure 2.1: Overview of studies on the price elasticities of market demand for air travel

in 0.27% lower demand. For beef the estimated elasticity is −0.75, implying that a 1% increase in the price for beef results in a 0.75% lower demand. Uncertainty in the estimates is reported with the 95% confidence intervals.

TABLE 1—US Price Elasticity Estimates, by Food and Beverage Category, from 1938–2007

Food and Beverage Category[a]	Absolute Value of Mean Price Elasticity Estimate (95% CI)	Range	No. of Estimates
Food away from home	0.81 (0.56, 1.07)	0.23-1.76	13
Soft drinks	0.79 (0.33, 1.24)	0.13-3.18	14
Juice	0.76 (0.55, 0.98)	0.33-1.77	14
Beef	0.75 (0.67, 0.83)	0.29-1.42	51
Pork	0.72 (0.66, 0.78)	0.17-1.23	49
Fruit	0.70 (0.41, 0.98)	0.16-3.02	20
Poultry	0.68 (0.44, 0.92)	0.16-2.72	23
Dairy	0.65 (0.46, 0.84)	0.19-1.16	13
Cereals	0.60 (0.43, 0.77)	0.07-1.67	24
Milk	0.59 (0.40, 0.79)	0.02-1.68	26
Vegetables	0.58 (0.44, 0.71)	0.21-1.11	20
Fish	0.50 (0.30, 0.69)	0.05-1.41	18
Fats/oils	0.48 (0.29, 0.66)	0.14-1.00	13
Cheese	0.44 (0.25, 0.63)	0.01-1.95	20
Sweets/sugars	0.34 (0.14, 0.53)	0.05-1.00	13
Eggs	0.27 (0.08, 0.45)	0.06-1.28	14

Note. Values were calculated based on the 160 studies reviewed. Absolute values of elasticity estimates are reported. The price elasticity of demand measures the percentage change in purchased quantity or demand with a 1% change in price.
[a]Including restaurant meals and fast food.

Figure 2.2: Price elasticity estimates for food categories.

2.2.3.2 Deriving consumer benefits

Total consumer benefits in the market are given by all the consumer benefits of those who buy the market good. An analogous continuous representation of total consumer benefits is given by the integral under the *inverse market demand function* from 0 to the equilibrium quantity:

$$CB = \int_0^{Q*} p(Q)dQ. \tag{2.33}$$

For example, when the inverse demand curve is given by $p(Q) = \frac{A}{B}e^{-\alpha Q}$, consumer benefits can be derived analytically by using the primitive function of

the inverse demand function:

$$CB = \int_0^{Q^*} \frac{A}{B} e^{-\alpha Q} = \left[-\frac{1}{\alpha} \frac{A}{B} e^{-\alpha Q} + z \right]_0^{Q^*} \tag{2.34}$$

$$= \left(-\frac{1}{\alpha} \frac{A}{B} e^{-\alpha Q^*} + z \right) - \left(-\frac{1}{\alpha} \frac{A}{B} e^{-\alpha 0} + z \right) \tag{2.35}$$

$$= \left(-\frac{1}{\alpha} \frac{A}{B} e^{-\alpha Q^*} \right) - \left(-\frac{1}{\alpha} \frac{A}{B} 1 \right) \tag{2.36}$$

$$= \left(-\frac{1}{\alpha} \frac{A}{B} e^{-\alpha Q^*} \right) + \frac{1}{\alpha} \frac{A}{B} \tag{2.37}$$

$$= -\frac{1}{\alpha} \frac{A}{B} \left[e^{-\alpha Q^*} - 1 \right] = \frac{1}{\alpha} \frac{A}{B} \left[1 - e^{-\alpha Q^*} \right]. \tag{2.38}$$

The process of the derivation of the primitive function can be viewed as reverse differentiation: the derivative of the primitive function gives back the original function (in our case the inverse demand function). It therefore can only be derived up to a particular constant z. The primitive function should be evaluated at the equilibrium demand $Q*$ and the point 0. The difference between these function evaluations gives the area under the inverse demand function from 0 to $Q*$. The third step shows that the constant z then drops out and is therefore irrelevant for the outcomes.

The final result shows that consumer benefits can be numerically estimated when the parameters α, $\frac{A}{B}$ and $Q*$ are known. From our earlier discussion on this inverse demand function, we had $\alpha = -\frac{1}{\beta_1} > 0$ and $\frac{A}{B} = e^{-\frac{\beta_0}{\beta_1}} > 0$. As $0 < e^{-\alpha Q*} < 1$. Therefore, in line with intuition, consumer benefits always will be positive.

Substituting the regression parameters in the expression for the consumer benefits gives an estimate for consumer benefits as a function of the observed/derived equilibrium quantity and the regression parameters of the market demand function:

$$\widehat{CB} = -\hat{\beta}_1 e^{-\frac{\beta_0}{\beta_1}} \left[1 - e^{-\frac{Q*}{\beta_1}} \right]. \tag{2.39}$$

Using regression techniques in combination with assumptions on the market inverse demand curve, economists are therefore able to estimate the consumer benefits of consumption of a market good. It is important to account for the dimensions of this expression. When consumption levels are given *per year* and prices are given *in euros*, the consumer benefits for the particular market are given *in euros per year*.

For example, consider the market for vegan chick nuggets. Assume that the total market quantity for these nuggets is given by $200,000$ pieces sold per year

in The Netherlands. The estimated parameters are given by $\hat{\beta}_1 = -0.08$ and $\beta_0 = 1.44$. Substituting in the equation for the consumer benefits gives:

$$\widehat{CB} = 0.08e^{-\frac{1.44}{-0.08}}\left[1 - e^{-\frac{200,000}{-0.08}}\right] = 5,252,798 \text{ euros per year.} \tag{2.40}$$

For this example, the market for vegan nuggets in the Netherlands therefore results in consumer benefits of about 5.25 million euros per year.

2.2.3.3 Deriving changes in consumer benefits.

Sometimes researchers are interested in changes in consumer benefits rather than the level of consumer benefits. This is especially relevant for particular policy recommendations/interventions where the change in value compared to doing nothing is analysed. Suppose there is a change in equilibrium quantity from Q^* to Q^{**}. The change in consumer benefits can then be derived using rules of integrals and is given by:

$$\Delta CB = \int_0^{Q^{**}} p(Q)dQ - \int_0^{Q^*} p(Q)dQ = \int_{Q^*}^{Q^{**}} p(Q)dQ. \tag{2.41}$$

Suppose we employ a stable Cobb-Douglas inverse demand curve over time which is derived in the previous section and defined as: $p(Q) = \frac{A}{A+B}\frac{Y}{Q}$. The change in consumer benefits can then be derived analytically:

$$\Delta CB = \int_{Q^*}^{Q^{**}} \frac{A}{A+B}\frac{Y}{Q}dQ = \left[\frac{A}{A+B}Y\ln Q = z\right]_{Q^*}^{Q^{**}} = \frac{A}{A+B}Y\left[\ln Q^{**} - \ln Q^*\right] \tag{2.42}$$

$$= \frac{A}{A+B}Y\left[\ln\left(\frac{Q^{**}}{Q^*}\right)\right]. \tag{2.43}$$

This uses the fact that the primitive function of $\frac{1}{Q}$ is given by $\ln(Q)$. The last step is taken using rules of working with logarithms: $ln(x) - ln(y) = ln(x/y)$. Again, the primitive function is derived of which the derivative gives back the inverse demand function. The change in consumer benefits is positive when $Q^{**} > Q^*$, and negative when $Q^{**} < Q^*$. For this particular specification of the utility function the change is proportionally increasing in the income level. In order to operationalise this equation, one has to account for the dimension of the quantity level and the dimension of the income variable. For the budget constraint the net spendable income is relevant. Suppose income in the regression is given by net spendable income per month and is equal to $Y = 3,500$ euros.

Furthermore, suppose there is an increase in the consumption of vegan nuggets from $200,000$ pieces per year in 2019 to $250,000$ pieces per year in 2020. For the

analysis of consumer benefits, we need the equality $\frac{A}{A+B} = e^{\beta_0}$, where β_0 is the constant of the regression. An econometrician estimates the following regression:

$$Q = \beta_0 + \beta_1 \ln p + \beta_2 \ln Y + \epsilon, \tag{2.44}$$

and finds $beta_0 = 5.2$, $\beta_1 = -1.02$ and $\beta_2 = 0.97$.

The constant is therefore estimated at $beta_0 = 5.2$. The parameters before the logged price and the logged income are not significantly different from 1. Therefore, we can assume Cobb-Douglas preferences describe choice behaviour correctly. The next step is to substitute the numerical values for β_0, Q^* and Q^{**} in the expression for the change in consumer benefits. This results in:

$$\Delta CB = e^{5.2} 3500 \left[\ln \frac{250000}{200000} \right] = 141574 \text{ euros per year.} \tag{2.45}$$

This example shows how regression results can be connected to empirical estimates of consumer benefits.

2.2.3.4 Deriving consumer surplus

Until now we have not accounted for the fact that the consumer also pays a price for the product. Consumer surplus is defined as the total consumer benefits minus the total consumer costs:

$$CS = \int^{Q_0^*} p(Q) dQ - p^* Q^*. \tag{2.46}$$

Here p^* is the equilibrium price that is observed and Q^* the equilibrium quantity. In order to calculate consumer surplus for our vegan chicken nuggets example, this requires input on the equilibrium price. Suppose the equilibrium price is equal to 2 euros per unit. Consumer surplus for The Netherlands is then given by:

$$CS = 5,252,798 - 2 \times 200,000 = 1,252,798 \text{ euros per year.} \tag{2.47}$$

To conclude: the mathematical analysis of consumer choice behaviour is useful to structure our thinking about consumer behaviour and the origins of the market inverse demand curve. This market inverse demand curve can be interpreted as the MWTP of a group of consumers with MWTPs sorted from high to low and can be estimated using regression techniques or studied using survey questions. Mathematical modelling can be employed to relate estimated parameters in regressions to the calculation of (changes) in consumer benefits where it is important to account for the dimensions of the variables.

2.3 Poducer behaviour and surplus

2.3.1 Producer cost functions and cost minimisation

The next step is to analyse the behaviour of producers in the market. This section discusses the classical economic model that helps to understand the economic aspects of decisions of producers. It is assumed that producers minimize their costs given a production target. We start simple by assuming that there are two inputs for the production function: capital K and labour L. The *production function* gives the production output as a function of the inputs: $f(K, L)$. The production costs for the firm are given by the *cost function* $C(K, L)$ which shows the costs for a given combination of inputs. This cost function captures the costs of capital—which are usually dependent on interest rates—and the costs of labour which can be interpreted as wages and/or overhead.

Suppose the firm seeks to produce a target quantity for the market equal to Q. According to the model, the firm then will minimize the costs with respect to capital K and labour L, but given the production constraint that the target quantity is equal to the production output. Formally, this can be written as:

$$\min_{K,L} C(K, L) \text{ subject to } f(K, L) = Q. \tag{2.48}$$

As there is a constrained involved, this minimisation problem can be analysed using Lagrangian techniques. This technique rewrites the constrained optimisation problem as an unconstrained optimisation problem. First, the Lagrangian function H needs to be defined. The first part of this Lagrangian function gives the goal function that needs to be minimized or maximized. The second part of this function adds a new variable λ to the optimisation problem and multiplies this variable with the constraint:

$$H = C(K, L) + \lambda(Q - f(K, L)). \tag{2.49}$$

The original optimisation problem of the firm is then mathematically reformulated as an unconstrained problem:

$$\min_{\lambda,K,L} H. \tag{2.50}$$

To solve this unconstrained minimisation problem, we need to analyse the familiar first-order conditions with respect to capital, labour and the new variable λ (which is the so-called Lagrangian multiplier). The three first-order conditions are given by:

$$\frac{\partial H}{\partial \lambda} = Q - f(K, L) = 0, \tag{2.51}$$
$$\frac{\partial H}{\partial K} = \frac{\partial C}{\partial K} - \lambda \frac{\partial f}{\partial K} = 0, \tag{2.52}$$
$$\frac{\partial H}{\partial L} = \frac{\partial C}{\partial L} - \lambda \frac{\partial f}{\partial L} = 0. \tag{2.53}$$
$$\tag{2.54}$$

The first first-order condition with respect to the Lagrangian multiplier guarantees that in equilibrium, the production target is always reached. The reason is that this condition states that the production target Q should be equal to the production level $f(K, L)$. The second condition guarantees optimal choice of capital by equation the marginal costs (or 'price') of capital with the marginal benefits. The third condition has a similar structure and equates the marginal costs of labour with the marginal benefits of labour. Assume a linear cost function with input prices equal to r for capital and ω for labour. When the cost function is linear in inputs and input prices, the second and the third first-order condition can be rewritten:

$$\frac{r}{\omega} = \frac{\frac{\partial f}{\partial K}}{\frac{\partial f}{\partial L}}. \tag{2.55}$$

Or in other words: the ratio of input prices is equal to the ratio of marginal productivities of the two respective inputs, which is the marginal rate of substitution. This result shares similarities with the result for consumer optimisation where the ratio of prices is equal to the ratio of marginal *utilities*.

Before moving on to an analytical example it is useful to reflect on the idea of the cost function. In the analysis it is assumed that the cost function only depends on the inputs. This assumption implies that firms neglect potential negative costs of production as the cost function itself is not a function of Q.

2.3.2 Specifying and interpreting the production function

To make things a bit less abstract we develop an example assuming a Cobb-Douglas type of production function and a linear cost function with input prices equal to r for capital and ω for labour. Before moving into technical derivations, it is useful to interpret the production function in more detail in order to better understand what it captures. The Cobb-Douglas type production function is given by:

$$f(K, L) = TK^{\alpha_K} L^{\alpha_L} \qquad T > 0. \tag{2.56}$$

This function has several properties. First, both inputs are *necessary* in order to have positive production. When there is no capital investment or no labour available, production is not possible and equals 0. Second, the parameter $T > 0$ captures the state of overall production technology of the firm. For a given combination of inputs K and L, a higher parameter T implies higher production levels. Improvements in overall technology therefore can result in higher production levels at the same level of inputs. Third, together with the parameter T, and the input levels K and L, α_K and α_L determine the marginal production impact of increasing input levels. It is plausible to assume that production is increasing in the input levels. The first derivatives with respect to the inputs show that production is increasing in the inputs:

$$\frac{\partial f}{\partial K} = \frac{\alpha_K}{K} T K^{\alpha_K} L^{\alpha_L} > 0 \tag{2.57}$$

$$\frac{\partial f}{\partial L} = \frac{\alpha_L}{L} T K^{\alpha_K} L^{\alpha_L} > 0 \tag{2.58}$$

$$\tag{2.59}$$

To have increasing production in the inputs K and L the parameters α_K and α_L should therefore be positive.

Fourth, the production function should plausibly be concave in the level of each production input. An additional unit of an input will have lower impacts when there is already a high level of the input present. This can be investigated by looking at the second derivatives:

$$\frac{\partial^2 f}{\partial K^2} = \frac{\alpha_K(\alpha_K - 1)}{K^2} T K^{\alpha_K} L^{\alpha_L} < 0 \tag{2.60}$$

$$\frac{\partial^2 f}{\partial L^2} = \frac{\alpha_L(\alpha_L - 1)}{L^2} T K^{\alpha_K} L^{\alpha_L} < 0 \tag{2.61}$$

$$\tag{2.62}$$

In order to have concave impacts we should make the assumptions $\alpha_K < 1$ and $\alpha_L < 1$. Fifth, the inputs in the Cobb-Douglas production function are complements. We can investigate this mathematically by looking at the cross-derivatives of the production function which are given by:

$$\frac{\partial^2 f}{\partial K \partial L} = \frac{\partial^2 f}{\partial L \partial K} = \frac{\alpha_K}{K} \frac{\alpha_L}{L} T K^{\alpha_K} L^{\alpha_L} > 0 \tag{2.63}$$

These cross-derivatives are positive showing that the marginal productivity of the inputs increases in the levels of the other inputs. Therefore, the inputs are complements.

Sixth, the substitution parameters α_L and α_K of the production function have a very useful economic interpretation. They are the elasticity of production with respect to the inputs. This elasticity is defined as the percentage change in production for a percentage change in one of the inputs. For example, for labour it is given by:

$$\frac{\partial f}{\partial L}\frac{L}{f(K,L)} = \frac{\alpha_L}{L}TK^{\alpha_K}L^{\alpha_L}\frac{L}{TK^{\alpha_K}L^{\alpha_L}} = \alpha_L. \tag{2.64}$$

Similarly, for capital we find:

$$\frac{\partial f}{\partial K}\frac{K}{f(K,L)} = \frac{\alpha_K}{K}TK^{\alpha_K}L^{\alpha_L}\frac{K}{TK^{\alpha_K}L^{\alpha_L}} = \alpha_K. \tag{2.65}$$

In the next section we will show that these elasticities play a key role in the determination of the shape of the cost function with respect to quantity produced.

Seventh, we can investigate whether there are increasing, decreasing or constant returns to scale in production levels. When we multiply all the inputs with a fixed factor $t > 1$ (for example multiplying the levels of the inputs with 2), the production function can be written as:

$$f(K,L) = T(tK)^{\alpha_K}(tL)^{\alpha_L} = t^{\alpha_K+\alpha_L}TK^{\alpha_K}L^{\alpha_L} \tag{2.66}$$

Multiplying the inputs with a factor t, therefore leads to a multiplication of production of $t^{\alpha_K+\alpha_L}$.[8] When $\alpha_K + \alpha_L < 1$, there will be decreasing returns to scale in production as a multiplication of the inputs with a factor t, leads to a multiplication of production less than t. When $\alpha_K + \alpha_L < 1$, there is constant returns to scale in production and multiplication of the inputs with t leads to multiplication of production levels with t. When $\alpha_K + \alpha_L > 1$, there will be increasing returns to scale in production as a multiplication of the inputs with a factor t, leads to a multiplication of production with more than t. Again the input elasticities play a key role here.

2.3.3 The cost function

The next step is to specify the cost function of the firm and derive the optimal cost function. We do this, because often only the costs of the firm can be observed (but not the optimal inputs). Furthermore, the cost optimal cost function plays a key role for the market behaviour of the firm.

It is assumed to be linear in expenditures on the inputs (for our case labour and capital): $c(K,L) = rK + \omega L$. This is a reasonable assumption unless the firm can negotiate lower prices for higher amounts of the inputs. In empirical work researchers sometimes specify input prices as a function of the total market

[8] We therefore say that the production function is homogenous of degree $\alpha_K + \alpha_L$.

demand for these inputs (see Pillai (2015)). Furthermore, it is assumed that costs related to production which fall outside the firm are not included. If the firm does take these costs into account, an additional term related to the target quantity can be added to the cost function and the choice of optimal target quantity can be included in the analysis.

The Lagrangian function that describes the constrained optimisation problem of the firm as an unconstrained optimisation problem can be written as:

$$H = rK + \omega L + \lambda(Q - TK^{\alpha_K}L^{\alpha_L}) \tag{2.67}$$

The three first-order conditions are given by:

$$\frac{\partial H}{\partial \lambda} = Q - TK^{\alpha_K}L^{\alpha_L} = 0, \tag{2.68}$$

$$\frac{\partial H}{\partial K} = r - \lambda\frac{\partial \alpha_K}{\partial K}TK^{\alpha_K}L^{\alpha_L} = 0, \tag{2.69}$$

$$\frac{\partial H}{\partial L} = \omega - \lambda\frac{\alpha_L}{\partial L}TK^{\alpha_K}L^{\alpha_L} = 0. \tag{2.70}$$

$$\tag{2.71}$$

The first condition ensures that the production target is reached. The second condition equates the marginal costs of capital to the marginal benefits of capital. The third condition equates the marginal costs of labour to the marginal benefits of labour. In Appendix A we show that the optimal capital and labour are given by:

$$K^* = \frac{Q}{T}^{\frac{1}{\alpha_K+\alpha_L}}\frac{\omega}{r}^{\frac{\alpha_L}{\alpha_K+\alpha_L}}\frac{\alpha_K}{\alpha_L}^{\frac{\alpha_L}{\alpha_K+\alpha_L}}, \tag{2.72}$$

$$L^* = \frac{Q}{T}^{\frac{1}{\alpha_K+\alpha_L}}\frac{r}{\omega}^{\frac{\alpha_K}{\alpha_K+\alpha_L}}\frac{\alpha_L}{\alpha_K}^{\frac{\alpha_K}{\alpha_K+\alpha_L}}, \tag{2.73}$$

$$\tag{2.74}$$

It might be useful for you to program these expressions in order to see how the target quantity, the technology parameters, the input prices and the input elasticities relate to the optimally chosen inputs. First, the optimal inputs are non-linear in the target quantity and the shape depends on the sum of the input elasticities. With increasing returns in production, the optimal inputs are convex in the target quantity. With constant returns, the optimal inputs are linear in the target quantity and with decreasing returns in production, the inputs are concave in the target quantity.

Second, the optimal inputs depend on the relative input prices, decrease in the own input price and increase in the input price of the other input. For the optimal choice of capital— $\frac{\alpha_L}{\alpha_K+\alpha_L}$ is the elasticity with respect to the input price r: when the interest rate increases with 1% the optimal capital will decrease with $\frac{\alpha_L}{\alpha_K+\alpha_L}$%. Furthermore, $\frac{\alpha_L}{\alpha_K+\alpha_L} > 0$is the elasticity of optimal capital with respect to labour prices. When labour prices are interpreted as wages this gives useful insights in how wage changes impact input substitution. Therefore, there is an important one-to-one relationship between the production elasticities and the elasticity of input prices.

Third, when technology improves and T increases, the optimal inputs decrease as production per unit of input is higher. Suppose production becomes z times more effective due to an increase in T from T to zT. Optimal inputs then change with a factor $\left(\frac{1}{z}\right)^{\frac{1}{\alpha_K+\alpha_L}}$. For example, when production becomes 50% more effective due to newly developed robot technology, and $\alpha_K + \alpha_L = 1$, this implies that the new inputs are $\left(\frac{1}{1.5}\right)^{\frac{1}{1}} = \frac{2}{3}$ of the old inputs. Robot technology therefore is expected to decrease both capital and labour.

Fourth, the optimal inputs relate in a complicated way to the input elasticities via the multiplication factor. For constant returns, $\alpha_K + \alpha_L = 1$ the expressions simplify and are given by:

$$K^* = \frac{Q}{T}\frac{\omega}{r}^{\alpha_L}\frac{\alpha_K}{\alpha_L}^{\alpha_L}, \tag{2.75}$$

$$L^* = \frac{Q}{T}\frac{r}{\omega}^{\alpha_K}\frac{\alpha_L}{\alpha_K}^{\alpha_K}, \tag{2.76}$$

$$\tag{2.77}$$

Substituting the optimal inputs in the cost function gives the optimal costs for the firm as a function of the technology parameter, the production elasticities and the input prices:

$$C(K^*, L^*) = rK^* + \omega L^* = \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}} \omega^{\frac{\alpha_L}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K}{\alpha_K+\alpha_L}} \left(\left(\frac{\alpha_K}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} + (\left(\frac{\alpha_L}{\alpha_K}\right)^{\frac{\alpha_K}{\alpha_K+\alpha_L}}\right). \tag{2.78}$$

Before moving on it is useful to draw some conclusions from this mathematical expression of the cost function.

First, the expression depends on the target quantity in potential non-linear ways. The shape of the costs function with respect to the target quantity is governed by $\frac{1}{\alpha_K+\alpha_L}$ and therefore depends on the sum of the input elasticities. For increasing returns in production, the sum of the input elasticities is larger than 1 and the costs are concave in the target quantity. The costs are convex

in the target quantity for decreasing returns in production. The special case of constant returns in production, $\alpha_K + \alpha_L = 1$, gives a linear cost function in the target quantity. The sum of the input elasticities of the production function therefore determine the shape of the cost function with respect to the quantity level.

Second, the derived cost function shows an inverse relation with the technology parameter T. The higher the level of production technology, the lower the (marginal) costs of production. The downward impact of technology improvements on costs is stronger when the sum of the parameters $\alpha_K + \alpha_L$. The elasticity for this parameter is given by $\frac{-1}{\alpha_K+\alpha_L}$, showing that for a 1% increase in the technology parameter, the costs decrease with $\frac{1}{\alpha_K+\alpha_L}$.

Third, the cost function depends in a non-linear way on the input prices r and ω. When wages are higher, costs of the firm will increase where the strength of the increase depends on the relative size of the technology parameters α_K and α_L. For example, the elasticity of costs with respect to wages is given by $\frac{\alpha_L}{\alpha_K+\alpha_L}$: for a 1% increase in the wages, costs of the firm increase with $\frac{\alpha_L}{\alpha_K+\alpha_L}$%. For capital, a 1% increase in the capital price leads to a $\frac{\alpha_K}{\alpha_K+\alpha_L}$% increase in the costs of the firm.

Fourth, the technology parameters α_K and α_L enter the cost function in a complicated non-linear way in the last term of the cost function. When researchers log-linearize the cost function this part ends up in the regression constant.

Fifth, from the cost function we can derive the average costs and the marginal costs:

$$AC = \frac{C(K^*, L^*)}{Q} = \left(\frac{1}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}} Q^{\frac{1-\alpha_K-\alpha_L}{\alpha_K+\alpha_L}} \omega^{\frac{\alpha_L}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K}{\alpha_K+\alpha_L}} \left(\left(\frac{\alpha_K}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} + (\left(\frac{\alpha_L}{\alpha_K}\right)^{\frac{\alpha_K}{\alpha_K+\alpha_L}}\right), \tag{2.79}$$

$$MC = \frac{\partial C(K^*, L^*)}{\partial Q} = \frac{1}{\alpha_K + \alpha_L} \left(\frac{1}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}} Q^{\frac{1-\alpha_K-\alpha_L}{\alpha_K+\alpha_L}} \omega^{\frac{\alpha_L}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K}{\alpha_K+\alpha_L}} \left(\left(\frac{\alpha_K}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} + (\left(\frac{\alpha_L}{\alpha_K}\right)^{\frac{\alpha_K}{\alpha_K+\alpha_L}} \tag{2.80}$$

Marginal and average costs are independent of the target quantity when there are constant returns to scale. For increasing returns, the average costs decrease in the target quantity and for decreasing returns average costs increase. For constant returns to scale the average and marginal costs curves are equal. This shows again the important role of the input elasticities.

At the end of this sub-section it is useful to discuss the producer results of pillar (1.) in relation to the third pillar of economic analysis: empiricism. There are two ways to investigate the structural parameters of the cost function. First, empirical economists have employed regression estimates of the Cobb-Douglas

production function using realized quantities of firms as a function of the firms' inputs. This can be operationalized by log-linearizing the production function:

$$\ln Q = \ln T + \alpha_K \ln K + \alpha_L \ln L + \epsilon \tag{2.81}$$

Such a regression requires firm data on production quantities and levels of inputs. Statistical tests can be performed in order see whether the cost function is significantly different from linear (in quantity Q). These tests use $\alpha_K + \alpha_L = 1$ as the null hypothesis.

The second strategy is to run regressions directly on the cost functions. For example, Pillai (2015)) estimates the average production cost function for the market for solar panels using average cost levels and data on input prices, realized quantities and input levels. Log-linearizing the average cost function gives the following equation:

$$\ln AC = -\frac{1}{\alpha_K + \alpha_L} + \frac{1 - \alpha_K - \alpha_L}{\alpha_K + \alpha_L} \ln Q + \frac{\alpha_L}{\alpha_K + \alpha_L} \ln \omega + \frac{\alpha_K}{\alpha_K + \alpha_L} \ln r + \tag{2.82}$$

$$\left(\left(\frac{\alpha_K}{\alpha_L} \right)^{\frac{\alpha_L}{\alpha_K + \alpha_L}} + \left(\frac{\alpha_L}{\alpha_K} \right)^{\frac{\alpha_K}{\alpha_K + \alpha_L}} \right) \tag{2.83}$$

The first and the last term of this equation can be combined in a regression constant.

The regression equation that is estimated is then given by:

$$\ln AC = \beta_0 + \beta_1 \ln Q + \beta_2 \ln \omega + \beta_3 \ln r + \epsilon. \tag{2.84}$$

Hypothesis testing can be performed to see whether parameters are in accordance with the Cobb-Douglas specification. For example, theory dictates that for Cobb-Douglas to hold we must have $\beta_2 + \beta_3 = 1$ and $\beta_1 + \beta_2 + \beta_3 = 0$. If these test fail, then the Cobb-Douglas production function is not a good approximation. Increasing returns imply that $\beta_1 < 0$, decreasing returns imply $\beta_1 > 0$ and constant returns in production imply $\beta_1 = 0$.

When Cobb-Douglas is appropriate, we can retrieve the structural parameters of the production function from the regression estimates:

$$\alpha_K + \alpha_L = \frac{1-\beta_1}{\beta_1}, \tag{2.85}$$

$$\alpha_L = \beta_2 \frac{1-\beta_1}{\beta_1}, \tag{2.86}$$

$$\alpha_K = \beta_3 \frac{1-\beta_1}{\beta_1}, \tag{2.87}$$

$$T = \frac{1-\beta_1}{\beta_1}\left[\left(\frac{\beta_3}{\beta_2}\right)^{\beta_2} + \left(\frac{\beta_2}{\beta_3}\right)^{\beta_3} - \beta_0\right] \tag{2.88}$$

2.3.4 Special case: constant average and marginal costs

The general cost function that is derived above allows us to derive simpler cases by making additional assumptions. When regression shows that it is plausible to assume that there are constant returns to scale in production, $\alpha_K + \alpha_L = 1$, and the cost function simplifies and can be written as:

$$C(K^*, L^*) = \frac{1}{T} Q r^{\alpha_K} \omega^{1-\alpha_K}. \tag{2.89}$$

Average costs are then given by $\frac{C(K^*,L^*)}{Q}$ and are constant given the assumptions made in this subsection. *Marginal costs* are given by the first derivative of the optimal costs with respect to the target quantity and are equal to the average costs:

$$MC = AC = \frac{\partial C(K^*, L^*)}{\partial Q} = \frac{C(K^*, L^*)}{Q} = \frac{1}{T} r^{\alpha_K} \omega^{1-\alpha_K}. \tag{2.90}$$

This simple expression accounts for the substitution of inputs of a firm and allows us to analyse exogenous changes in marginal costs due to changes in technology (via T), strength of substitution between inputs (via α_K) and changing input prices (via r and ω). Because—given the assumptions—optimal costs are linear in the target quantity, marginal costs for the firm are constant per unit of production. However, this does not imply that marginal costs cannot be different for different firms as the input elasticities and the technology parameter might be different for different firms.

2.3.5 Empirical examples

This section discusses two examples. First, we discuss one of the key sustainability innovations in the energy market: the production of solar panels. A second example is given for the production function of housing. Both examples are related to the key sustainability issues of climate change and spatial planning.

When there are exogenous improvements in production technology over time it is expected that the technology parameter T will increase over time. Indeed, such a result has been found in the literature on solar panels. Pillai (2015) finds an average 21% annual reduction in average production costs over the time period 2005–2012.[9]

For solar panels, it is expected that capital, materials and labour are the main inputs. According to Pillai (2015) (p.289), capital and materials are the main inputs and firms substitute these inputs in order to optimize production. The input prices for materials and capital therefore make an important contribution on the average and marginal costs of solar panels (see Pillai (2015), Table 4). For example, the decreasing input price for polysilicon material inputs contributes to 7% of the cost efficiency improvement over the time period 2005–2012. This result shows that input prices can be key drivers for average cost reductions of sustainable energy solutions and that the discussion on producer behaviour is related to recent research on sustainability innovations.

The second example that we discuss is related to sustainable spatial planning and more specifically to the production function of housing which is studied by Epple, Gordon, and Sieg (2010). They estimated the production function for new housing for the greater metropolitan area of Pittsburgh. Their Cobb-Douglas function includes land inputs R and non-land inputs M such as building materials and capital (note that in their paper they use L for land inputs, but we already used L for labour, so use R instead). Firms can for example choose to substitute land and materials by building houses with multiple floors. $epple2010new estimate the following production function:[10]

$$f(R, M) = 1.38R^{0.144}M^{0.856}. \tag{2.91}$$

The elasticity of production with respect to land is therefore given by 0.144 which means that for a 1% increase in the amount of land inputs, housing production increases with 0.144%. Because the elasticities add up to 1, there are constant returns to scale. The marginal cost function and the average cost function for the metropolitan Pittsburgh area are constant and can be derived as a function of the input prices for land (p_R) and non-land inputs (p_M):

$$MC = AC = \frac{\partial C(K^*, L^*)}{\partial Q} = \frac{C(K^*, L^*)}{Q} = \frac{1}{1.38}p_R^{0.144}p_M^{1-0.144}. \tag{2.92}$$

Usually the price of land is different for different locations resulting in marginal costs of building houses which are potentially different over space. Another recent study by Combes, Duranton, and Gobillon (2021) estimate the production function for newly built single-family homes for France:

[9] They also show that average costs are decreasing in quantity levels.

[10] It is not entirely clear from their paper what the dimension of the production function is and what the constant 1.38 means.

> "or newly built single-family homes in France, the production function for housing is close to constant returns and is well, though not perfectly, approximated by a Cobb-Douglas function with a capital elasticity of 0.65." (Combes et al. 2021, abstract)

They show that the Cobb-Douglas production function gives a reasonable (but not perfect) approximation and that the capital elasticity is estimated at 0.65. The elasticity estimate for land is around 0.35, as the paper shows constant returns is a reasonable approximation. This elasticity estimate for land is higher than the value found for Pittsburgh. These two examples show how closely linked recent empirical results are to the theory of production functions and cost functions.

The theory of production functions can also be linked to recent developments on circular economy. In the context of housing production there can be shortage of building materials. It is expected that this will impact the input prices for building materials and thereby leads to input substitution and higher marginal production costs of housing. When the housing market works in a more circular way, materials of old houses will be re-used and re-offered on the market. This implies that the total supply of housing production inputs will be higher when competitive circular material use is possible.

2.3.6 From firm costs to market inverse supply and supply

The next step is to relate the firm marginal costs to the market inverse supply curve. Again, the aggregation of marginal costs curves into an inverse supply curve is not trivial, but the intuition is similar as with consumer analysis. Aggregation can be done by sorting heterogeneous firms according to their marginal costs per unit of the market good. This gives the market inverse supply curve $s(Q)$ for the market good which express what the price should be for which an amount Q of the market good is supplied. Marginal costs for the firm can be constant per unit of production, but this does not imply that there is no heterogeneity in marginal costs for different firms. For example, the firms might have similar input prices, but different input elasticities and technology parameters. The inverse of the inverse supply curve is the supply curve $Q^S(p)$, which shows how much is supplied at the market for a given level of the market price p. Empirical economists regularly estimate the price elasticity of supply which is defined as:

$$\kappa = \frac{\partial Q^S}{\partial p} \frac{p}{Q^{S(p)}}. \tag{2.93}$$

For a 1% increase in the market price, the market supply will change with κ%. When the slope of the supply function $\frac{\partial Q^S}{\partial p}$ is very large, the supply function is steeply increasing in the price and supply is elastic. Supply then responds strongly to price changes. An empirical example of supply elasticities can be

given using the work of Green, Malpezzi, and Mayo (2005) on housing supply in the US. The table in Figure 2.3 is from their paper and shows price elasticities of supply in the range of -0.30 (Miami) up to 29.9 (Dallas). These estimates are therefore widely varying over space.

Not all estimates are significant (confidence intervals are not reported). For Pittsburgh, the elasticity is estimated at 1.43 and significant at the 1% level. This implies that a 1% increase in the price of housing will lead to 1.43% increase in the housing supply in Pittsburgh. For Atlanta, housing supply is more elastic: a 1% increase in the price of housing leads to an estimated 21.6% increase in housing supply. Later we will see how elasticities are related to policies related to prices.

2.3.7 Producer surplus

For the analysis of producer surplus, it is assumed that all firms ask the same price for their product.[11] Total revenues in the market are then given by p^*Q^*. Given these assumptions, there will be firms in the market who make a profit. Low-marginal cost firms do not compete other firms out of the market by offering lower prices and firms do not propose customer specific prices. In the absence of fixed costs, total costs in the market are given by the area under the supply curve. Producer surplus can be mathematically described as follows:

$$PS = p^*Q^* - \int_0^{Q^*} s(Q)dQ. \tag{2.94}$$

Using information on the marginal cost curve we can derive producer surplus using integral calculus. Suppose the inverse supply curve is independent of the market quantity and given by:

$$s(Q) = \frac{B}{T} r^{\alpha_K} \omega^{1-\alpha_L}. \tag{2.95}$$

Here, B is a proportional constant that captures the aggregation of firms.

Substituting in the expression for producer surplus then results in:

[11] Here it becomes a bit tricky when we have heterogeneous firms as there is no real price competition going on despite the fact that low marginal costs firms can offer lower prices than high marginal cost firms. We deliberately ignore profit maximizing behaviour, imperfect substitutability between products among other issues. In theory, free entry and competition would lead to marginal costs equal to the *marginal costs of the firm with the lowest costs* and profits equal to 0. Despite all these issues, the model might serve as a reasonable approximation for medium term analysis of economic surplus.

TABLE 1—HOUSING SUPPLY ELASTICITIES BASED UPON LAGGED PRICE CHANGE

MSA	Elasticity (with significance)
Dallas	29.9
Tampa-St. Petersburg	27.4
Phoenix	21.7
Atlanta	21.6*
Charlotte	17.0*
Oklahoma City	13.7*
Columbus	13.5
Houston	12.8
Denver	11.4
Indianapolis	11.0**
Kansas City	11.0**
Grand Rap	10.8**
Cincinnati	8.25**
Tulsa	8.25
San Antonio	8.23
Portland	7.14**
St. Louis	6.89**
Akron	6.64**
Memphis	5.63
Baltimore	5.52
Birmingham	5.33**
San Diego	5.33
Rochester	5.25**
Syracuse	5.01**
Detroit	4.74**
Salt Lake City	4.69
Orlando	4.50
Milwaukee	4.45**
Minneapolis	4.21
Los Angeles	3.73*
Philadelphia	3.09**
Hartford	2.85**
Buffalo	2.84**
Chicago	2.48
Fort Lauderdale	2.23
Providence	2.10**
Boston	1.77**
Albany	1.55
Pittsburgh	1.43**
Toledo	0.83
San Jose	0.33
San Francisco	0.14
New Orleans	0.06
Honolulu	−0.10
Miami	−0.30

* Statistically significant at the 5-percent level.
** Statistically significant at the 1-percent level.

Figure 2.3: Price elasticities of housing supply

$$PS = p^*Q^* - \int_0^{Q^*} \frac{B}{T} r^{\alpha_K} \omega^{1-\alpha_L} dQ = p^*Q * \left[Q \frac{B}{T} r^{\alpha_K} \omega^{1-\alpha_L} + z \right]_0^{Q^*} \tag{2.96}$$

$$= p^*Q^* - Q^* \frac{B}{T} r^{\alpha_K} \omega^{1-\alpha_L}. \tag{2.97}$$

This requires input on the equilibrium price and quantity, the technology parameter, the input prices and the production elasticity of substitution.

2.4 Analysis of economic surplus

2.4.1 Introduction

Now that we have completed the analysis of the choice behaviour of consumers and consumer value in markets (section 2), and the choice behaviour of producers and producer value in markets (section 3), we continue our analysis on the second pillar of economics: the analysis of equilibrium. The assumption of equilibrium ensures that the interaction of consumers and producers in the market is properly accounted for and gives new information for the researcher. Not all consumers will buy the market good and not all potential firms will participate in selling goods on the market. Some might be interested but will not receive sufficient gains from trade.

2.4.2 Equilibrium

Demand and supply interact which may lead to equilibrium in markets. The equilibrium intuition is governed by the following two equalities:

$$p(Q^*) = p^* = s(Q^*). \tag{2.98}$$

The first equality $p(Q^*) = p^*$ states that the marginal benefits of the equilibrium consumer at Q^* are equal to the equilibrium price. Customers at the left of Q^* will buy the good because their marginal WTP is higher than the equilibrium price. Customers at the right side of Q^* will not participate in the market, because their marginal WTP is lower than the equilibrium price.

The second equality $p^* = s(Q^*)$ states that the equilibrium price should be equal to the marginal costs of production for the equilibrium quantity produced. Firms at the left of Q^* have lower marginal costs than the equilibrium price and will therefore continue to sell their goods. Firms at the right of Q^* have higher marginal costs than the equilibrium price and will therefore choose not to produce the good.

Table 2.1: Demand and supply elasticities and model parameters

Case description	Inverse demand
Case I: Exponential inverse demand and exponential inverse supply	$p(Q) = Ae^{-\alpha Q}$ Parameter assun
Case II: Constant elasticity inverse demand and inverse supply	$p(Q) = AQ^{-\alpha}$ Parameter assum
Case III: Logged inverse demand and inverse supply	$p(Q) = A - \alpha \ln Q$ Parameter as
Case IV: linear inverse demand and inverse supply	$p(Q) = A - \alpha Q$ Regression (line

Table 2.2: Calibrated model parameters for commonly used inverse demand and inverse supply functions

Case description	Inverse demand
Case I: Exponential inverse demand and exponential inverse supply	$p(Q) = Ae^{-\alpha Q}$ Parameter assun
Case II: Constant elasticity inverse demand and inverse supply	$p(Q) = AQ^{-\alpha}$ Parameter assum
Case III: Logged inverse demand and inverse supply	$p(Q) = A - \alpha \ln Q$ Parameter as
Case IV: linear inverse demand and inverse supply	$p(Q) = A - \alpha Q$ Regression (line

2.4.3 Calibration of equilibrium

How to fit the conceptual model to a real market in order to do justice to the third pillar: empiricism? That is the challenge of calibration in this section. Using the equilibrium equations, it is possible to recover two unknown parameters of the demand and supply function. If one also has information about the equilibrium elasticities, it is possible to calibrate four model parameters. This is the reason that the estimation of equilibrium elasticities is an important task of empirical economists. In the previous section we have seen estimates for these elasticities for different kinds of market goods. The tables below can be used for calibration of the (inverse) market demand and supply curves given this knowledge on the estimated (equilibrium) elasticities.

Table 2.2 gives the calibrated parameters for the inverse demand and inverse supply functions based on information on the equilibrium elasticities and the equilibrium price and quantity. Equilibrium elasticities can be obtained from the literature or from an empirical analysis of demand and supply data. Equilibrium prices and quantities can be observed for the market that you are interested in.

The parameters A and B can be calibrated using the identities for the inverse demand and inverse supply at the equilibrium. The parameters a and b can be calibrated using the equations we have for the price elasticities for specified demand and supply functions (see Table 2.1). To illustrate how a market can be calibrated using elasticities from the literature we provide an example of case I: exponential inverse demand and inverse supply functions. Table 2.3 shows

Table 2.3: Input for the calibration (numbers are illustrative and not from empirical studies)

Variable explanation	Value
Price elasticity of demand ϵ^*	-0.7
Price elasticity of supply κ^*	1.1
Equilibrium price per piece p^*	2.9 euro
Equilibrium quantity per year Q^*	$200,000$ pieces per year

Table 2.4: Calibrated model parameters

Variable explanation	Value
Parameter of the inverse demand function α	$\alpha = \frac{1}{\epsilon^*}\frac{1}{Q^*} = \frac{1}{0.8}\frac{1}{200{,}000} = \frac{}{16}$
Parameter of the inverse demand function A (maximum MWTP in the sample)	$A = p^* e^{\alpha Q^*} = 2.9e^{\frac{200{,}000}{160{,}000}} \approx$
Parameter of the inverse supply function b	$b = \frac{1}{\kappa^*}\frac{1}{Q^*} = \frac{1}{1.1}\frac{1}{200{,}000} = \frac{}{22}$
Parameter of the inverse supply function B (minimum marginal costs in the sample)	$B = p^* e^{bQ^*} = 2.9e^{-\frac{200{,}000}{220{,}000}} \approx$

information we have about the equilibrium elasticities for vegan nuggets.

Using the first row of Table 2.1 we then calibrate the parameters of the inverse demand and the inverse supply function. These are given in Table 2.4:

Using the calibrated inverse demand and inverse supply function one can derive the consumer benefits, the producer costs and economic surplus as we will see in the next section. The important insight here is that model parameters can be calibrated using empirical observations of equilibrium prices and quantity and empirical estimates of the equilibrium elasticities. This provides the bridge between theoretical and conceptual models and empirical observations.

2.4.4 Analysis of economic surplus

Economic surplus (ES) generated in the market is given by the sum of consumer surplus and producer surplus. It seeks to combine pillar (iii) with a particular view on normativity (pillar (iv)): for the calculation of economic surplus it is assumed that:

1. The choices of market actors express their value attached to the market products;
2. Choice value can be fully counted as economic surplus because the choices of market actors are 'good'.

We come back to these assumptions in Chapters @ref(erroreconsurplus and 4.

Mathematically, economic surplus is defined by:

$$ES = CS + PS = \underbrace{\int_0^{Q^*} p(Q)dQ - p^*Q^*}_{\text{Consumer surplus}} + \underbrace{p^*Q^* - \int_0^{Q^*} s(Q)dQ}_{\text{Producer surplus}} \tag{2.99}$$

$$= \underbrace{\int_0^{Q^*} p(Q)dQ}_{\text{Total consumer benefits}} - \underbrace{\int_0^{Q^*} s(Q)dQ}_{\text{Total consumer costs}} . \tag{2.100}$$

The first line shows the first way to express economic surplus: it is the sum of consumer surplus and producer surplus. Consumer surplus is the area under the inverse demand curve minus the consumer spending on the market good. Producer surplus is the producer revenues minus the total producer costs. Producer costs are equal to the area under the inverse supply curve (because it is assumed that fixed costs are 0).

The advantage of this way of expressing economic surplus is that it is clear how the surplus is divided between consumers and producers. Sometimes national governments are only interested in the consumer surplus for a particular product for inhabitants of their own country and not in the producer surplus of firms in other countries. Therefore, it might be useful to report both CS and PS and to state where the benefits fall.

There is a second way to express economic surplus, which is given by the second line of the equation. Because payments of consumers are revenues for producers, the second and the third term cancel out resulting in an equation which states that consumer surplus is equal to the consumer benefits minus the producer costs. Interestingly, one does not have to observe the equilibrium prices in order to calculate economic surplus as long as one knows the inverse demand and inverse supply curves and the equilibrium quantity.

At this point it is also useful to investigate the point for which economic surplus is optimal. Differentiating economic surplus with respect to the equilibrium quantity gives:

$$\frac{\partial ES}{\partial Q^*} = p(Q^*) - s(Q^*) = 0 \tag{2.101}$$

This is equal to the equilibrium condition which states that marginal benefits are equal to marginal costs. The expression shows that economic surplus in the market is optimal when marginal consumer benefits are equal to marginal costs in equilibrium. In the absence of other costs and considerations, one can conclude with Adam Smith from this equation that free trade will result in the maximum economic surplus. However, one should keep in mind that together with the assumptions on rationality, equilibrium is imposed rather than empirically supported using data analysis and that particular pillar (iv) assumptions on normativity are made that might not hold in practice.

Integral calculus can be employed to derive economic surplus for particular specifications of the inverse demand and the inverse supply curve. We will illustrate this for an example. Suppose the inverse supply curve is exponentially increasing in quantity and given by:

$$s(Q) = Be^{bQ}, \quad B > 0, \quad b > 0. \tag{2.102}$$

Furthermore, assume that the inverse demand curve is exponential and given by:

$$p(Q) = Ae^{-\alpha Q}, \quad A > 0, \quad \alpha > 0. \tag{2.103}$$

Substitution in the expression for economic surplus gives:

$$\begin{aligned} ES = & \underbrace{\int_0^{Q^*} Ae^{-\alpha Q} dQ}_{\text{Total consumer benefits}} - \underbrace{\int_0^{Q^*} Be^{bQ} dQ}_{\text{Total consumer costs}} && (2.104) \\ = & \left[-\frac{1}{\alpha} Ae^{-\alpha Q} + z \right]_0^{Q^*} - \left[\frac{1}{b} Be^{bQ} + w \right]_0^{Q^*} && (2.105) \\ = & -\frac{1}{\alpha} A \left(e^{-\alpha Q^*} - 1 \right) - \frac{1}{b} B \left(e^{bQ^*} - 1 \right), && (2.106) \\ = & \frac{1}{\alpha} A \left(1 - e^{-\alpha Q^*} \right) - \frac{1}{b} B \left(e^{bQ^*} - 1 \right). && (2.107) \end{aligned}$$

This shows that economic surplus in the market can be written as a function of the observed equilibrium quantity and the structural parameters of the inverse demand and the inverse supply function. In the previous sub-section, we calibrated these structural parameters using information on equilibrium prices and equilibrium demand and the equilibrium demand and supply elasticities (see Table 2.4). We can now use these parameters to derive the consumer benefits, the producer costs and the total economic surplus in the market for vegan nuggets. Table 2.5 provides an extension of Table 2.4 and shows the results. We leave it to the reader to develop examples for the other cases 2–4.

2.5 External costs and economic surplus

2.5.1 Introduction

In the previous section we investigated mathematically the economic surplus that markets generate. This economic surplus is generated assuming free trade possibilities and equilibrium and is only related to the decisions of consumers and producers. It therefore does not account for potential other costs and benefits that fall outside the market and does not include other considerations related to rationality and normativity.

Table 2.5: Economic surplus (example)

Variable explanation	Value
Consumer benefits (euros per year)	$\frac{1}{\alpha}A[1-e^{-\alpha Q^*}] = 160,000 \times 10.12 \times [1-e^{\frac{200,000}{160,000}}] = 1,155,519$ eu
Producer costs (euros per year)	$\frac{1}{b}B(e^{bQ^*}-1) = 220,000 \times 1.17 \times (e^{\frac{200,000}{220,000}}+1)$ 380,956 eurors
Consumer surplus (euros per year)	$CS = CB - CC$ $\frac{1}{\alpha}A[1-e^{-\alpha Q^*}] - p^*Q^* = 1,155,519 - 580,000$
Producer surplus (euros per year)	$PS = PB - PC$ $p^*Q^* - \frac{1}{b}B(e^{bQ^*}-1) = 580,000 - 380,956 = 1$
Economic surplus (euros per year)	$ES = CB - PC = 774,563$ euros $= CS + PS = 774,563$ euros

More specifically, and related to sustainability, a challenge for the model developed in the previous section can be environmental costs that fall outside of the market. Examples of negative *external costs* of meat production/consumption for the environment are CO2 pollution, animal welfare, soil erosion, water pollution and air pollution (see Godfray et al. 2018). External costs provide euro values of these negative impacts in order to make the impacts comparable with the market surplus of consumers and producers.

Economists have argued that these costs should be added to economic surplus in order to give an honest account of the surplus that markets generate. This addition of *external costs* entails a particular pillar (iv) assumption on responsibility which will be discussed in Chapter 4. In order to account for external costs (EC), we add the term $e(Q^*)$ to the economic surplus. Economic surplus is defined as:

$$ES = CS + PS - EC \tag{2.108}$$

$$= \underbrace{\int_0^{Q^*} p(Q)dQ}_{\text{Total consumer benefits}} - \underbrace{\underbrace{\int_0^{Q^*} s(Q)dQ}_{\text{Total consumer costs}} - \underbrace{e(Q^*)}_{\text{Total external costs}}}_{\text{Total social costs}}. \tag{2.109}$$

Besides this, we also define the marginal external costs $MEC(Q) = \frac{\partial e(Q)}{\partial Q}$, which show—loosely speaking—with how much total external costs change for a change in the quantity. The total social costs are then defined as the producer costs and the external costs. The marginal social costs are given by the first derivative of the total social costs. This first derivative is equal to $s(Q) + MEC(Q)$ which signifies that marginal external costs are economically speaking just another cost component above the producer costs. Again, we can investigate optimality of the extended economic surplus function by investigating the first derivative:

$$\frac{\partial ES}{\partial Q^*} = p(Q^*) - s(Q^* - MEC(Q^*) = 0. \tag{2.110}$$

This condition is different from the condition in the previous section because of the added external costs. The condition shows that—given the model assumptions—optimal economic surplus cannot be reached without intervention in the market because the behaviour of consumers and producers leads to $p(Q^*) - s(Q^*) = 0$. When marginal external costs are positive, the derivative is negative, showing that improvements in economic surplus can be obtained when demand decreases. In short, this means that when positive (marginal) external costs are ignored by market actors the prices for the market goods are too low. When there are marginal external benefits, $MEC(Q^*) < 0$, market prices are too high.

2.5.2 General analysis of the consumer externality tax

One way to deal with environmental externalities is to set a consumer tax. Mathematical analysis can be employed to investigate this consumer tax. The regulator sets a tax τ and this will result in a new equilibrium price p^{*R} and a new equilibrium demand Q^{*R}. The tax will impact the equilibrium equation for consumers as now their marginal benefits should be higher or equal to the equilibrium price plus the tax. For the equilibrium consumer we have:

$$p(Q^{*R}) = p^{*R} + \tau. \tag{2.111}$$

For the equilibrium producer, there is no change and the marginal costs should be equal to the equilibrium price:

$$s(Q^{*R}) = p^{*R}. \tag{2.112}$$

The price of the producer can therefore be treated as given for the regulator. The regulator seeks to optimize welfare given the first constraint and keeping in mind that in the regulated equilibrium inverse supply is equal to the supply price. The Lagrangian is given by:[12]

$$L = \underbrace{\int_0^{Q^{*R}} p(Q)dQ}_{\text{Total consumer benefits}} - \underbrace{\int_0^{Q^{*R}} s(Q)dQ - e(Q^{*R})}_{\text{Total social costs}} + \underbrace{\lambda(p(Q^{*R}) - p^{*R} - \tau)}_{\text{Equilibrium constraint}}. \tag{2.113}$$

The first-order conditions are given by:

[12]This methodology might look a bit extensive as one can directly obtain the result using the optimal welfare condition and plugging in the equilibrium condition. Nevertheless, it helps to move to more complicated analyses. See for example: Verhoef, Nijkamp, and Rietveld (1996).

$$\frac{\partial L}{\partial Q^{*R}} = p(Q^{*R}) - s(Q^{*R}) - MEC(Q^{*R}) + \lambda \frac{\partial p(Q)}{\partial Q} = 0 \tag{2.114}$$

$$\frac{\partial L}{\partial \lambda} = p(Q^{*R}) - p^{*R} - \tau = 0 \tag{2.115}$$

$$\frac{\partial L}{\partial \tau} = \lambda = 0 \tag{2.116}$$

Substituting the second and third condition in the first condition as well as $s(Q^{*R}) = p^{*R}$ gives:

$$p^{*R} - \tau - p^{*R} - MEC(Q^{*R}) + 0\frac{\partial p(Q)}{\partial Q} = 0 \tag{2.117}$$

This results in the following expression for the consumer tax:

$$\tau = MEC(Q^{*R}) \tag{2.118}$$

The consumer tax should therefore be equal to the marginal external costs *in the regulated equilibrium*. This is a well-known result in the field of public and environmental economics. The tax ensures that in equilibrium the optimality condition for maximizing economic surplus (including external costs) holds. In the absence of external costs, the tax will be 0.

When marginal external costs are constant per unit of consumption, the tax will be independent of the equilibrium quantity. Total external costs can then be defined as $e(Q) = mQ$ and the consumer tax is given by:

$$\tau = m \tag{2.119}$$

An example for meat production can be given based on the work of Van Drunen, Van Beukering, and Aiking (2010).

> The total (editor: marginal) external costs for conventional pork are estimated to be at least **€2.06 per kg** for an average consumer price of €6.69 (PVE, 2009), or 31%. In this, animal welfare is the main factor, followed by biodiversity, animal disease and climate change.

Van Drunen, Van Beukering, and Aiking (2010) estimate the marginal external costs at 2.06 euros per kilogram for conventional pork which is about 1/3 of the current equilibrium price. Interestingly, animal welfare is the largest component in the external costs. Given the analytical analysis this would result in a consumer tax of $m = 2.06$ euros per kilogram for conventional pork. From the analytical model we know this will improve economic surplus. Nevertheless, it is interesting to analyse with how much the surplus will change.

2.5.3 Consumer tax for specific inverse demand and supply functions

This general analysis can be linked to concrete inverse demand and inverse supply functions. Suppose these are given by:

$$p(Q) = Ae^{-\alpha Q}, s(Q) = MC \tag{2.120}$$

Assume that the marginal external costs are constant per unit and therefore:

$$e(Q) = mQ. \tag{2.121}$$

From the analysis in the previous section we already know that the tax in the regulated equilibrium is equal to the marginal external costs m.

We want to study what the predicted resulting equilibrium demand is in the tax-regulated optimum. In order to investigate this research question, we will look at the equilibrium conditions for consumers and producers:

$$Ae^{-\alpha Q^{*R}} = p^{*R} + \tau = p^{*R} + m \tag{2.122}$$

For the equilibrium producer, there is no change and the marginal costs should be equal to the equilibrium price:

$$MC = p^{*R} \tag{2.123}$$

This gives us two equations and two unknowns. Substituting the second equation into the first gives:

$$Ae^{-\alpha Q^{*R}} = MC + m \tag{2.124}$$

The assumption of constant marginal costs helps to derive analytical results. Without this assumption one often has to solve the equation numerically. Solving for the equilibrium quantity in the regulated equilibrium gives:

$$e^{-\alpha Q^{*R}} = \frac{MC + M}{A} \tag{2.125}$$

$$-\alpha Q^{*R} = \ln\left[\frac{MC + M}{A}\right] \tag{2.126}$$

$$Q^{*R} = -\frac{1}{\alpha}\ln\left[\frac{MC + M}{A}\right] = \frac{1}{\alpha}\ln\left[\frac{A}{MC + M}\right] \tag{2.127}$$

This requires the intuitive additional assumption that the marginal social costs $MC + m$ should be lower than the maximum willingness to pay of the first consumer A.

Because marginal costs are assumed to be constant, the equilibrium price remains at MC and therefore $MC = p^{*R}$, in the regulated equilibrium. With these results one can estimate the change in economic surplus in the regulated and in the unregulated equilibrium by substituting the expression for the quantity in the economic surplus function. We leave this as an exercise for the reader.

Some important assumptions are made in the analysis. First, the model suggests that the only way the consumer can avoid external costs is by not choosing to buy the product. Despite potential knowledge that consumers have, external costs are not incorporated somewhere in the decision of the consumer to choose the market good. Whether this is a reasonable account of behavioural decisions requires further investigation using for example surveys and interviews. It can be that the consumer already adds the external costs to the price that is paid when making the purchase decision, even when there is no externality tax. As consumers are increasingly aware of the impacts of consumption on climate this might be realistic. If this would be the case, the equilibrium that we as economists observe is the socially optimal equilibrium and externality taxes would lead to lower welfare. Suppose consumer partly take the external costs eQ into account when buying their good. Each consumer adds additional costs $\theta e(Q^{*R})$ to the supply price when making the purchase decision with $0 \leq \theta \leq 1$. This will impact the equilibrium condition but not the value derived from the consumption and production of the market good:

$$L = \underbrace{\int_0^{Q^*R} p(Q)dQ}_{\text{Total consumer benefits}} - \underbrace{\int_0^{Q^*R} s(Q)dQ - e(Q^{*R})}_{\text{Total social costs}} + \underbrace{\lambda(p(Q^{*R}) - p^{*R} - \theta e(Q^{*R}) - \tau)}_{\text{Equilibrium constraint}}. \tag{2.128}$$

Using this Lagrangian one can derive the optimal tax:

$$\tau = (1 - \theta)MEC(Q^{*R}) \tag{2.129}$$

This expression confirms the intuition: when individuals fully consider marginal external costs when purchasing the good, $\theta = 1$, and the tax should be 0. When individuals ignore marginal external costs in their purchase decision, $\theta = 0$, and the tax should be equal to the marginal external costs. Surveys can be used to assess which share of external costs is already accounted for when individuals make their decisions.

Second, sometimes there are other ways in which consumers take responsibility for external costs, such as real monetary compensation of climate damage. For these cases it is expected that the consumers trade the compensation costs with a direct

utility benefit related to taking responsibility. When the compensation costs that are paid really reduce damage, they can be counted as benefits in the economic surplus function. The externality tax is then lower compared to the one proposed above. A way to deal with this is to include only *uncompensated* external costs in the economic surplus function. When θ is the share of compensated external costs this leads to the same expression for the tax as for the previous example.

Third, it is useful to interpret $e(Q^*)$ in more detail. Because our analysis is a partial equilibrium analysis, there can be two types of goods that generate externalities: the market good and the outside good. Strictly speaking total external costs are then a function of the consumption level of *both* goods and can be defined as $e(Q^*, G^*) = e(Q^*, Y - (p^* + \tau)Q^*)$. In practice, the introduction of a tax implies substitution of the market good to the outside good due to higher price levels, leading to higher consumption levels of the outside good. Furthermore, the tax itself has a direct impact on the consumption level of the outside good. Such a rise in consumption of the outside good can lead to additional external costs. Usually it is assumed that external costs of goods on all other markets are correctly priced. However, such an assumption is likely to be unrealistic.

When unregulated external costs are present for the outside good these can be properly accounted for in partial equilibrium analysis. When we assume that the externality costs for the market good and the outside good are additive the external costs can be defined as $e(Q^*) = e_Q(Q^*) + e_G(Y - (p^* + \tau)Q^*)$. Here, the first term is increasing in the market good consumption and the second term is decreasing when there are positive external costs for consumption of the outside good. The second term gives the external costs of rising income levels. For sustainability analysis, one can find estimates for the external costs of the outside good by relating spendable income levels to emissions of CO2 equivalents multiplied with the price of these CO2 equivalents. The analytical analysis is more complicated, but suggests that the recommended tax should be lower when positive marginal external costs of the outside good are present.

One of the implications of this discussion is that policy makers should target their regulations to markets which have higher marginal external costs than the marginal external costs of the average consumption basket of consumers. Otherwise there is the risk that a tax will lead to substitution of the market good to the outside goods which has higher marginal external costs.

Another lesson to be learned from this third point is that it is best to start with taxation of the most polluting market good. For example, the external costs for particular meat types that are close substitutes are higher than for others. For example, starting taxation with the most polluting meat type will therefore lead to lowest risks of regulatory failure when an externality tax is introduced.

On the other hand, one could also argue that externality taxes in other markets should also be introduced. However, it is questionable whether such a 'Pigouvian state' with all markets regulated with optimal taxes is politically feasible and

Table 2.6: Equilibrium prices and quantities with and without regulation assuming a tax that is proportional to p^{*R}

Case description	Inverse demand
Case I: Exponential inverse demand and exponential inverse supply	$p(Q) = Ae^{-\alpha Q}$ Parameter assun
Case II: Constant elasticity inverse demand and inverse supply	$p(Q) = AQ^{-\alpha}$ Parameter assum
Case III: Logged inverse demand and inverse supply	$p(Q) = A - \alpha \ln Q$ Parameter as
Case IV: linear inverse demand and inverse supply	$p(Q) = A - \alpha Q$ Regression (line

desirable.

2.5.4 Stylised solutions for the case when marginal external costs are proportional to the equilibrium price.

It is possible to obtain stylised analytical solutions when it is assumed that the marginal external costs are a proportion of the equilibrium price in the regulated equilibrium. This practical assumption is sometimes made for a quick approximation of welfare effects for applied analysis. We can then write $MEC = c_e p^{*R} = \tau$. The equilibrium conditions can then be written as:

$$p(Q^{*R}) = p^{*R} + \tau = p^{*R}(1 + c_e) \tag{2.130}$$

For the equilibrium producer, marginal costs should be equal to the equilibrium price:

$$s(Q^{*R}) = p^{*R}. \tag{2.131}$$

The challenge is the to find a solution for Q^{*R} for the following equation:

$$p(Q^{*R}) = s^{*R}(1 + c_e) \tag{2.132}$$

Analytical expressions can then be obtained for common combinations of inverse demand and inverse supply functions. Table 2.6 gives the specifications of inverse demand and inverse supply, the equilibrium values for the unregulated equilibrium, and the equilibrium values for the regulated equilibrium. The parameter values can be obtained using the calibration procedure discussed earlier.

2.5.5 Analysis of changes in economic surplus due to externality taxation

An externality tax results in changes in economic surplus. These changes can be summarized in three parts: changes in consumer surplus, changes in producer surplus and changes in external costs. Because taxes are assumed to be transfers and have no impact on the inverse demand curve, these drop out of the economic surplus function. Mathematically the change in economic surplus can be expressed as:

$$\Delta ES = ES(Q^{**}) - ES(Q^{*}) = \Delta CS + \Delta PS + \Delta EC \tag{2.133}$$

$$= -\int_{Q^{**}}^{Q^{*}} p(Q)dQ + \int_{Q^{**}}^{Q^{*}} s(Q)dQ + e(Q^{*}) - e(Q^{**}) \tag{2.134}$$

Because of lower equilibrium demand, consumer surplus decreases, producer costs decrease and external costs also decrease.

An analytical example can be provided using a Cobb-Douglas based inverse demand curve and constant marginal costs of production. Assume external costs are constant per unit and equal to $e(Q) = mQ$. The change in economic surplus can then be written as:

$$\Delta ES = -\int_{Q^{**}}^{Q^{*}} \frac{A}{A+B}\frac{1}{Q} + \int_{Q^{**}}^{Q^{*}} MCdQ + m(Q^{*} - Q^{**}) \tag{2.135}$$

$$= -\frac{A}{A+B}\ln\left[\frac{Q^{*}}{Q^{**}}\right] + MC(Q^{*} - Q^{**}) + m(Q^{*} - Q^{**}) \tag{2.136}$$

This requires the calculation of the new equilibrium quantity after the tax has been implemented. The equilibrium tax will be equal to the marginal external costs m. The supply price in the regulated equilibrium will not change and will be equal to MC. The equilibrium quantity after the tax is implemented is then given by:

$$p^{**} + \tau = \frac{A}{A+B}\frac{1}{Q^{**}} \tag{2.137}$$

$$MC + m = \frac{A}{A+B}\frac{1}{Q^{**}} \tag{2.138}$$

$$Q^{**} = \frac{A}{A+B}\frac{1}{MC+m} \tag{2.139}$$

The equilibrium quantity without regulation is given by the demand function:

$$Q^* = \frac{A}{A+B}\frac{1}{p} = \frac{A}{A+B}\frac{1}{MC} \tag{2.140}$$

Substituting these expressions in the equation for the change in welfare gives:

$$\Delta ES = \frac{A}{A+B}\ln\left[\frac{MC+m}{MC}\right] + (MC+m)\frac{A}{A+B}\left(\frac{1}{MC} - \frac{1}{MC+m}\right) \tag{2.141}$$

$$= \frac{A}{A+B}\ln\left[\frac{MC+m}{MC}\right] + \frac{A}{A+B}\ln\left[\frac{MC+m}{MC} - 1\right] \tag{2.142}$$

$$= \frac{A}{A+B}\ln\left[\frac{MC+m}{MC}\right] + \frac{A}{A+B}m \tag{2.143}$$

$$= \frac{A}{A+B}\left(\ln\left[\frac{MC+m}{MC}\right] + m\right) \tag{2.144}$$

This can be illustrated using an example. Suppose that the marginal costs of production of a particular type of meat are equal to 6 euro per kilogram. Consumption is equal to 2 million kilograms per year. Furthermore, assume that the external costs are estimated at 2 euros per kilogram. From the equilibrium condition we have $\frac{A}{A+B} = Q * MC = 4,000,000$. Substituting these values gives the change in economic surplus when an optimal tax is implemented:

$$\Delta ES = 4,000,000\left(\ln\left[\frac{6}{6+2}\right] +\right) = 6,849,272 \text{ euros per year.} \tag{2.145}$$

This example has showed how—given the model assumptions—one can estimate a quantitative value for the change in economic surplus resulting from a consumer tax.

2.6 Conclusion

This chapter shows what the behavioural underpinnings of supply and demand in markets are and how supply and demand functions can be calibrated using elasticity estimates and observations of market prices and quantities. We have also showed how economic surplus in markets can be estimated based on the inverse demand and inverse supply functions. Furthermore, an application to consumer externality taxation is given. The models that were developed are stylised and can be used as a natural starting point for further analytical and empirical development.

Chapter 3

Behavioural Error and Economic Surplus

3.1 Introduction

This chapter investigates the implications of behavioural errors for the calculation of economic surplus. The sub-field of behavioural economics started around 1950 with the pioneering works of Herbert Simon and Ward Edwards. The aim was to develop a novel perspective on rationality and choice behaviour. Laibson and List (2015) define behavioural economics as follows:

> Behavioural economics uses variants of traditional economic assumptions (often with a psychological motivation) to explain and predict behaviour, and to provide policy prescriptions (p.385).

From their perspective, modern behavioural economics therefore can be viewed as an extension of the neo-classical perspective. Laibson and List (2015) suggest that behavioural economics refines three of the key core principles of classical economics (p.386). We add to this list how the fourth pillar 'normativity' changes when behavioural errors are present.

(i) *Optimization*: market actors try to choose their best (feasible) option. They employ heuristics or a decision goal function for choosing amounts (utility, costs etc.).
(ii) *Equilibrium*: market actors try to choose their best feasible option when interacting with others. This results in a behavioural equilibrium.
(iii) *Empiricism*: behavioural models need to be tested with data.
(iv) *Experienced based normativity*: experiences (not choices) of market actors generate value and experienced value of market actors is equal to value for society.

Compared to the neo-classical paradigm, behavioural economics changes the first core pillar 'optimization' by adding the word 'try'. In classical economic models it is usually assumed that market actors behave rationally: choices of market actors are the result of a rational optimization procedure in the sense that choices track utility value. Behavioural economics has operationalized this idea by replacing the direct utility function by other heuristics or by other kinds of direct utility functions. An example of a consumer heuristic is: buy Q of the market good when it rains and $Q + 2$ of the market good when it does not rain. Such a heuristic is independent of the valuation or utility considerations and is therefore difficult to connect to analysis of economic surplus without further assumptions. Examples of different conceptualization of utility functions are utility functions that account for biased processing of probabilities and loss aversion (Tversky and Kahneman 1992).

The second pillar states that equilibrium in markets might be affected by behavioural errors as market actors interact with each other and is a logical result of adding behavioural error at the first pillar. According to behavioural economists, consumers and producers seek to optimize, but make behavioural errors. This in turn has impacts on how market actors interact with each other. When consumers and producers make systematic mistakes, this can impact equilibrium prices and quantities, but more in-depth analysis is needed to show what the impacts of behavourial errors is on equilibrium outcomes.

The third pillar states that particular assumptions on the rationality of choices made by neo-classical economists should be backed up by empirical analysis. It has been argued by neo-classical economists that choices are made as if market actors are rational. Friedman (1953) wrote about this in his famous essay 'The methodology of Positive Economics':

> Consider the problem of predicting the shots made by an expert billiard player. It seems not at all unreasonable that excellent predictions would be yielded by the hypothesis that the billiard player made his shots **as if** he knew the complicated mathematical formulas that would give the optimum directions of travel, could estimate accurately by eye the angles, etc., describing the location of the balls, could make lightning calculations from the formulas, and could then make the balls travel in the direction indicated by the formulas. Our confidence in this hypothesis is not based on the belief that billiard players, even expert ones, can or do go through the process described; it derives rather from the belief that, unless in some way or other they were capable of reaching essentially the same result, they would not in fact be expert billiard players.
>
> It is only a short step from these examples to the economic hypothesis that under a wide range of circumstances individual firms behave as if they were seeking rationally to maximize their expected returns (generally if misleadingly called "profits") and had full knowledge of the data needed to succeed in this attempt; as if, that is, they knew

> the relevant cost and demand functions, calculated marginal cost and marginal revenue from all actions open to them, and pushed each line of action to the point at which the relevant marginal cost and marginal revenues were equal (Friedman, 1953).

Using the analogy with the expert billiard player, Friedman (1953) argues that there can be a complicated mathematical description of market choices that correctly *describes* behaviour even when market actors themselves are not aware of the complexities of their own behaviour. Why then, is additional behavioural economic analysis useful?

An important reason for behavioural economists to delve into the peculiarities of individual choices is that this might lead to a better *understanding* of the underlying mechanisms of choice behaviour. After all, scientists are allowed to be curious and are free to formulate hypotheses about the behaviour of individuals. A key purpose of science is to seek understanding and behavioural economists contribute to this with their empirical research.

For *prediction* of choices, understanding of the mechanisms of choice is not needed. 'As if' rationality can be a sufficient assumption as more complicated and realistic behavioural models might not always lead to better predictions. Prediction of choice can be tackled using econometric techniques (regression, machine learning) and does not require any underlying behavioural model of choice that describes the mechanisms. For example, one can investigate the change in the consumption of solar panels resulting from subsidies using regression techniques with demand as a dependent variable and subsidies as an independent variable. Utility maximisation and cost minimisation are not needed for these types of impact analyses but might give suggestions about what kind of control variables one can include in the regression.

The scope of the essays of Friedman is only positive or descriptive economics.[1] What Friedman (1953) did not discuss in his essay is that for normative policy analysis as if rationality is not a sufficient assumption. This is because these analyses require a model about how market actors determine value. For example, the question whether the subsidy for solar panels resulted in added value for society cannot be answered using models of prediction alone. Suppose subsidies result in a significant impact on the demand for solar panels. When effectiveness would be the only purpose the recommendation would be to have very high subsidies in order to increase the demand for solar panels.

To conclude our discussion about pillar (*iii*): the goal of the analysis (understanding, prediction, normative policy analysis) therefore determines whether *as if* rationality is an appropriate assumption. It also determines which research strategy is appropriate. When estimating economic value—and relate market outcomes to appropriate policy interventions—there is a normative component in the analysis, implying that the assumption of *as if* rationality is not sufficient.

[1] Without normative recommendations, the field of economics reduces to the field of econometrics.

Pillar (*iv*) shows that there is also a change in normative perspective when using behavioural economic models for policy analysis. The reason is that there can be a divergence between the value 'calculation' at choice and the experienced value of market actors. This has potential impacts on the calculation of economic value and policy recommendations. Customers can be more (less) satisfied with a product than they expected at the moment of choice. Firms might overestimate (underestimate) the marginal costs of production. These biases can have impacts on equilibrium outcomes and also on the estimation of economic surplus when one assumes that it is experienced value which counts.

To illustrate the importance of behavioural errors for the economics of sustainability, it is interesting to look at an example from the literature. Busse et al. (2015) find that:

> The choice to purchase a convertible or a four-wheel-drive is highly dependent on the weather at the time of purchase in a way that is inconsistent with classical utility theory. We consider a range of rational explanations for the empirical effects we find, but none can explain fully the effects we estimate.

Interestingly, Busse et al. (2015) show that even for expensive purchases in real markets—such as a new car—particular valuation biases stemming from the choice context can occur. In other words, the inverse demand curve that we have learned to estimate with statistical techniques potentially captures contextual valuations which might be irrelevant for the experienced value of the product. This leads to a divergence between choice value and experienced value.

The remainder of these notes are organised as follows. Section 2 discusses the conceptual argument against neo-classical valuation and also discusses common responses. Section 3 uses choice behavioural modelling to arrive at an inverse demand curve that captures behavioural error. This section helps to develop mathematical narratives of choice that capture some of the essential ingredients of behavioural economics without delving to deeply into psychological details. Section 4 develops a simple behavioural economic surplus model graphically and mathematically and shows the implications of behavioural errors on the calculation of economic surplus. Section 5 and 6 continue by analysing the implications for externality taxation by investigating changes in economic surplus. Section 7 concludes.

3.2 Arguments against neo-classical valuation and responses

3.2.1 Introduction

The quantitative and methodological implications of behavioural economics for the calculation of economic surplus in markets are often not so clear. Such a calculation is needed as an input for practical cost-benefit assessments that

seek to account for behavioural error. One of the methodological worries of cost-benefit practitioners who employ the micro-economic toolbox of the neo-classical economists is that behavioural errors lead to *systematic* flaws in policy recommendations because surplus calculations are incorrect. Before discussing these economic implications of behavioural error, we will discuss the qualitative argument in more detail. It has the following structure:

(A) Market actors make behavioural errors when valuing market goods;
(B) Therefore, the valuations derived from the choice-based neo-classical modelling approach are incorrect;
(C) As a result, neo-classical economic analyses of policies are mistaken as economic surplus is calculated in an incorrect way.

3.2.2 Discussion of Premise A

3.2.2.1 Introduction

Most researchers will grant that market actors make some behavioural errors in valuation and accept premise (A). However, five objections against the move from (A) to (B) are often raised. These objections are the 'learning objection' (sub-section B.2), the 'size objection' (sub-section B.3), the 'generalisability objection' (sub-section B.4), the 'knowledge objection' (sub-section B.5) and the 'competition objection' (sub-section B.6).

3.2.2.2 The learning objection

The first objection is the 'learning objection' which comes in three parts. Individuals learn to behave rationally because they:

(i) learn from repeated choices (individual learning);
(ii) learn from the behavioural errors and choices of others (social learning);
(iii) learn from information from other sources.

To illustrate learning with an example: A family household is likely to be able to estimate the amount of biological milk they need after some repetition. Or in the words of Friedman (1953):

> "they figure it out what their optimal quantity is without knowing any of the mathematical formulas that the researcher uses to describe behavioural choice. When the household buys too much milk there will be leftovers for the next week. When there is not enough milk the household has to substitute to water or other beverages."

Such an argument is less convincing for less regular choices such as the purchase of a car, a new house or a holiday. For these choices social learning might be a relevant aspect. For example, for tourism destination choices, reviews of campings and hotels can serve as an input for valuations and can lead to more optimal behaviour of consumers under the assumption that these reviews are

truthful. Customers can also learn from information provided by other sources such as science, customer protection agencies, media sources or the government.

In order to counter the learning objection, one has to show that in the presence of individual repetition, social interaction and information provision, individuals still make systematic mistakes. This is quite challenging to investigate empirically and often open for investigation for particular market goods.

3.2.2.3 The size objection

The second objection against premise (A) is the 'size objection' which challenges the size of the behavioural error. When behavioural errors are present these might not be quantitatively relevant from a personal or societal perspective as the inverse demand curve which captures the WTP at choice, has not been impacted substantially. This objection has been countered by results from studies that investigated the size of behavioural error. For example, Allcott, Lockwood, and Taubinsky (2019) predict in a study on sugar-sweetened beverages in the US that—given their assumptions—consumers consume 31% too much sugar-sweetened beverages on average.[2] This suggests that overconsumption in this market is substantial and quantitatively relevant. Furthermore, other studies have investigated food waste of consumers in order to estimate sub-optimal choice behaviour. These studies suggest that there is structural overconsumption of food.

3.2.2.4 Generalisability of empirical results

The third objection is more conceptual and is called the 'generalisability objection'. It is targeted at the generalisability of empirical and experimental results and also applies to neo-classical empirical results. Does the result of Allcott, Lockwood, and Taubinsky (2019) only holds for the US or not? Does the result also apply to other goods? When for some particular markets in some countries behavioural errors are quantitatively relevant this does not imply that behavioural errors are relevant for all markets. According to this third objection, premise (A) has to be studied on a case-by-case basis.

Researchers have used meta-analyses to partly counter this objection. These meta-analyses investigate and summarize the results of many empirical studies and seek to explain structural differences in results using variables such as population characteristics, study design and setup, year of the study etc. An example of such a meta study is the study of Cadario and Chandon (2020) who investigate the effectiveness of healthy eating nudges. They employ the definition of Leonard (2008):

> Nudges are defined by Leonard (2008) as "any aspect of the choice architecture that alters people's behavior in a predictable way (1) without forbidding any options or (2) significantly changing their

[2] See for example Visschers, Wickli, and Siegrist (2016).

> economic incentives. Putting fruit at eye level counts as a nudge; banning junk food does not."

Eating nudges therefore seek to change behavioural eating choices by changing the choice architecture without using economic incentives or traditional educational tools such as cooking workshops Cadario and Chandon (2020). When nudges show that people change their behaviour for a small change in choice architecture this signals something about potential behavioural error in their original choice. Cadario and Chandon (2020) find that the type of intervention has a substantial impact on the effectiveness of eating nudges. They consider three types of interventions: cognitive, affective and behavioural (p.468). Cognitive interventions seek to provide knowledge about food choices. These nudges have lower impacts than affective interventions which are targeted at feelings about food. Affective interventions in turn have lower impacts than behavioural interventions which are targeted at changes in behaviour (p.477). For example, Figure 3.1 in the paper shows that size enhancements such as smaller plates have the largest impacts Cadario and Chandon (2020).

Figure 3.1: Figure from @cadario2020healthy

Jachimowicz et al. (2019) is another example of a meta study on the impact of changing default options. Controlling for a number of variables related to channels and study characteristics they show that changing defaults is mainly effective in the consumer domain, but less in the environmental domain (which is the domain where policy makers seek to foster pro-environmental behaviour). Their result therefore suggests that changing the choice architecture is mainly beneficial for society when performed by producers of market goods rather than the government.

These meta study examples show that the generalisability objection can be partly mitigated by providing systematic overviews of studies on the empirical impacts of behavioural interventions and error. As the field of empirical behavioural economics is growing rapidly, it is expected that more meta studies will be available in the next decade.

3.2.2.5 Knowledge objection

The fourth objection is the '*knowledge objection*', which stems from an epistemic concern at the sight of the researcher: is it possible to know as a researcher whether a market actor makes a mistake in valuation? Is it really possible to determine the difference between experienced and decision utility? Premise (A) depends on the assumption that there is some 'true' experienced utility concept: one can only determine what a valuation error is by contrasting a descriptive utility function with another 'true' experienced utility function.

Consumers might have reasons for choice that are unobserved by the researcher. For example, it might be that households waste food because they are afraid to run out of food or want to keep food in the fridge for unexpected guests. Without asking about this, it is hard to infer from food choices only whether food waste is the result of behavioural errors in valuation. As Infante, Lecouteux, and Sugden (2016) write: [according to the behavioural economics view] "the inner rational agent is trapped in a psychological shell". The process of using the 'true' preferences rather than the choice-based preference of individuals for economic surplus analysis is called ''preference purification' (Infante, Lecouteux, and Sugden 2016) The assumption that there are true preferences to be satisfied has motivated the use of nudging approaches in order to help individuals to make 'better' choices.

Four methodological approaches have been offered in the literature to provide more evidence on the relationship between choice and experienced utility (Chetty 2015). The first approach is the 'direct measurement approach':

3.2.2.5.1 Direct measurement approach:measure experienced utility using self-reported happiness This approach uses complementary survey techniques to measure whether experienced utility deviates from choice utility. The researcher has to make a choice whether the directly measured experienced utility captures the true (optimal) utility or whether the choice-based utility is the true utility. Often it is unclear how the two relate. Classical economists usually prefer preferences based on revealed choices. Other researchers trust the self-reported happiness surveys more. Direct measurement approaches also can include qualitative reasons behind choices that might give useful information about the 'why' of choices (see for example Mouter, Koster, and Dekker 2021).

The second approach discussed by Chetty (2015) is the rational context approach:

3.2.2.5.2 Rational context approach: use revealed preference measurements in a context where agents are known to make choices that maximize their experienced utilities; (these are called 'sufficient statistics') When the utility is measured correctly in a context where the individual behaves rational, this experienced utility can be compared with measurements of decision utility in other contexts where behavioural errors are present in order to determine the size of behavioural errors. For example, when researchers know that there are contextual effects of the weather on car purchases (Busse et al. 2015), researchers might come up with a particular context, where weather does not influence preferences. Note however, that the researcher still has to make a choice what counts as a relevant context where the individual behaves rational.

The third approach is the structural modelling approach.

3.2.2.5.3 Structural modelling approach: build a structural model of the difference between decision and experienced utility. This approach uses mathematical modelling techniques with inputs from psychology and calibration based on empirical research. Examples of this approach will be given in Section II on behavioural modelling.

Allcott, Lockwood, and Taubinsky (2019) add a fourth strategy which is the rational reference group approach.

3.2.2.5.4 Rational reference group approach: consumption is compared to particular reference groups who have full information on the product and its impacts. This approach deals with informational deficiencies of market actors. In Allcott, Lockwood, and Taubinsky (2019), the reference group are the dietarians who are assumed to have full information on the impact of sugar sweetened beverages on health. Here the researcher has to make a choice which particular group can be viewed as the group having full information. This assumes that the choices of other individuals are not based on full information and that this is the reason for a divergence in consumption between the two groups. More sophisticated analysis can add control variables to explain differences in consumption levels between the groups related to for example income, education or age. This concludes the discussion on the fourth objection whether it is possible to know as a researcher that market actors make mistakes.

3.2.2.6 The competition objection

The fifth objection against premise (A) is the 'competition objection. Friedman (1953) highlights that when a firm makes systematic mistakes, it would be competed out of the market by firms that behave more rationally. Systematic monetary mistakes can be costly and thereby subject to the forces of the market. This objection has been countered by offering two types of responses.

First, such an argumentative strategy cannot deal with *systematic* mistakes of all

producers and is also invalid for the systematic mistakes that consumers make when calculating their marginal benefits. Consumers do not always pay the price for the mistakes they make when estimating their marginal benefits and are therefore not competed out of the market. Producers might all overestimate marginal costs of production leading to less entry in the market. Second, for durable and experiential goods such as cars, houses and holidays it is harder for consumers to track quantity left-overs in order to arrive at their optimal quantity. Third, not all mistakes are publicly available and known. Therefore, evaluation mistakes can remain in the presence of competition.

3.2.3 Discussion of premise (B)

3.2.3.1 Introduction

The second premise (B) in the argument concludes from the presence of behavioural errors that marginal valuations are incorrect. For consumers this would imply that the maximum willingness to pay for a particular market good is measured incorrectly. This maximum marginal willingness to pay (MWTP) is the ratio of the marginal utility of consumption and the marginal utility of income.

3.2.3.2 Examples of optimization with behavioural error that do not impact MWTP

A strategy to counter this objection is to give examples of optimization procedures that include behavioural errors (so premise (A) is granted), but do not result in a different MWTP (so premise (B) does not result).

Suppose the direct utility with behavioural error is defined as: $U^d(Q,G) = \epsilon \times U(Q,G)$, where $\epsilon > 0$. The marginal utilities are then scaled with a factor ϵ and therefore the ratio of marginal utilities is unaffected by behavioural error. When consumers make *proportional* errors in the evaluation of the direct utility, these mistakes will drop out and the MWTP is still correctly measured.

A similar result can be obtained for *additive* errors in direct choice utility: $U^d(Q,G) = U^e(Q,G) + \epsilon$. The error will drop out as the *marginal* utilities govern market behaviour and choice of consumers as these show up in the first-order conditions for utility maximization. For multiplicative and additive errors in direct utility, the MWTPs are therefore correctly estimated. Premise (A) is granted, but premise (B) does not follow.

A more general third example can be given by assuming that the decision utility function is given by $U^d(Q,G) = V(U^e(Q,G))$, where V is an increasing function of the experienced utility $U^e(Q,G)$. This implies that when experienced utility increases (decreases), choice utility also increases (decreases). Using the budget constraint, the direct choice utility can be written as:

$$U^d(Q,G) = V(U^e(Q, y - pQ)). \tag{3.1}$$

The first derivative for optimal choice of the market good is given by:

$$\frac{\partial U^d}{\partial V}(U_Q^e - pU_Q^e) = 0. \tag{3.2}$$

As $\frac{\partial U^d}{\partial V} > 0$ this results in:

$$p_d = \frac{U_Q^e}{U_G^e}. \tag{3.3}$$

This shows that for choice utility functions that are increasing in the experienced utility, the experienced MWTP coincides with the rational MWTP. Although this mathematical result is quite general, it is not entirely clear what it means from a behavioural perspective and whether it makes psychological sense. Usually the choice to buy the good is made first and the experience comes afterwards. What this discussion shows is that premise (B) does not necessarily follow from premise (A). Rational MWTPs can in principle be correctly measured in the presence of behavioural error.

3.2.4 Discussion of premise C

D. Discussion of premise C. Premise (C) can follow from premise (B) but requires additional analysis which will be provided in sections III and IV. Errors in evaluation of value can potentially lead to a discrepancy in the value used for decision making and the experienced value of consuming a good or service. Consumer and producer surplus can therefore be affected by behavioural errors.

When both consumers and producers make behavioural errors or when externalities are present the impacts on economic surplus are less obvious. Without further assumptions it is unclear whether behavioural errors impact economic surplus negatively or positively.

3.3 Behavioural choice models with errors

3.3.1 Introduction

This section discusses parsimonious models of consumer choice that include behavioural biases. The focus is not on the underlying cognitive, neurological or physiological processes of why biases might arise. This is because the aim is to study the impact of behavioural error on economic surplus calculations rather than to discuss psychological aspects of behaviour that are relevant for the prediction of market choices. In line with the previous section, two kinds of direct utility functions are applied to illustrate the implications of behavioural error for economic evaluation and policy decision:

- **Direct decision utility function**: the direct decision utility for consuming a quantity Q of a market good and a quantity G of the outside good is given by $U^d(Q, G)$. It includes all potential behavioural biases and errors an individual can make when calculating the value for decision making. The function can also be a descriptive device for a particular heuristic decision approach of an individual.*

- **Direct experienced utility function**: the direct experienced utility function for consuming a quantity Q of a market good and a quantity G of an outside good is given by $U^e(Q, G)$. It includes the numerical evaluation for this consumption for the case of perfect information and knowledge resulting in no biases in the evaluation. It therefore represents the 'true' value of consumption.*

Classical economic theories usually assume that the decision utility function is equal to the experienced utility function: $U^d(Q, G) = U^e(Q, G)$. The recent interest in behavioural economics has led to attempts to provide extensions for this model for public economics (see Bernheim and Taubinsky 2018).

This section discusses the implication of such extensions for the inverse demand curve that we estimate using market data. Section 3.3.2 provides an example for a model where decision utility and experienced utility are fully separated. Section 3.3.3 develops a model where individuals satisfice rather than optimize. Section 3.3.4 develops a model where individuals calculate utilities on the basis of the weighted average of a decision utility function and an experienced utility function. Section 3.3.5 develops a model that employs direct utility weights.

3.3.2 Approach 1: separated decision and experienced utility

We first investigate the implications of behavioural error for optimal utility and the inverse demand function when decision and experienced utility are fully separated. This implies that experienced utility has no relation at all to the choice utility. For this sub-section, the following direct linear decision utility function is assumed:

$$U^d(Q, G) = U^d(Q, y - pQ) \tag{3.4}$$

where $G = Y - pQ$, stems from the budget constraint. Behavioural errors in valuation can therefore occur for both the market good (the decision value of consumption) and the composite good (the decision value of remaining money). Taking the first derivative of the decision utility function gives the first order-condition for decision utility maximisation:

$$U^d_Q - U^d_G p = 0. \tag{3.5}$$

Here, $U_Q^d \equiv \frac{\partial U^d(Q)}{\partial Q} > 0$ is the marginal decision utility for consuming the market good and $U_G^d \equiv \frac{\partial U^d(Q)}{\partial G} > 0$ the marginal utility of the outside good which is equal to the marginal utility of income in equilibrium. Using the first-order condition one can derive the inverse demand curve:

$$p_d = \frac{U_Q^d}{U_G^d}, \tag{3.6}$$

which is the well-known ratio of marginal decision utilities.

The question is: does it matter? Do consumers loose utility because of the behavioural errors in evaluation? And will this impact economic valuation estimates? Or more formally: are behavioural errors a sufficient condition for biased policy evaluation? In order to answer this question, we have to analyse the counterfactual case of perfect rationality. Suppose that the consumer would be perfectly rational. For that case the optimal the inverse demand is given by:

$$p_e = \frac{U_Q^e}{U_G^e}, \tag{3.7}$$

This looks very much like the result of Eq. (3.6). The only thing that changes in the expression for the inverse demand are that we now have a ratio of the marginal *experienced* utilities.

The marginal choice utilities result in demand choice Q which is in turn evaluated by the experienced utility function. For policy analysis, the only thing that matters in the end is whether the demand Q that is chosen is optimal. This demand determines the final consumer benefits based on the experienced willingness to pay in Eq. (3.7).

Consider the following case. John has a very complicated and biased decision utility function and makes all kinds of errors in the marginal utility of income and in the marginal utility derived from buying solar panels. However, this results in a consumption level of the good which is exactly equal to the case of the decision based on experienced inverse demand (Eq. (3.6). Therefore $Q_e^* = Q_d^*$. The implication for John is that in terms of experienced utility—and in terms of economic value—there is no impact of the behavioural errors of John on the optimal experienced utility.

Only when there is a divergence between Q_e^* and Q_d^*. there is an impact on economic value and for these cases the question is: how much? The answer to that question of course depends on the divergence between the marginal decision utility and the marginal experienced utility as captured by the respective inverse demand curves. When the experienced utility function is very flat around the optimal demand, sub-optimal choices of demand due to behavioural errors do not impact valuations a lot.[3] This leads us to an intermediate conclusion:

[3]You can make a graph with a concave experienced utility function, optimal demand and sub-optimal demand to visualize this.

Behavioural errors are not a sufficient condition for errors in the calculation of economic value. This implies that the assumption of perfect rationality of consumers is not a necessary condition for a correct calculation of economic value. The quantitative impact of sub-optimal choice of demand depends on the steepness of the experienced utility function around the optimum. This is captured by the second-order derivative of the experienced utility function.

Or in other words: it depends. Practically speaking the flatness of the experienced utility curve can be investigated by asking question about changes in experienced value such as happiness when a particular good or service is consumed in higher or lower levels.

3.3.3 Approach 2: satisficing and the inverse demand curve

One of the founding fathers of behavioural economics, Herbert Simon suggested that persons are not optimizing, but satisficing (Simon 1957). In the context of market analysis satisficing means that market actors consume and produce quantities that give 'good enough' utility or costs rather than optimal utility or costs. In mathematical terms this implies that persons do not set the first-order condition for utility maximisation or cost minimisation exactly to 0. The point of satisficing is not that decision utility and experienced utility are different, but rather that the first pillar 'optimisation' is incorrect. This gives satisficing models of behaviour a distinct flavour.

Suppose a representative individual has an income level Y and spends money on the market good Q and some composite or outside good G. The price of the composite good is assumed to be exogenous and is normalised to 1. The price of the market good is given by p. The consumer can choose Q and G but has to account for the budget constraint.

For the rational case the first-order condition for utility maximisation results in Eq. (3.3) . This is the well-known result that the inverse demand curve is defined by the ratio of two marginal experienced utilities. For specified direct experienced utility functions, the first-order condition can also be solved to obtain the optimal quantity Q^*.

Now suppose the representative consumer is satisficing rather than optimizing. This potentially leads to sub-optimal demand choices. Mathematically speaking, satisficing can be modelled by adding an error term in the first-order condition. This error term $e^{-aE}\epsilon$ is related to the effort provision of a consumer E. The parameter $a > 0$ governs the strength of the impact of effort on the error. When the individual provides infinite effort, $E \to \infty$ this error term will vanish to 0. When the consumer provides no effort, $e^{-aE}\epsilon = \epsilon$. The parameter ϵ is assumed to be bounded from below and above. It can be interpreted as the size of the error for consumers who do not provide any effort. The satisficing condition can then be written by:

$$-U_G^e p + U_Q^e + e^{-aE}\epsilon = 0. \tag{3.8}$$

The inverse choice demand function is then given by:

$$p_d = \frac{U_Q^e + e^{-aE}\epsilon}{U_G^e} = \frac{U_Q^e}{U_G^e} + \frac{e^{-aE}\epsilon}{U_G^e}. \tag{3.9}$$

Compared to the experienced inverse demand curve (Eq. (3.6)), satisficing therefore results in an inverse choice demand curve which shifts inwards or outwards with $\frac{U_Q^e}{U_G^e}$. The sign and size of this shift depends on the effort provision, the size and sign of the error and the marginal utility of income. The rational model is a limiting case that results for $\epsilon = 0$, and/or $E \to \infty$.

Until now it is assumed in this subsection that the consumer is not aware of the cognitive biases implying that it is assumed that effort provision is exogenous. An open question in this model is therefore how optimal effort can be determined assuming that an individual makes a trade-off between the benefits of effort provision and the costs of sub-optimal consumption. As in the previous section, it is likely that the curvature of the utility function around the optimum plays a role here. This curvature is captured by the second derivate of the utility function. When this second derivative is strongly negative, sub-optimal consumption has a large impact on utility. For these cases, it is then more likely that the individual will provide more effort when making choices. When the utility function is very flat around the optimum, consumption mistakes resulting from satisficing are not very costly for the consumer and therefore it is likely that the effort that is provided by the individual to reduce error is low.

3.3.4 Approach 3: System I and system II thinking

A third approach to model behavioural error is to define decision utility as the weighted average of System I and System II utility. This terminology was made popular by Kahneman (2011), and is based in the work of Stanovich and West (2000). Loosely speaking, System I can be interpreted as the brain system which is heuristic and makes fast decisions. System II is the slow thinking system which makes decisions upon deliberation and reflection. Instead of choosing to model behaviour with one system or the other, it makes sense to account for both systems when writing down a direct utility function. Therefore, we define direct utility as the weighted average of system I and system II perceived utility:

$$U^d(Q, G) = \varphi U^I(Q, G) + (1 - \varphi)U^{II}(Q, G). \tag{3.10}$$

This specification of utility assumes the individual has at least some partial knowledge about the experienced utility when making the decision, but is also partly governed by the 'heuristic' System I decision utility which is assumed to

be irrelevant for evaluation purposes. The parameter $0 \leq \varphi \leq 1$, governs how system I utility and system II utility are weighted. When $\varphi = 1$, system I utility is completely governing the decision of the consumer. When $\varphi = 0$, we obtain the rational or neo-classical case where decision utility is equal to experienced utility.

When we assume that experienced utility is equal to system II utility, $U^{II}(Q, G) = U^e(Q, G)$, and we obtain:

$$U^d(Q, G) = \varphi U^I(Q, G) + (1 - \varphi)U^e(Q, G). \tag{3.11}$$

Substituting $G = Y - pQ$, and taking the total derivative gives the first-order condition for utility maximisation:

$$-p\varphi U_G^I + \varphi U_Q^I - p(1 - \varphi)U_G^e + (1 - \varphi)U_Q^e = 0 \tag{3.12}$$

Collecting the terms around the price gives:

$$-p(\varphi U_G^I + (1 - \varphi)U_G^e) + \varphi U_Q^I + (1 - \varphi)U_Q^e = 0 \tag{3.13}$$

Solving for the price gives the implicit solution for the inverse demand curve:

$$p_d = \frac{\varphi U_Q^I + (1 - \varphi)U_Q^e}{\varphi U_G^I + (1 - \varphi)U_G^e} \tag{3.14}$$

The inverse demand curve can therefore be expressed as a ratio of weighted averaged system I and system II marginal utilities. When $\varphi = 1$, the inverse demand curve is completely independent of the decision utility and the model of section II.A arises. When $\varphi = 0$, we obtain the rational or neo-classical case where inverse demand is the ratio of marginal experienced utilities.

Some progress can be made by making additional assumptions on how experienced marginal utility and system I marginal utilities relate. Suppose that these marginal utilities are proportionally related so that we can write:

$$U_Q^I = c_Q U_Q^e, \tag{3.15}$$

and:

$$U_G^I = c_G U_G^e. \tag{3.16}$$

When consumers are unaware of the potential benefits of additional consumption of the good, $c_Q < 1$. When they are too optimistic about the consumption benefits of the good, $c_Q > 1$. Similarly, consumers can be unaware of the potential of using the money for other purposes leading to $c_G < 1$. When consumers are too optimistic about the potential of spending their money otherwise, $c_G > 1$.

Substituting the expressions for the decision utility functions into the expression for the inverse demand curve gives:

$$p_d = \left(\frac{\varphi c_Q + (1-\varphi)}{\varphi c_G + (1-\varphi)} \right) \frac{U_Q^e}{U_G^e} \equiv b_c p_e \tag{3.17}$$

Given the assumptions in Eq. (3.15) and (3.16), the experienced demand curve p_e is shifted with a *proportional constant* factor $b_c = \frac{\varphi c_Q + (1-\varphi)}{\varphi c_G + (1-\varphi)}$. This proportional shift is very useful for the analysis of economic value and gives a structural relationship between the experienced and the decision inverse demand curve. When $c_Q < c_G$, there will be an *underestimation* of marginal benefits as $b_c < 1$. When $c_Q > c_G$, there will be an *overestimation* of marginal benefits, and $b_c > 1$. The case $c_Q = c_G$ is the special case where behavioural errors are present, but without impacts for the marginal benefits of consumers as the marginal utilities are shifted with the same proportion.

This again highlights the important fact that the MWTP is a ratio of marginal utilities. Behavioural errors can occur when consumers are not rational about the marginal utility of consumption ('benefits of additional consumption'), or the marginal utility of remaining money ('benefits of additional income'). Note that it is assumed that the parameter $0 \leq \varphi \leq 1$ is exogenous. For more complicated models of behaviour, this parameter can be made dependent on the degree of information and the effort provision of the individual.

3.3.5 Approach 4: direct utility weights.

The fourth approach uses structural modelling and incorporates direct utility weights that capture misperceptions in the benefits in consumption of the market good and the outside good. The advantage of this approach is that the conceptualization of behavioural error is very transparent and that adjustments of existing specifications of direct utility can be used. We will illustrate this approach using an exponential-linear direct utility function defined as:

$$U^d(Q,G) = w_1 \frac{A}{a} \left(1 - e^{-aQ}\right) + w_2 BG \tag{3.18}$$

In Eq. (3.18) w_1 is a perception weight related to the direct utility of consumption of the market good. Benefits of the market good are overestimated when $w_1 > 1$, and underestimated when $w_1 < 1$. In a similar way, $w_2 > 1$ describes the case of overestimation of the benefits of remaining money and $w_2 < 1$ the case of underestimation of the benefits of remaining money.

Substituting the amount for the outside good $G = Y - pQ$ and taking the total derivative gives:

$$\frac{dU^d}{dQ} = w_1 A e^{-aQ} - p w_2 B = 0. \tag{3.19}$$

The inverse choice demand function including behavioural errors is then given by:

$$p_d = \frac{w_1}{w_2}\frac{A}{B}e^{-aQ}. \tag{3.20}$$

The experienced inverse demand results for $w_1 = w_2 = 1$. The experienced inverse demand curve lies above the decision inverse demand curve when proportional errors in the evaluation of the market good utility are larger than proportional errors in the evaluation of the benefits of remaining money ($w_1 < w_2$).

Again, this model shows a plausible behavioural underpinning for proportional shifts in the inverse demand function as we can define $b_c = \frac{w_1}{w_2}$. When there are no income effects this approach also works for more complicated specifications of the direct utility function. In the presence of income effects things become more complicated. The interested reader might try this using a Cobb-Douglas direct utility function.

3.3.6 Conclusion

This section has showed that for plausible behavioural utility functions one can obtain proportional shifts in the inverse demand function due to behavioural error. There are several ways to obtain more information about the size of behavioural error. First, error can be operationalised by asking customers' MWTP before and after consuming the good. When the two diverge this can be a sign that behavioural error is present.

Second, other methods can be employed to obtain the weights in the direct utility function. For example, one can ask customers to grade the market good before and after consumption using satisfaction surveys. When a customer grades the good in utility due to experience. The ratio between decision and experienced utility is then given by 1/1.25 and this provides an argument to set $w_1 = 0.8$. Whether it is possible to disentangle satisfaction related to price payments and actual use of the good is an interesting open question.

Third, one can investigate whether consumers think they have chosen the optimal amount of the good after experiencing the good.

3.4 Behavioural errors and economic social surplus

3.4.1 Implications of behavioural errors for consumer surplus

This section shows what the implications of behavioural error are for the calculation of economic social surplus. For simplicity it is assumed in this section that consumers pay a price p^* for the good and that this price will not change as a result of behavioural error. Define $p_d(Q)$ as the decision inverse demand

curve and $p_e(Q)$ as the experienced inverse demand curve. Figure 3.2 shows the implication of behavioural errors for experienced consumer surplus.

Figure 3.2: implications of behavioural errors on consumer surplus

The left panel of Figure 3.2 shows the case of *underestimation* of the marginal value of an additional unit consumption: the decision marginal willingness to pay is lower than the experienced marginal willingness to pay. From the previous section we know that there can multiple sources for this underestimation.

As expected, this underestimation of value results in an equilibrium quantity level which is lower than the optimal level as $Q_e^* > Q_d^*$. This will lead to a loss in consumer surplus. But how large is this loss compared to the case of full rationality?

The consumer surplus for the left panel is equal to the area under the experienced inverse demand curve up to the equilibrium demand Q_d^*. When behavioural errors are present, the consumer benefits are therefore equal to $A + C + D$. The consumer costs are equal to the area D. Consumer surplus in the presence of behavioural error is then equal to consumer benefits minus consumer costs: $A + C$.

When consumers are perfectly rational, the consumer surplus is again equal

to the area under the experienced inverse demand curve. As the marginal willingness to pay for the good is higher for perfect rationality, the equilibrium consumption is higher and equal to Q_e^*. This results in consumer benefits equal to $A + B + C + D + E + F$, consumer costs equal to $D + E + F$ and consumer surplus equal to $A + B + C$. Behavioural errors therefore lead to a loss in consumer surplus equal to B. This area is equal to the net benefits of those who do not participate in the market because they underestimate the marginal value of consumption.

Next, consider the second case of overestimation of the marginal value of consumption as depicted by the right panel of Figure 3.2. Again, the optimal amount is not consumed, but now the market demand is higher than optimal. Consumer surplus is equal to the area under the experienced inverse demand curve $(A + C + E)$, minus the consumer costs $(C + E + D)$, resulting in the area $A - D$.

In the presence of perfect rationality, the consumption will be equal to Q_e^*. This leads to consumer surplus equal to A. When there is an overestimation of the marginal benefits of consumption of the market good, there is a *loss* in consumer surplus equal to the area D. This area is equal to the net loss in surplus because of the additional entry of consumers due to misperception of marginal benefits. The experienced marginal willingness to pay for these consumers is lower than the price and therefore their net contribution to total consumer surplus is negative. From this discussion on Figure 3.2 we arrive at the following observation related to consumer surplus:

When optimal market demand is not equal to decision market demand, consumer surplus is negatively affected by behavioural errors. The bias goes only via the channel of higher or lower consumption of the market good.

What is true at the individual level is therefore also true at the market level. This is because for the consumers who do not change behaviour in the market when they become more rational there is no change in their consumption decision. These consumers buy the good anyways, regardless of whether their marginal decision willingness to pay is biased or perfectly rational. Therefore, their experienced marginal benefits will not change when they become more rational. Even when the decision marginal willingness to pay is much more complicated graphically, it is only the equilibrium demand together with the experienced inverse demand curve that determines what the size of consumer surplus is and what the size of the losses are related to behavioural error. Behavioural errors therefore mainly impact the decisions of the consumers who are close to the equilibrium level: the so-called marginal consumers.[4] Providing information to those who are in doubt is therefore most relevant.

Figure 3.3 shows the particular case of a market where the equilibrium consumption demand is 'coincidentally' equal to the optimal demand. This case

[4]Note that marginal does not mean unimportant here, but rather consumers who are at the margin of changing their choice.

is important as it shows that behavioural error is not a sufficient condition for losses in calculations of consumer surplus. The consumer surplus for rational and boundedly rational choice is in both cases equal to the area A.

Figure 3.3: Behavioural error without losses in consumer surplus

A mathematical definition of economic social surplus in the presence of behavioural error is given by:

$$CS = \int_0^{Q_d^*} p_e(Q)dQ - p^* Q_d^* \tag{3.21}$$

The first part of Eq. (3.21) are the consumer benefits which are given by the integral under the experienced inverse demand curve up to the equilibrium decision demand. The second part are the consumer costs which are equal to the equilibrium decision demand multiplied by the price. In line with Figure 3.3, this expression is equal to the case of perfectly rational behaviour when the

equilibrium demand with behavioural error is equal to the optimal demand with rationality: $Q_d^* = Q_e^*$.

A survey question for consumers to investigate their behavioural errors is therefore: Do you consume too much or too little of a particular good upon reflection? The answers might give an indication of whether market choice demand is above or below rational demand based on the experienced marginal benefits.

The results depend on the assumption that equilibrium prices do not adjust due to behavioural errors of consumers. The next sub-sections will analyse this assumption in more detail.

3.4.2 The impact of behaviour error on supply decisions and producer surplus

The next step is to gain insights on the impact of behavioural errors of firms on supply decisions. These firms are assumed to minimize costs given a production target resulting in a total cost function that is potentially a function of the production level. It is assumed that fixed costs are zero. Total costs in the market are given by the area under the inverse experienced supply curve. Figure 3.4 shows the implications of behavioural errors of firms on market equilibrium.

The left panel of Figure 3.4 shows the decision inverse supply function $s_d(Q)$ which can be interpreted as the perceived marginal production costs for a group of firms. Because consumers can make behavioural errors, equilibrium arises from the intersection of this decision inverse supply curve with the inverse decision demand curve $p_d(Q)$. This results in an equilibrium price equal to p_d^*. The equilibrium quantity and the equilibrium price are lower compared to the case of perfect rationality of firms. The right panel shows the opposite situation, where firms make behavioural errors leading to a lower equilibrium price and higher equilibrium demand compared to the case of perfectly rational firms.

What are the implications of behavioural errors of firms for producer surplus? The left panel of Figure 3.4 shows the situation where the marginal costs of production are overestimated. The total producer revenues in the market are then equal to the area $A + B + C + D + E$. The total producer costs are equal to the area under the experienced supply function which is equal to E. Producer surplus in the presence of behavioural error, is then equal to $A + B + C + D$.

For the case of perfectly rational firm decisions, demand will increase resulting in producer benefits equal to $C + D + E + G + H$. Producer costs are equal to $E + H$ and therefore producer surplus is equal to $C + D + G$. The difference with the bounded rationality case is then given by $G - A - B$, which is smaller or larger than 0 depending on the change in demand. Overestimation of marginal costs can therefore lead to an increase in producer surplus in the market (but at the expense of consumer surplus as prices increase).

The right panel shows the case where the marginal costs are underestimated by firms. This leads to a lower price and higher equilibrium demand. Producer

Figure 3.4: The implications of behavioural errors of firms on market equilibrium. Note that the figures schematically summarize the argument and potentially amplify the difference between $s_d(Q)$ and $s_e(Q)$ for clarity reasons. In reality the errors might be smaller or larger

surplus is equal to the producer revenues $(C+D+E+I+J)$ minus the producer costs $(B+D+E+F+G+H+I+J)$ and is therefore equal to $C-B-F-G-H$.

Producer surplus can also be written in mathematical terms:

$$PS = p_d^* Q_d^* - \int_0^{Q_d^*} s_e(Q)dQ. \tag{3.22}$$

The first part of Eq. (3.22) are the producer revenues, whereas the second part are the total experienced or real costs in the market (as fixed costs are assumed to be equal to 0). When $Q_d^* = Q_e^*$ producer surplus coincides with the rational case.

Before moving on it is useful to reflect a bit more on the nature of behavioural error for producers. A difference with consumers is that for particular sectors, producer surplus is really monetary: both the benefits and the costs are of a monetary nature. Consider, for example the case of solar panels. It is reasonable to assume that firms that make solar panels are well aware of the costs of production and their benefits of selling the solar panels. This is not the case for consumers: consumer spending is monetary, but the inverse demand is a monetized preference summarized in the maximum marginal willingness to pay. A preference is less tangible and might lack the direct 'feedback' of monetary spending and revenues.

Many firms in the right panel of Figure 3.4 will make a structural loss as the experienced supply curve lies above the equilibrium price. At some point these firms literally pay the price for their behavioural errors. Only the firms with $s_e(Q) < p_d^*$ will make positive profits. For an equilibrium in the market, it is therefore more plausible that the situation of the left panel occurs: firms overestimate marginal costs, for example because they are not minimizing expected costs, but do something else.

This argumentation might be less convincing for sectors where supply decisions use inputs which are non-monetary, less tangible and/or more difficult to process by producers.

Consider for example supply decisions on sustainability research. When scientists have to estimate the amount of labour they invest in research proposals they can make mistakes and underestimate the marginal costs of research. When we observe structural extra work by scientists this might be related to behavioural biases related to labour supply and time: the decision supply curve is then below the experienced supply curve.

3.4.3 Behavioural error and total economic surplus

This sub-section brings together the insights of the previous sections and shows the total economic surplus when behavioural errors are present at the demand and the supply side. It thereby investigates how equilibrium economic value is affected by behavioural errors. We can distinguish 4 conceptual cases:

I. Producers overestimate marginal costs and consumers overestimate marginal benefits;
II. Producers overestimate marginal costs and consumers underestimate marginal benefits;
III. Producers underestimate marginal costs and consumers overestimate marginal benefits;
IV. Producers underestimate marginal costs and consumers underestimate marginal benefits

Only case I will be discussed graphically and the other cases are left to the reader. But before doing this, we add up consumer surplus and producer surplus to obtain a mathematical expression for total economic surplus in the presence of behavioural error:

$$ES = CS + PS = \int_0^{Q_d^*} p_e(Q)dQ - p^* Q_d^* + p_d^* Q_d^* - \int_0^{Q_d^*} s_e(Q)dQ \tag{3.23}$$

$$= \int_0^{Q_d^*} p_e(Q)dQ - \int_0^{Q_d^*} s_e(Q)dQ \tag{3.24}$$

This mathematical expression holds for all conceptual cases. The well-known insight that remains here is that payments of consumers are revenues of producers. When regulators are not interested where the economic surplus ends up this 'dampens' the impact of behavioural errors of consumers (producers) as these are partly mitigated by benefits for producers (consumers).[5] Compared to the rational case, the only thing that changes is Q_d^* which is in line with the earlier observation that behavioural errors materialize in sub-optimal choice of demand.

From this economic surplus equation, we can also learn that there is exactly one point for which economic surplus is optimal. Differentiating ES with respect to Q_d^* gives $p_e(Q_d^*) - s_e(Q_d^*) = 0$, which is satisfied when $Q_d^* = Q_e^*$ as it then shows the equilibrium condition for perfect rationality.[6]

Figure 3.5 shows the impact of behavioural errors for two possible constellations of case I. Equilibrium arises at the point where the decision inverse supply curve is equal to the decision inverse demand curve.

The left panel shows the case where equilibrium demand is higher than optimal. The right panel shows the case where equilibrium demand Q_d^* is lower than optimal. The red area indicates the economic surplus loss because of behavioural

[5]Following classical economic analysis, we ignore the potential normative relevant issue of the source/reason of payments and benefits. Consumers can benefit from behavioural errors of producers and producers can benefit from behavioural errors of consumers.

[6]The second derivative of ES is equal to the slope of the inverse demand curve (negative) minus the slope of the supply curve (positive) which implies that economic surplus is strictly concave in Q. Therefore, there is only one solution for optimal economic surplus which economic surplus is optimal when experienced inverse demand is downward and the experienced supply curve is not downward sloping.

Figure 3.5: Loss of economic social surplus because of behavioral errors for Case I

errors. This area occurs in the left panel because for the consumers between Q_e^* and Q_d^* the experienced marginal costs of production are higher than their experienced marginal benefits. In the right panel it occurs because for the consumers between Q_d^* and Q_e^* the marginal costs of production are lower than the marginal experienced willingness to pay. More economic surplus could be gained when demand moves closer to the optimal demand. In the absence of externalities, the *economic surplus* loss resulting from behavioral errors can be written as the difference between the two experienced curves:

$$\Delta ES = \int_{Q_e^*}^{Q_d^*} p_e(Q)dQ - \int_{Q_e^*}^{Q_d^*} s_e(Q)dQ \tag{3.25}$$

This formula exactly represents the red 'triangles' in Figure 3.5. The economic surplus loss formula is the same for all cases I–IV introduced above. It can be operationalised by assuming particular forms for the experienced inverse demand and inverse supply curves. The figure and the formula show that the economic surplus loss is governed by the difference between Q_d^* and Q_e^*.

Figure 3.6 shows the special case where behavioural errors are present without losses in economic surplus. When choice demand is equal to optimal demand, no economic surplus losses occur despite the fact that consumers and producers make behavioural errors. Nevertheless, this figure shows that behavioural errors lead to higher equilibrium prices compared to the optimal case.

The increase in prices results in an increase in producer surplus for those firms producing on the market. The higher prices for consumers do not impact economic surplus as these are transfers from consumers to the producers. Despite the fact that behavioural errors are made at both sides of the market and that prices are higher compared to the rational case; economic surplus is at the optimal level. Higher prices resulting from behavioural error at the consumer level are therefore not sufficient to conclude that economic surplus goes down.

3.5 Behavioural error and policy recommendations

3.5.1 Behavioural errors and information provision

Figure 3.6 suggests a paradoxical issue when behavioural errors are made and information is provided to consumers or producers in order to help them to make more rational choices: economic surplus can go down. The popular view of behavioural policy intervention interprets errors as a particular form of externalities as they lead to a sub-optimal demand choice in markets (Chetty (2015), figure 9).

As the impacts of behavioural error do not fall upon other actors in and outside the market, but can be viewed as internal costs for the market actors themselves, these are usually referred to as *internalities*. Examples of internalities can

Figure 3.6: Behavioural error without economic surplus losses

be health effects related to cycling, smoking, drinking alcohol or eating meat. When individuals underestimate the impacts of their consumption on their own health, this is a negative internality. One of the policy interventions to combat internalities is the provision of information (cognitive interventions), the change of feelings about the product (affective interventions) and programs to change behaviour (behavioural interventions).

Suppose information is provided to firms in Figure 3.6 and the decision inverse supply curve becomes equal to the experienced inverse supply curve. The result of this is that equilibrium decision demand Q_d^* will move upwards. Because equilibrium demand then moves further away from the socially optimal demand, an economic surplus loss will occur.

A similar result is obtained for providing information to consumers: when consumers in Figure 3.6 become perfectly rational due to information, equilibrium decision demand becomes lower than the optimal demand. For the case shown in Figure 3.6, this results in an economic surplus *loss* compared to the situation without a policy. This somewhat paradoxical result only occurs when there are behavioral errors at the demand *and* the supply side. When errors are made only by producers or only by consumers, providing (costless) information is beneficial because equilibrium decision demand moves into the direction of the optimal demand. We can summarize this analysis in the following observation about behavioural error internalities:

Costless information increases economic surplus when it moves equilibrium demand closer to the optimal demand. This is always the case when only one side of the market makes behavioral errors (either consumers or producers).

There is of course a key normative discussion point here that requires attention. There might be an intrinsic value of providing information for more rational choices. This intrinsic value is now excluded from the analysis. However, when one accepts that this intrinsic value is utilitarian, it can be included as a benefit that adds positively to economic surplus when information is provided. This might result in provision of information even when this does not enhance economic surplus.

3.5.2 Pricing of internalities

Besides information provision, one can also use a price instrument to regulate internalities. This price instrument requires input on the relationship between the decision and the experienced demand and supply curves. Suppose we can define the decision inverse demand curve as a proportion of the experienced inverse demand curve: $p_d(Q) = b_c p_e(Q)$. Section 3.3 gave a mathematical behavioural underpinning for such a proportional relationship. The parameter b_c has a useful interpretation: when behavioral errors result in an inverse demand curve with (on average) 10% higher maximum willingness to pay for consumers, $b_c = 1.1$. Similarly, we can write $s_d(Q) = b_p s_e(Q)$ for producers, where b_p indicates how strong the shift in the experienced marginal costs is due to behavioral errors.

For example, when $b_p = 1.05$, producers overestimate the marginal costs of production by 5%.

Now suppose a regulator introduces a consumer tax τ in order to regulate error internalities. The regulator seeks to optimize economic surplus subject to the constraint that marginal decision benefits are equal to the supply price plus the consumer tax.

Errors will impact the equilibrium demand via the equilibrium conditions. Optimizing economic surplus using Lagrangian techniques results in a tax equal to (see Appendix B):

$$\tau = p^{*R} \frac{b_c - b_p}{b_p}. \tag{3.26}$$

In this expression, p^{*R} is the *equilibrium supply price* in the regulated equilibrium. This supply price does not include the tax. For the case of constant marginal costs, firms do not adjust their price after the tax is introduced and $p^* = p^{*R} = MC$. The consumer price paid is equalto $p^{*R} + \tau$. The tax is therefore proportional to the equilibrium supply price and can deal with all four conceptual cases discussed in section 3.4.3. Three numerical examples are provided below.

Suppose consumers overestimate benefits with 10% and producers are perfectly rational. Marginal costs are independent of the quantity and equal to $p^* = p^{*R} = MC = 10$ euros. Then the tax will be equal to $\tau = p^{*R} \frac{1.1-1.0}{1.0} = 10\frac{1}{10} = 1$ euro. This tax is equal to 10% of the equilibrium supply price in the regulated equilibrium.

When producers overestimate marginal costs with 5% and consumers overestimate marginal benefits with 10%, the tax will be equal to $\tau = p^{*R} \frac{1.1-1.05}{1.05} = 10\frac{0.05}{1.05} = 0.47$ euro. This tax is equal to 4.7% of the equilibrium supply price in the regulated equilibrium.

Suppose, producers underestimate the marginal costs of doing research with 10% and consumers overestimate the benefits with 5%. This results in a tax equal to: $\tau = p^{*R} \frac{1.05-0.90}{0.90} = 10\frac{1}{6}$, which is about 17% of the equilibrium supply price in the regulated equilibrium. These three examples show how to do *what-if* analysis when the exact size of behavioural error is unknown.

Table 3.1 gives an overview of the cases discussed in section 3.4.3, the parameter assumptions and the implications for the sign of the tax. When the sign is negative a subsidy is justified to increase demand. For Cases I and IV, both a tax or a subsidy can apply depending on the sign of the difference $b_c - b_p$. This is because equilibrium demand can be lower, higher or equal to the optimal demand. For Case II there has to be a subsidy in order to move the equilibrium demand closer to the optimum. Case III shows the opposite result and always results in a tax. As discussed before, cases I and II are more likely to occur than cases III and IV.

For the analysis the assumption is made that the parameters b_c and b_p do not change when a particular policy is implemented. A behavioral interpretation

Table 3.1: Behavioral error and the sign of an error internality tax

Case	Case description	Parameter assum
I	Producers overestimate marginal costs and consumers overestimate marginal benefits;	$b_c > 1; b_p > 1$
II	Producers overestimate marginal costs and consumers underestimate marginal benefits;	$b_c < 1; b_p > 1$
III	Producers underestimate marginal costs and consumers overestimate marginal benefits;	$b_c > 1; b_p < 1$
IV	Producers underestimate costs and consumers underestimate marginal benefits;	$b_c < 1; b_p < 1$

could be that consumers and producers are not aware of the behavioral decision errors they make. When the parameters b_c and b_p are behavioral choice parameters things become more complicated as consumers and producers then optimally gather information themselves to reduce their behavioural errors. Nudges, information provision and taxes might have an impact on this behavioral process leading to more complicated tax expressions.

3.6 Behavioral errors and pricing of consumption externalities

3.6.1 Introduction

Consumption externalities occur in markets when the experienced marginal social costs are higher than the experienced marginal costs. There can be so-called negative or positive side effects of consumption that are not captured in the market behavior of producers and consumers. The negative side effects usually gain most attention by policy makers.

Suppose external costs are given by $e(Q)$ and marginal external costs by $MEC \equiv \frac{\partial e(Q)}{\partial Q}$. Then marginal experienced social costs are given by $s_e(Q) + MEC$. A tax is then justified in order to move to the socially optimal equilibrium, where marginal experienced consumer benefits are equal to marginal experienced social costs. But before analyzing this we ask ourselves the question: how do behavioral errors impact economic surplus losses due to external costs?

Figure 3.7 indicates the situation where marginal costs are overestimated by firms as $s_d(Q) > s_e(Q)$. Consumers are assumed to be fully rational. Equilibrium demand in the unregulated equilibrium is below the socially optimal demand Q_s^*. Therefore, the red area indicates the potential economic surplus gain from a taxation policy that results in real marginal social costs equal to marginal private experienced benefits.

A policy that leads to better information for suppliers will bring the decision supply curve $s_d(Q)$ closer to the experienced supply curve $s_e(Q)$. This leads to *higher* equilibrium consumption. Because of the external costs this will lead to an additional economic surplus *loss* equal to the orange area. An internality

Figure 3.7: economic surplus losses, externalities and behavioral errors

and an externality interact with each other: behavioral errors and the marginal external costs. This might lead to suboptimal policies when one of these effects is ignored when formulating a policy.

Figure 3.8 shows another conceptual case where equilibrium demand is higher compared to the fully rational case. In this figure it is assumed that producers behave fully rational and consumers overestimate the marginal benefits of consumption. Information provision to consumers can lead to an economic surplus gain equal to the orange area. An *additional* policy that taxes the externality will lead to an economic surplus gain equal to the red area. Internalities and externalities can also be regulated with a single tax. Such a policy would lead to an economic surplus gain equal to the red plus the orange area.

Figure 3.8: economic surplus losses, externalities and behavioral errors

Ignoring behavioural errors in decision making leads to potential loss of economic surplus. There are three possible conceptual situations each having particular implications for the assessment of policy success.

First, there is the potential to obtain a higher gain in economic surplus as the regulated equilibrium quantity ends up to be lower than the socially optimal quantity $Q_s^{**} > Q_s^*$. The tax could have been set higher, but because of the regulator did not account for the behavioural errors it is lower than optimal. Such a result is in general acceptable as the policy moves demand closer to the social optimum.

Second, it could be that the regulated equilibrium quantity is higher than the socially optimal equilibrium quantity: $Q_s^{**} < Q_s^*$. The tax is too high and this will likely result in opposition.

Third, there can be cases where a tax is not beneficial at all because behavioral errors lead to the socially optimal demand: $Q_s^{**} - Q_s^*$ Behavioral errors than coincidentally result in equilibrium demand equal to socially optimal demand.

This discussion suggests that there are strong qualitative similarities with the discussion in section 3.4.3 (around Figure 3.5). In Figure 3.5 it was found that policies that move equilibrium demand closer to the demand under perfect rationality increase economic surplus. This qualitative conclusion remains true in the presence of consumption externalities, but now optimal (social) demand Q_s^* is the relevant benchmark. When externality taxes move the equilibrium demand closer to Q_s^* these taxes are beneficial for society (according to the model).

3.6.2 A combined externality-internality tax

We can approach the analysis in an analytical way by optimizing the economic surplus function (including external costs $e(Q)$) with respect to the tax. Let $MEC = \frac{\partial e(Q)}{\partial Q}$ be the external costs for example related to pollution. The tax then seeks to correct simultaneously for the internality and the externality. Equilibrium is determined by the intersection of the decision inverse demand and the decisions inverse supply curve. The optimal tax is then given by:

$$\tau = p^{*R}\frac{b_c - b_p}{b_p} + b_c MEC. \tag{3.27}$$

The first part is isomorphic ('equal of form') to the earlier derived internality tax expression of Eq. (3.26). In the presence of external costs an additional term is added to this tax expression which multiplies b_c with the marginal external costs.

When consumers overestimate marginal benefits, $b_c > 1$, the recommended tax is higher compared to the case of perfectly rational consumers. The reason is that the observed equilibrium quantity is below the quantity under perfect rationality. When consumers underestimate marginal benefits, $b_c < 1$, the externality tax should be lower compared to the perfectly rational case. Behavioural errors of suppliers do not have a direct impact on the second part of the tax expression.

When the tax is targeted at suppliers instead of consumers this would not be the case.

3.7 Discussion and conclusion

There are enough examples from the recent literature that market actors can make valuation errors or use heuristics when they are making choices. The lessons of behavioural economists are therefore important inputs for those who work on applied economic policy analysis. Economists who seek to make assessments have the task to investigate whether their assessment and policy recommendations are robust for behavioural errors. Asking consumers whether they consume the right amount of a particular good upon reflection in surveys might help to gain information about the size of behavioural error biases.

The discussion in the previous sections has showed under fairly general assumptions what the implications of behavioural errors in partial equilibrium could be. However, still input is needed on the experienced supply and demand curves. If it is not possible to measure these curves, one still cannot derive the economic surplus effects of policies without making additional assumptions about the approximate proportional relationship between the decision and experienced demand and supply curves.

In order to develop stylized cases, this chapter assumes proportional behavioural errors which are helpful for interpretation and rough quantitative estimates of how behavioural errors impact recommendations. In the absence of empirical information, such a *what-if* analysis is often the best option available for practical economists doing cost-benefits analysis for public policy.

For externality tax recommendations the sign of the impact of behavioural errors on externality taxes is not clear beforehand. This implies that economists should be cautious in their recommendations when knowledge and evidence about the size and sign of behavioural errors is lacking for particular markets, unless one is willing to assume that the neo-classical model is the correct model for a particular market by default.

As section 3.2 shows, from a normative perspective the regulation of internalities is still controversial as there can be many objections to the idea of regulating and observing structural behavioural error that are counterintuitive at first sight. For example, should a meat tax include the potential health effects or not when individuals deliberately make their choices? Should cycling be subsidized when there are also accident externalities present? Again, the use of surveys to investigate perceptions of what is internal and external for consumers can help here. Furthermore, analysis of the acceptance of particular policies targeted at internalities might be useful. The case of smoking might provide a good benchmark example of regulating internalities where freedom of choice is limited by addiction to the market good. Whether such an argument is also applicable for other market goods is open for investigation.

Chapter 4

Moral Considerations and Economic Surplus[1] [2]

4.1 Introduction

These lecture notes introduce another definition of economic surplus ('the economic pie') that allows for moral considerations. It specifically investigates moral considerations related to the normative validity of counting costs and benefits of consumers and producers and external costs as economic social surplus. Why is this important? In the first week we have seen that the classical perspective of economic value rests on four pillars: (*i*) optimisation, (*ii*) equilibrium, (*iii*) empiricism and (*iv*) choice-based normativity. The implication of this perspective is that choices of market actors perfectly track value for society.

The second week provided an extended perspective using insights from psychology. According to this perspective there can be a difference between choice-based value and experienced based value. Given particular assumptions it can be justified that experienced value is the right metric to measure instead of choice-based value. As we have seen, this has implications for policy interventions related to information provision and externality taxation.

This chapter questions the results of chapter 2 from a different angle: there can be moral considerations related to the choices of market actors that lead to the question whether it is justified to include the value of consumers, producers and

[1] These lecture notes are based on the paper 'Moral Considerations and Economic Surplus' (working paper). Please do not distribute these lecture notes without permission.

[2] The author gratefully acknowledges the contribution of the Dutch Organization for Scientific Research (NWO Responsible Innovation grant—313-99-333). Furthermore, the content of these lecture notes benefited from the authors' participation in a seed money project from the Amsterdam Sustainability Institute. I am indebted to many people inside and outside academia for their invaluable thoughts and responses.

the external costs completely in the economic surplus calculation.

Before moving on, it is useful to illustrate why moral considerations are important. Consider a pork producer who is active in the market for pork. The government has set specific rules for pork production for those who want to become active in this market. With the aim of lowering marginal costs, the pork producer ignores some of these rules related to animal welfare without providing good reasons and without changing external costs. The producer rationally trades off lower marginal costs with the probability of getting a fine when being caught. According to the economic model of value developed in chapter 2 this leads to an *increase* in producer surplus in the market of pork and thereby into an increase is economic surplus. The example suggests that there are normative limits for behaviour by consumers and producers in markets that might not be captured in behavioural choice.

Or consider a philosopher of animal welfare who considers the eating of pork to be morally problematic because pigs are sentient beings 'like us'. If the reasoning of this philosopher is correct, counting consumer benefits fully as social surplus is likely to be questionable.

These two examples at least illustrate that the choice-based normativity perspective (chapter 2) and experienced-based normativity (chapter 3) might not be shared by all persons and groups in society. In these lecture notes we therefore change the 4th pillar 'normativity' from choice-based to morally-based evaluation.

How to deal with this in cost-benefit analysis?

The key question to investigate here is: "What should count as 'better or worse off'?" (Atkinson 2009; Sandel 2013). This normative question can be investigated from different viewpoints: the individual level, the community level and the societal level.

At the *individual level*, consumers and producers make moral commitments to their own perspective on the good life. They might want to eat less meat, fly less and might want to lower their carbon emissions by making more sustainable choices. However, these moral commitments and targets are not always reached leading to a discrepancy between the individuals' perspective on the good life and their actual choices. Most persons are aware of this gap between their goals and actual behaviour in their personal lives. For example, sometimes individuals make promises to particular sustainability targets to others but do not keep these promises. Another example is the observation of flight shame which suggests that individuals have an independent reflective moral attitude (conscience) towards their own behaviour. Even when choices (Chapter 2) and experiences (Chapter 3) are valuable and decisions are made to buy the product, this suggests that these choices might not be evaluated as truly 'good' by the individual.

At the *community level*, individuals participate in communities who determine (partly or fully) what counts as valuable. Community membership can be voluntary or non-voluntary and membership of communities results in particular

duties and customs. Communities entail particular minimum standards of behaviour for consumers and producers based on the community values. These standards can be viewed as partly independent of the preferences that are expressed by consumers and producers in their day-to-day market choices. For example, an activist participating in a 'climate change community' is likely to be held accountable for behavioural choices: only 'telling the truth' is likely not sufficient according the community standards.

At the *societal level*, individuals participate in the common project of living together. Governments can formulate particular minimal sustainability standards for market participants which in turn can lead to particular duties and customs for market actors. This was illustrated with the example of the pork producer above. Furthermore, governments can ask market participants to make concrete commitments in terms of sustainability. Goals can act as targets for moral excellence and result in obligations.

In order to count something as valuable for society these different viewpoints can play a role and can result in a discrepancy between choice-based value and morally weighted social value. In line with the results of Chapter 3 this can imply that the social optimum is not reached even when there are no external costs. Note that the different viewpoints do not restrict choice of market actors but seek to foster a plural perspective on what should count as valuable.

The remainder of this chapter is structured as follows. First, Section II gives an overview of several strategies that are employed to deal with moral considerations when valuing market outcomes. This section motivates why an adjustment of economic surplus is needed when moral considerations are present. Section III develops a novel analytical specification of economic surplus and sections IV and V discuss the implications for pricing. Examples are provided along the way. Section VI concludes.

4.2 Dealing with moral considerations at the valuation stage

4.2.1 Introduction

This section discusses four general approaches to deal with moral considerations when analysing (changes) in economic social surplus. The first approach is the ethical checkbox approach (section II.B), the second approach is the economizing ethics approach (section II.C), the third approach is the ethicizing economics approach (section II.D). The fourth approach is the qualitative valuation approach (section II.E). These approaches are not necessarily perfect substitutes but can be employed in a complementary way.

4.2.2 Ethical checkbox approach

The first approach to deal with moral considerations is the 'ethical checkbox' approach (ECH). This approach analyses moral considerations independently of economic surplus assessments using a list of *necessary* ethical requirements or rules for the behaviour of market actors. It is closely related to the 'independent doctrine' which views (micro-)economics as separate from ethics (see High 1985, 4–6, for a brief history). The advantage of the ECH approach is that it gives room to the realm of moral and ethical reasoning for guiding market practices. This circumvents that ethical considerations are fully overtaken by utilitarian and game-theoretic arguments of economists when policy recommendations are given. When moral considerations are of a different kind than (monetized) utilitarian consumer value or producer profits and transcend consumer and producer choice behaviour such a separation seems useful and justified. When the ethical checkboxes are ticked, micro-economists can employ the standard definition of economic social surplus which is—in the absence of externalities—assumed to be equal to the sum of consumer and producer surplus.

A disadvantage of the ECH approach is that it ignores the loss of economic social surplus resulting from consumers and producers in the market who do not act according to the ethical checkbox. The example of the pork producer who violated the rules illustrates this. Furthermore, the ECH approach also ignores other moral considerations that are not part of the checkbox but are nevertheless important for individual market actors, the communities they participate in or for society. This is because the standard specification of economic social surplus is based on choice-based normativity (pillar (*iv*)). The ECH approach assumes that costs and benefits should be counted without further normative consideration by the individual, the community or society. Consumer and producer moral responsibility do not enter the equations insofar this responsibility is not fully expressed in the particular market under consideration. The normative evaluation of public policies for markets is of course affected by this implicit assumption of moral *excellence* of market actors. As a result, it is questionable whether the assumed economic social surplus function correctly measures economic social value in the presence of (remaining) moral considerations.

4.2.3 Economizing ethics approach

The second approach to deal with moral considerations when calculating economic social surplus is the 'economizing ethics' approach (ECE). This approach views moral considerations as (monetized) utilitarian internalities or externalities and has been applied to analyse repugnant markets, morality and identity, public policy decisions with an ethical component and regulations in the social domain (Lazear 2000; Bénabou and Tirole 2011; Roth 2018; Chorus et al. 2018; Mouter, Koster, and Dekker 2021).

A good aspect of ECE is that it might enhance understanding of individuals choice behaviour when moral considerations for private or public decisions are

present (Chorus 2015). Furthermore, it frankly acknowledges that ignoring moral considerations in markets can lead to a loss of utilitarian value for people inside and outside these markets. Posner and Sunstein (2017) (p. 1812) write:

> Our simple answer, put too briefly, is that on welfarist grounds, moral commitments can matter, and that when people would suffer a welfare loss when their moral commitments are violated, regulators should ask: how much are people willing to pay to honor those commitments?

And in another section of this paper Posner and Sunstein (2017) (p. 1840) write: "*Our major goal is to acknowledge rather than to resolve the measurement problem and to insist on the basic principle: people* ***experience*** *welfare losses from social outcomes that offend their moral commitments, even if those outcomes do not involve their own wealth or health. Private willingness to pay is the best way to measure those losses.*" The ECE approach therefore assumes that all experiences should count including experiences of moral losses or gains by third parties due to the market choices of consumers and producers.

ECE might also explain descriptively why there are societal laws that put restrictions on market trade. This is because the loss of monetized moral value by third parties can go beyond the value of market transactions, but is related to the actual market behaviour of consumers, producers and the regulator. According to the ECE approach, including the monetized utilitarian externalities stemming from moral considerations thereby can lead to better estimates of (changes in) economic surplus. Furthermore, an advantage of the ECE approach for economists is that it can be operationalized using empirical choice data in combination with the micro-econometric toolbox of applied economists. It is possible to ask persons what their willingness to accept is for violating a particular moral consideration in order to measure the external effects.

A disadvantage of the ECE approach is that it often fails to make normative sense for assessments of decisions outside markets. Without acknowledging distinct social domains with their corresponding distinct core values, rules and customs, it feels like bringing the hockey-stick of economic efficiency to join a field where teams play basketball.

4.2.4 Ethicizing economics approach

The third approach is the 'ethicizing economics' (ETE) approach which is explored in this paper. There are two different ways the ETE approach has been implemented in the literature. First, in the climate change literature, this approach has been pursued to deal with distributional effects using a novel specification of the social welfare function that employs a single ethical perspective such as prioritarianism instead of ECE (Spash 1997).

The second ETE approach which is pursued in these lecture notes, seeks to operationalise diverse normative ideas about what is socially valuable by re-specifying

economic social surplus using ethical parameters. Such a normative *diversity* might even be present at the level of the individual as a mixture of deontological, virtue and consequential arguments play a role in the reflection on behavioural choices of market actors (see Van Staveren 2007; White 2009). Economists are well equipped to study consequences, but other moral considerations related to duties and virtues have received less attention. What kinds of considerations can we think of?

Deontology is the study of duties:

> The word deontology derives from the Greek words for duty (deon) and science (or study) of (logos). In contemporary moral philosophy, deontology is one of those kinds of normative theories regarding which choices are morally required, forbidden, or permitted. In other words, deontology falls within the domain of moral theories that guide and assess our choices of what we ought to do (deontic theories).Stanford Encyclopedia of Philosophy

Deontological moral considerations therefore are therefore related to what market actors ought (not) to do and not what they are actually doing. This type of moral considerations results in particular duties. For now, it is sufficient to notice that such moral considerations can exist for particular markets of interest.

Moral considerations related to virtue focus on moral character:

> It may, initially, be identified as the one that emphasizes the virtues, or moral character, in contrast to the approach that emphasizes duties or rules (deontology) or that emphasizes the consequences of actions (consequentialism).Stanford Encyclopedia of Philosophy

Besides moral character, virtue ethics also focusses on motivations and intentions behind the choices of market actors. According to this perspective, choosing polluting goods for utilitarian reasons is not sufficient when there are no appropriate motivations or reasons given.

The framework for cost-benefit analysis developed in Chapter 2 does not deal with these moral considerations and leaves out normative reflection at the individual, community and societal level in the calculation stage. For example, a virtue ethicist focussing on sustainability practices might want to count the benefits of flying to Barcelona for the funeral of a family member differently than the marginal benefits of a person who goes 5 times per year on holidays by air travel 'just for fun'. Motivations can matter for counting benefits and costs according to this view.

The ETE approach acknowledges that the mathematical specification of economic social surplus is a moral enterprise itself and that it can impact economists' normative recommendations for optimal economic policies. Furthermore, the plural version of it acknowledges that there are multiple reasonable normative positions each resulting in a particular specification of the economic 'pie'. Count-

ing producer costs and benefits fully—as we did in week Chapter 2—is just one of the normative perspectives.

Working directly on the economic social surplus function seems to be more in line with the arguments developed by moral philosophers from schools and traditions that focus on virtues and duties. The intuition of the difference between the ECE approach and the ETE approach is illustrated by structuring the argument. The ECE approach proceeds as follows:

(A) Consumer choices of market good Q are morally questionable for sound moral considerations R.
(B) Consumer benefits for Q are not adjusted in the economic surplus.
(C) Additional external costs are counted for violation of moral considerations R by consumers.

This follows the ideas of Posner and Sunstein (2017) who propose to monetize external costs and results in a model that is equivalent to the model in Chapter 2 with additional external cost terms.

The ETE approach instead, deals directly with the moral consideration when specifying the economic surplus:

(A) Consumer choices of market good Q are morally questionable for sound moral considerations R.
(B) Then in order to account for moral consideration R in studies of the market in Q, it is reasonable to adjust the consumer benefits for Q downwards using an ethical parameter.

Both approaches account for moral considerations but do this in a different manner depending on whether researchers find ECE or ETE more appropriate.

4.2.5 Qualitative valuation approach

The fourth approach to deal with moral considerations is the qualitative valuation (QUV) approach. For this approach the quantitative calculation of exact changes in economic surplus is less relevant. The QUV approach can highlight which values are relevant for normative policy decisions on markets or decisions in other spheres of society (Klamer 2017; Wempe and Frooman 2018).

It is valuable for at least six reasons. First, it can enhance reciprocal understanding of those involved in valuation by fostering the conversation about the values underlying regulatory and market choices. Second, it acknowledges that normative preferences for public policies can partly be socially constructed for example by reading, role switching or by a deliberative process. Third, it is helpful for articulating which moral considerations are relevant for market valuation and why. Fourth, QUV can be helpful for the selection of competent persons for policy decisions. Fifth, an advantage of the QUV approach can be that the conversation is open and not dictated by a pre-defined analytical valuation method. Sixth, qualitative motivations developed by participants in the QUV

approach can be helpful for policy makers articulating their motivations for their public policies (see for example Mouter, Koster, and Dekker 2021, sec. 5).

A disadvantage of the QUV approach is that no *quantitative* estimation of impacts and value are provided that allow for a comparison of different economic policy interventions in markets including the status quo of doing nothing. Therefore, quantitative questions such as: 'how high should an externality tax on good X be?' or 'what is the loss of economic social value of changes in the market for X?' cannot be answered using the QUV approach only. A second disadvantage is that it is harder to reach consensus, as heterogeneous opinions are sometimes hard to weigh when they are based on qualitative statements only. A third disadvantage is that the outcomes of the QUV approach depend on the quality of expression of the people who are involved. This might disadvantage those with less rhetorical and verbal qualities.

4.2.6 Contributions of the remainder of this chapter

There is no good reason to assume *a priori* that the assumption of moral excellence of consumers, producers and regulators in markets is valid. Furthermore, different groups in society or different experts might define moral excellence in different ways, and in some markets the assumption of moral excellence might be more valid than in others. According to Sen (2000) (p. 952):

> sensible cost-benefit analysis demands something beyond the mainstream method, in particular, the invoking of explicit social choice judgments that take us beyond market-centered valuation.

The aim of the remainder of this chapter is to offer a novel specification for the economic pie that allows for diverse normative considerations related to economic social surplus. The methodological approach allows for the inclusion of three kinds of moral considerations. Moral considerations related to:

(i) marginal consumer benefits;
(ii) producer marginal costs;
(iii) external costs.

In order to show the impacts for economic policy recommendations, the new conceptualisation of economic social surplus is applied to the optimal taxation of a consumption externality. Such an exercise is useful to gain intuition on how moral considerations and ethical parameters can impact standard economic policy recommendations. The externality taxes seek to efficiently price positive or negative side effects resulting from market transactions while still allowing for trade in the market. But what is defined as 'efficient' of course depends on the definition of the economic pie by the researcher. Something, we already discovered in Chapter 3.

4.2.7 Scope

Before moving on, it is good practice to think about the scope of the model: where can it be applied, and for which decisions can it not be applied? The model of this week is applicable for the analysis of the taxation of a consumption externality in existing markets with remaining moral considerations. It is not applicable for the analysis of whether a market should be opened or closed by a government. This class of policy decisions leads to other kinds of moral considerations related to offering choice opportunities. This raises fundamental freedom questions of which the analysis is beyond the scope of this academic course.

Furthermore, distributional concerns, price and tax discrimination, bounded rationality of market actors (Chapter 3.), strategic firm interactions, limited information, general equilibrium and bargaining between consumers and producers are ignored. This is done for clarity of exposition and argumentation and to stay close to how practical cost-benefit is often operationalized.

4.3 Redefining the economic pie

4.3.1 The advisory committee

This section redefines the economic 'pie' for a stylised partial equilibrium model of a single good or service mathematically and graphically. Suppose that an advisory committee is summoned that has N members. Assume that each member n of this advisory committee receives a weight $\omega_n \geq 0$ and that these weights sum up to 1. Depending on the kind of market, the committee could consist of the full population of a region, city, country or the world, a group of representative elected officials, representatives of particular communities, representatives of responsible advisory institutions, a group of (scientific) experts or a weighted combination of all these groups. Membership of the committee can be based on experience, knowledge and expertise and/or on the degree that the expert represents a community or population of interest or simply on the status of being a citizen.

We will not discuss how this advisory committee *should* look like, what the requirements of membership should be and what the procedural process of deliberation should be. The mathematical description also abstracts away from potential strategic considerations within the advisory committee.

It is assumed that all normative positions of the committee members are in principle correct or internally consistent and are based on correct information about the consequences of behaviour. This pragmatic assumption is inconsistent with the observation that in the presence of objective moral considerations there can be mutual exclusive positions where given the underlying (meta-)ethical assumptions the moral reasoning of only one or several of the committee members

is correct.[3]

4.3.2 Adjusted consumer surplus

Total monetized consumer benefits for the market are denoted by $\int_0^{Q^*} p(Q)dQ$ where Q is the demand for the good or service. The marginal consumer benefits are given by $p(Q)$. This market inverse demand function reflects the maximum marginal willingness to pay for a group of anonymous consumers and is assumed to be independent of consumers' income levels. In equilibrium, marginal consumer benefits are equal to the private costs, resulting in an equilibrium demand of Q^*. It is assumed that all consumers all have the same private access costs and pay the same equilibrium price p^* for the good or service. From a normative perspective the assessment of marginal consumption benefits of the consumer as reflected by the consumers' marginal willingness to pay deserves more investigation.

When there are moral considerations related to consumers' marginal benefits, member n of the committee can adjust the consumer benefits with an ethical weight $0 \leq \alpha_n \leq 1$. This adjustment can be based on moral considerations and reasons related to consumer responsibility. Committees can use survey data and interviews to learn more about the reasons and intentions of consumers or investigate whether the consumption choice has been made under free conditions or is the result of addiction. The adjustment of marginal consumer benefits does not entail that choice *should* be restricted or forbidden.

Moral considerations relating to consumer behaviour can lead to scaling down consumer benefits in particular markets. Eq. (4.1) adjusts the consumer surplus for moral considerations related to consumer marginal benefits:

$$ACS = \sum_{n=1}^{N} \omega_n \left(\int_0^{Q^*} \alpha_n p(Q)dQ - p^* Q^* \right) = \bar{\alpha} \int_0^{Q^*} p(Q)dQ - p^* Q^*, \quad (4.1)$$

where $\bar{\alpha} = \sum_{n=1}^{N} \omega_n \alpha_n$ is the weighted average sum over all committee members for the ethical parameters for consumer benefits.[4] In the presence of moral considerations, consumer benefits are therefore adjusted downwards. Adjusted consumer surplus estimates in markets with moral considerations will therefore decrease in comparison to the case where moral excellence of consumers is assumed ($\bar{\alpha} = 1$).

At this point it is helpful to provide an illustration how the model can be operationalized for practical policy analysis. Suppose an advisory committee on

[3]See Singh et al. (2020) for a recent discussion on expert agreement, disagreement and the (non-)existence of objective moral facts.

[4]An additional assumption is required for a valid model: at least one committee member needs to have $\alpha_n > 0$. It is not necessary to have a weighted average: a majority voting decision on the weights is also possible, although this fully leaves out the normative perspectives of particular minorities in the advisory committee.

Table 4.1: Representatives and weights. Numbers are chosen for illustration.

Committee member	Decision weight	Recommended weight for consumer ma
Citizen representative.	$\omega_1 = 0.1$	$\alpha_1 = 0.6$
Next generation representative.	$\omega_2 = 0.1$	$\alpha_2 = 0.5$
Consumer representative.	$\omega_3 = 0.1$	$\alpha_3 = 1$
Producer representative.	$\omega_4 = 0.1$	$\alpha_4 = 1$
Governmental representative (chair of the committee).	$\omega_5 = 0.1$	$\alpha_5 = 0.6$
Economist (specialization: environmental economics)	$\omega_6 = 0.1$	$\alpha_6 = 1$
Philosopher (specialization: animal ethics)	$\omega_7 = 0.1$	$\alpha_7 = 0.7$
Ecologist (specialization: biodiversity)	$\omega_8 = 0.1$	$\alpha_8 = 0.6$
Climate activist representative.	$\omega_0 = 0.1$	$\alpha_9 = 0$
Animal representative.	$\omega_{10} = 0.1$	$\alpha_{10} = 0$

the economic value of the pork market is summoned. There are 10 representatives invited for this committee. Each of them has received an equal weight $\omega_n = 0.1$.

Table 4.1 below provides an overview of the ethical weights the representatives propose for the marginal benefits of consumers and their reasons for choosing these weights. It includes a wide range of voices and reasons from society, institutes and advocacy groups.

The average ethical weight for consumer marginal benefits is given by the following weighted sum:

$$\bar{\alpha} = \sum_{n=1}^{N} \omega_n \alpha_n = 0.1 \sum_{n=1}^{10} \alpha_n = 0.6 \tag{4.2}$$

This shows that the average consumer weight results in counting 60% of the consumer marginal benefits. According to this committee, the counted consumer benefits will therefore be 40% lower compared to the model of Chapter 2. (Note that numbers and reasons are chosen for illustration).

It might be helpful for you to reflect on the reasons and their relationship to the ethical weights. Do you think the reasons behind the weights are convincing? Which expert is closest to your own opinion on the topic? Do you think there are other reasons that should be included? What do you think about assigning the experts an equal weight 0.1?

Figure 4.1 summarizes this section with a graphical illustration of the impacts of averaged ethical weights for consumer surplus in markets. It shows that market equilibrium is established at the point where the inverse demand is equal to the supply price p^*. In the presence of moral considerations related to consumer benefits, the inverse demand applied for evaluation shifts inwards

with a factor $\bar{\alpha} < 1$. As equilibrium demand remains at Q^*, the counted total consumer benefits are equal to $A + B$ and are $C + D$ lower compared to the case $\bar{\alpha} = 1$ which is the classical case of Chapter 2. Consumer surplus is given by $(A + B) - (B + D) = A - D$. This is lower compared to the consumer surplus calculation that ignores moral considerations $(A + C)$.

Figure 4.1: Moral considerations and adjusted consumer surplus

4.3.3 Adjusted producer surplus

Producer behaviour is described by a market inverse supply function $s(Q)$ that shows the minimum price for which a total of Q units are supplied by a group of anonymous firms. The area under this inverse supply function up to the equilibrium demand gives the total costs for all firms in the market. It is assumed that fixed costs are zero. Given the assumptions made, the area above this inverse supply curve up to the equilibrium price then measures the total net profits in the market. As in the basic micro-economic model, firms are assumed to be price-takers and general equilibrium and potential market power considerations are ignored. Furthermore, it is assumed that costs for production inputs can be counted as real societal costs instead of transfers to other firms or workers.

Moral considerations related to the marginal costs can be related to the investments of firms to produce goods or services. When there are moral considerations

related to capital costs, labour, natural resources, (raw) materials and information/data used for production, these producer costs can be adjusted upwards by committee members applying an ethical weight $\delta > 1$.

Moral considerations related to capital costs can be present when it is unclear where the money comes from that is used to finance the business. When money comes from illegal or doubtful criminal sources or when there is evidence of financial fraudulent behaviour this gives a reason to adjust the marginal cost curve upwards (see Reurink (2018) for an overview of such moral considerations).

There are many examples of moral considerations to correct producer costs upwards related to labour. Firms might have lower production costs because they use forced labour or child labour of children who cannot go to school (probably indirectly via subcontracting). Suppose that the counterfactual option of fair production would be to produce at a marginal cost curve that is 5% higher. When committee member n considers subcontracted forced and/or child labour to be morally problematic (s)he can adjust the total costs using $\delta_n = 1.05$. In line with the ETE approach, this correction does not include the external costs of forced labour but seeks to account for the intrinsic moral quality of the production process of the firms in the market that make use of this labour.

Moral considerations related to resources and raw materials can be important as well. For example, when materials of products cannot be re-used in the production process this might be viewed as intrinsically less valuable for ecological oriented committee members. Taking the moral considerations related to (iii) and (iv) into account, the adjusted (equilibrium) producer surplus can be written as:

$$APS = \sum_{n=1}^{N} \omega_n \left(p^* Q^* - \int_0^{Q^*} \delta_N s(Q) dQ \right) = p^* Q^* - \bar{\delta} \int_0^{Q^*} s(Q) dQ, \quad (4.3)$$

where $\bar{\delta} = \sum_{n=1}^{N} \omega_n \delta_n$ are the weighted average sums of the producers' related ethical weights over all representatives. Moral considerations lead to higher producer costs compared to the standard model of Chapter 2. In highly controversial markets there is no reason why this adjusted producer surplus cannot be negative.

At this point it is useful to illustrate the model in this subsection using an example. Table 4.2 summarizes the weights and the reasons of the representatives. All numbers are chosen for illustrative purposes.

The average weight for the marginal costs is given by:

$$\bar{\delta} = \sum_{n=1}^{N} \omega_n \delta_n = 0.1 \sum_{n=1}^{10} \delta_n = 2.228. \quad (4.4)$$

This implies that counted marginal costs are about 223% higher than the real

Table 4.2: Scaling of marginal costs of production. All numbers are chosen for illustration.

Committee member	Decision weight	Recommended weight for c
Citizen representative.	$\omega_1 = 0.1$	$\delta_1 = 1.1$
Next generation representative.	$\omega_2 = 0.1$	$\delta_2 = 1.1$
Consumer representative.	$\omega_3 = 0.1$	$\delta_3 = 1$
Producer representative.	$\omega_4 = 0.1$	$\delta_4 = 1$
Governmental representative (chair of the committee).	$\omega_5 = 0.1$	$\delta_5 = 2.2$
Economist (specialization: environmental economics)	$\omega_6 = 0.1$	$\delta_6 = 1$
Philosopher (specialization: animal ethics)	$\omega_7 = 0.1$	$\delta_7 = 1.25$
Ecologist (specialization: biodiversity)	$\omega_8 = 0.1$	$\delta_8 = 1.8$
Climate activist representative.	$\omega_9 = 0.1$	$\delta_9 = 1.8$
Animal representative.	$\omega_{10} = 0.1$	$\delta_{10} = 10$

marginal costs. It is useful for you to reflect on the different reasons behind the chosen ethical weights. Do you consider their reasoning to be valid? Are there other reasons that should have been brought to the table? Which representative most closely matches your own perspective?

Figure 4.2 summarizes this sub-section with a graphical illustration of the impacts of moral weights on the determination of producer surplus in markets. The figure shows that equilibrium is established at the point where the market inverse supply curve $s(Q)$ is equal to the equilibrium price as firms with higher marginal production costs will not produce for a lower price. In the presence of moral considerations, the market inverse supply curve applied for evaluation shifts upwards with a factor $\bar{\delta} > 1$. As equilibrium demand remains at Q^* regardless of how costs are counted, the counted producer costs in the economic social surplus are then $E + G + H$. This is $E + G$ higher compared to the case where $\bar{\delta} = 1$. Total counted producer surplus is equal to $F - E$. This is $G + E$ lower compared to the neo-classical case.

A key difference with the model developed in Chapter 3 is that the producers do not adjust prices according to the normative standard that is set at the evaluation stage. Therefore $s(Q)$ is used to determine the equilibrium price (and not $\bar{\delta}s(Q)$).

4.3.4 Adjusted economic surplus

In the absence of external costs, the adjusted equilibrium social surplus is given by the sum of the adjusted consumer and producer surplus:

$$AES = ACS + APS = \bar{\alpha} \int_0^{Q^*} p(Q)dQ - \bar{\delta} \int_0^{Q^*} s(Q)dQ. \tag{4.5}$$

Figure 4.2: Moral considerations and producer surplus.

Eq. (4.5) shows that moral considerations bring down the consumer benefits and raise the producer costs. Figure 4.3 shows the welfare loss that follows from the assumption of moral excellence of consumers and producers. Equilibrium demand Q^* results from the intersection of the inverse demand curve $p(Q)$ and the inverse supply curve $s(Q)$. Q^{**} would be the socially optimal quantity as it is determined by the intersection of weighted marginal consumer benefits and weighted marginal producer costs.

Figure 4.3: Welfare loss following from the moral excellence assumption

The red 'triangle' then indicates the loss of welfare due to the moral excellence assumption. As the demand in equilibrium is higher than optimal, there is a loss of surplus for the consumers between Q^* and Q^{**}, as for these consumers counted consumer costs are higher than counted consumer benefits. The key insight can be summarized as:

> According to the model, the assumption of moral excellence of market actors leads to structural overconsumption in markets as the socially optimal demand is always lower than the equilibrium demand in the unregulated situation."

The result in Figure 4.3 is qualitatively different compared to behavioral welfare economic models with behavioral errors made by consumers and producers. These behavioral errors can lead to higher or lower 'experienced' inverse demand curves and to higher orlower 'experienced' inverse supply curves. As we have seen in Chapter 3, this can result both in equal, lower or higher equilibrium

demand than the socially optimal demand under perfect rationality. This is not the case for the conceptual model developed for this chapter: because of the moral excellence assumption optimal demand will always be lower than observed demand.

4.4 Taxation and moral consideration

Because market surplus is not equal to the socially optimal surplus a regulator can correct for this using a price instrument. Such a price instrument does not interfere with the freedom to consume or produce and does not have to be a substitute for using policy measures to combat remaining injustice in the market. The assumption that the ethical parameters are independent of the equilibrium price and quantity helps to obtain some stylized results.

Appendix C.1 shows that the equilibrium tax expression that optimizes the adjusted social surplus Eq. (4.5) is given by:

$$\tau = p^{*R}\frac{\bar{\delta} - \bar{\alpha}}{\bar{\alpha}}. \tag{4.6}$$

The first observation is that this tax is non-negative as the equilibrium price in the regulated equilibrium (p^{*R}) is positive and $\bar{\delta} \geq 1 \geq \bar{\alpha} \geq 0$. In the presence of moral considerations, it is therefore economically justified to impose a positive tax that corrects for the fact that in the unregulated equilibrium, consumers and producers in the market ignore moral considerations. This observation is consistent with the discussion around Figure 4.3: socially optimal demand is below equilibrium demand in the unregulated equilibrium.

Second, because the ethical weights enter proportionally, the tax in Eq. (4.6) is proportional to the equilibrium price p^{*R} in the regulated equilibrium. This result does not hold anymore when ethical weights are functions of the demand.

Third, moral considerations related to consumer benefits raise the tax as decreasing $\bar{\alpha}\}$ is equivalent with shifting the adjusted market inverse demand curve inwards. The level of the tax increases in $\bar{\delta}\}$—as intuitively—moral considerations related to producer costs are equivalent with shifting the market inverse supply curve upwards, leading to a lower socially optimal equilibrium demand.

Fourth, a stylized example is useful to gain some intuition about the potential quantitative importance of moral considerations. Using the recommended ethical weights in Tables 4.1 and 4.2 results in a tax recommendation of:

$$\tau = p^{*R}\frac{2.228 - 0.6}{0.6} = 2.71p^{*R)}. \tag{4.7}$$

This is a substantial tax equal to 271% of the equilibrium price, but based on the numbers chosen for illustrative purposes in Tables 4.1 and 4.2. Moral

considerations therefore potentially play a large role in policy decisions.

Fifth, the tax has a similar analytical structure as the tax developed in Chapter @ref{erroreconsurplus}. Assuming $\bar{\alpha} = \frac{1}{b_c}$ and $\bar{\delta} = \frac{1}{b_p}$, gives the result of week @ref{erroreconsurplus}.. The reason is that moral considerations scale the inverse demand and the inverse supply curve in as shown in Figure 4.3 and behavioural errors scale these curves in a similar proportional way. Note however, that the reasons for doing so are entirely different. Furthermore, prices do not adjust to the ethical weights. This is different in Chapter @ref{erroreconsurplus} where prices adjust with the behavioural errors of consumers and producers.

4.5 Externality taxation and moral considerations

4.5.1 Moral considerations related to the kind of externality

The last category of moral considerations that this chapter deals with stems from the application to the taxation of externalities by a regulator: the kind of external effect determines whether it should be regulated using a tax.

This category of considerations does justice to the intuition and to discussions on good governance in political science: for some external effects imposing a tax is more reasonable than for others. Intrinsic considerations, social customs, governance goals, value and externality tolerance and moral reasons might therefore 'overrule' (part of the) utilitarian external costs. As Medema (2007) (p. 339) discussing Mill writes:

> Indeed, even where government intervention may be warranted, the appropriate form of such action will vary, depending on the **nature** of the market failure." (PK: bold emphasis added).

Members of the advisory committee therefore have to decide on their weight for the external costs in the economic social surplus to do justice to these moral considerations. Define $0 \leq \eta_n \leq 1$ as the weight attached to the externality. Then the weighted external costs over all committee members are given by $\bar{\eta}e(Q)$, where $\bar{\eta} = \sum_{n=1}^{N} \omega_n \eta_n$ is the average weight over all committee members. Again, it is assumed for simplicity that the ethical weights η_n of committee members do not depend on the equilibrium quantity. An example for the motivated choice of ethical weights by the representatives is given in Table 4.3.

The average weight for the external costs is given by:

$$\bar{\eta} = \sum_{n=1}^{N} \omega_n \eta_n = 0.1 \sum_{n=1}^{10} \eta_n = 0.81. \tag{4.8}$$

Table 4.3: Scaling of the external costs. All numbers are chosen for illustration.

Committee member	Decision weight	Recommended weight external costs
Citizen representative.	$\omega_1 = 0.1$	$\eta_1 = 0.8$
Next generation representative.	$\omega_2 = 0.1$	$\eta_2 = 1$
Consumer representative.	$\omega_3 = 0.1$	$\eta_3 = 0.5$
Producer representative.	$\omega_4 = 0.1$	$\eta_4 = 0$
Governmental representative (chair of the committee).	$\omega_5 = 0.1$	$\eta_5 = 0.81$
Economist (specialization: environmental economics)	$\omega_6 = 0.1$	$\eta_6 = 1$
Philosopher (specialization: animal ethics)	$\omega_7 = 0.1$	$\eta_7 = 1$
Ecologist (specialization: biodiversity)	$\omega_8 = 0.1$	$\eta_8 = 1$
Climate activist representative.	$\omega_9 = 0.1$	$\eta_9 = 1$
Animal representative.	$\omega_{10} = 0.1$	$\eta_{10} = 1$

This implies that counted external costs are about 19% lower than the real marginal costs. Note that the numbers in Table 4.3 are chosen for illustration.

4.5.2 Application: pricing consumer externalities that are external to the market

The insights of Sections III.A and III.B can be used to adjust the economic social surplus and to investigate the implications for the optimal taxation of consumption externalities. For simplicity it is assumed that these externalities do not affect the generalized prices of consumers in the market. Adding adjusted external costs leads to the following adjusted economic social surplus:[5]

$$AES = \bar{\alpha} \int_0^{Q^*} p(Q)dQ - \bar{\delta} \int_0^{Q^*} s(Q)dQ - \bar{\eta} e(Q^*). \tag{4.9}$$

The implicit equilibrium expression for the first-best externality tax is given by (see Appendix C.2 for the interested reader):

$$\tau = p^{*R} \frac{\bar{\delta} - \bar{\alpha}}{\bar{\alpha}} + \frac{\bar{\eta}}{\bar{\alpha}} MEC, \tag{4.10}$$

where p^{*R} is the price in the regulated equilibrium. This tax expression has two analytical terms. The first component is equal to the tax Eq. (4.6) derived in Section II.

[5]Eq. (4.9) excludes monetized moral utilitarian external costs to remain consistent with the ethicizing economics approach as discussed in Section II. In principle these moral externalities could be added, but only for those members in the committee who consider the ECE approach to be more appropriate. This would result in a blended approach of the ECE and ETE approaches. For reasons of clarity and the argument given in Section II, this blended approach is not pursued in this paper.

The second term in Eq. (4.10) is related to the external costs. It multiplies the marginal external costs in the regulated equilibrium MEC with a factor $\frac{\bar{\eta}}{\bar{\alpha}}$, which is equal to $\bar{\eta}$ in the absence of moral considerations related to consumer benefits and governmental taxation. Suppose the external costs are fully counted implying $\bar{\eta} = 1$. This results in the following simplified tax expression:

$$\tau = p^{*R}\frac{\bar{\delta} - \bar{\alpha}}{\bar{\alpha}} + \frac{1}{\bar{\alpha}}MEC, \tag{4.11}$$

Counting consumer benefits for 50% then implies a recommended externality tax which is 100% higher compared to the recommended standard Pigouvian tax that internalizes marginal external costs. When giving recommendations for pricing consumption externalities, moral considerations can therefore interact in a non-linear way and potentially have substantial quantitative implications.

4.6 Conclusion and discussion

This chapter shows that moral considerations can have implications for economic policy valuation as these considerations go to the root of the normative matter: the specification of economic social surplus or the so-called economic 'pie'. This specification is at the heart of the concept of economic efficiency operationalized in micro-economic policy analysis. When externalities are absent, Section III shows that—from an economic perspective—taxation is justified because 'free' markets do not lead to the optimization of economic social surplus adjusted for moral considerations. This contrasts with the idea:

> ... that competitive equilibrium is efficient since the time of Marshall (Lazear 2000, 102).

Section IV extends the model to account for moral considerations related to external costs. This section also shows how the expressions for first-best Pigouvian consumption taxes are affected by moral considerations. When there are only moral considerations related to consumer benefits, this can have substantial impact on recommended externality taxes: half the consumer slice, twice the externality price. Moral considerations therefore interact with externality pricing and are not simply additive monetized external costs.

Ignoring moral considerations when calculating economic social surplus implicitly assumes moral excellence of consumers and producers. This leads to the ignorance of other normative perspectives in the valuation stage of a policy intervention in a market. This insight is especially relevant for the current use of practical cost-benefit analysis for public policies, because for these decisions moral considerations that transcend market valuations of consumers and producers are often relevant. Whether a proposed policy can be labeled as 'efficient' or 'optimal' therefore depends on the specification of the economic pie, which in turn depends on the underlying ethical assumptions that are made.

The proposed adjustment of economic social surplus in this paper is pragmatic in that it allows for a wide variety of normative views related to consumer, producer and governmental behavior. For the purpose of public policy evaluation such an impartial calculation of economic social surplus using a weighted average of normative perspectives seems reasonable, but of course this pragmatism is a normative perspective as well.

One could argue that in the proposed model non-utilitarian moral considerations related to duties of consumers and producers are implicitly monetized via the taxation rules. This is a valid concern, but comparatively speaking—and as stated in the introduction—the plural ethicizing economics approach is preferred over the approach that ignores these moral considerations altogether when calculating economic surplus ('ethical checkbox') or the approach that explicitly monetize the ignorance of these moral considerations ('economizing ethics') as it *directly* deals with the moral considerations when specifying the economic 'pie' without taking an external costs 'detour'. Furthermore, the proposed calculation of economic surplus still allows for an ethical checkbox or other social and law considerations that can overrule or complement tax recommendations based on economic social surplus calculations. The use of moral externality taxes is therefore not a substitute for action to reduce grave and tragic injustice and external costs in markets.

A benefit of the approach is that room is given for considerations related to voice and exit (Hirschman 1970). 'Exit' considerations are captured by market choices of consumers and producers. If they do not want to produce or consume the good they can choose something else. 'Exit' is properly accounted for in the behavioural choices of consumers and producers. 'Voice' considerations in our context are related to the different normative perspectives in society related to what should be counted as valuable. The ethical weights are helpful instruments to give different voices a place in the normative stage of the analysis.

Chapter 5

Regression Analysis in the Social Sciences

5.1 Introduction

Why should we have something as applied econometrics in the social sciences? That is because we have theories as in Chapters (2)–(4) and those theories contain variables such as in the direct utility function of equation (2.7):

$$U(Q, G) = \frac{A}{\alpha}\left(1 - e^{-\alpha Q}\right) + BG. \tag{5.1}$$

Here, the quantities for the market good Q and the outside good G are considered to be *known*—also often referred to as data. In theoretical work they are fictional or sometimes simulated. A certain amount of consumption of the market good and the outside good lead then some level of utility which economists want to maximise.

This second part of the course is therefore about using data to quantify (socio-economic) parameters. Moreover, we focus on measuring **causal** effects, instead of mere correlations. Note, that in an ideal world, we would like conducting experiments as to measure a causal relation of a phenomenon X on Y. However, we almost always only have observational data on, for example demand and prices. Therefore, the second part of this course and syllabus deals with difficulties arising from using observational data to estimate these causal effects and to rewrite models as (2.7) such that we can actually *use* data to tease out values for—in this case—A and α.

This chapter is organised as follows. The next section addresses the problem of finding a *relation* between some X and some Y. Here, we follow an example from the well-known textbook of @stock2003introduction where we look at the

relation between school class size and school class performance. At the same time, we also introduce some `STATA` commands. To do so, this section deals as well with the statistical framework that is needed for applied econometrics. Note that we assume that the reader already had a course in introductionary statistics and provide only the basic concepts most important for this course in Appendix D.

5.2 So, what is the problem?

As explained in the introduction above, applied econometrics aims to give the policy maker well-informed, and evidence based, values for variables she needs. She needs these variables basically for two separate things:

1. **Causal inference**: The policy maker wants to assess the effect of a change in one variable (typically called X) on another variables (often called Y).
2. **Prediction**: If you know what variable X is, what should Y then be?

Nowadays, most applied econometric techniques are concerned with causal inference, not so much with prediction. Even more, the techniques often applied are beneficial for correct causal inference, but might harm prediction. However, see that without correct causal inference (so knowing the **true** causal effect) prediction is always cumbersome. That is why current methods such as machine learning methods first focus on finding the correct causal mechanism (even without sometimes specifying what they may be) and then optimize prediction.

Thus, finding (causal) mechanism helps the scientist or policy maker in assessing the outcomes of a particular (policy) intervention. In the economic realm one could think about trying to assess the following quantities:

- To what extent do people eat less meat if we increase the prices with 1% (using a meat-tax)?
- If we increase Dutch dikes with one meter, how much less flood risk will there be?
- How much do classes perform better is we reduce class-sizes with one student?

5.2.1 A first encounter with `STATA`

For this section, we will focus on the last question. And this is an important question for policy as teachers are costly, but parents value school performance very highly. To start answering this question we *use* data. In `STATA` data is in a specific format, named `.dta`. Note that this format is not in a text format and cannot be read with the human eye. To start using `STATA` we first need to set the working directory to the appropriate directory[1]. You can do this by using the dropdown menu `File` in `STATA` and click on `Change working directly` but you can also do this with the following command:

[1] I suggest that for this—and every other course—you have a specific directory

```
cd  "/Users/tomba/Dropbox/Thomas/Colleges/M&T_AED_2022/"
```

Note here that this file path is used on an Apple or Linux system. On a Windows system you need forward slashes. It is also good to have subdirectories in your project/course folder for e.g. slides, data and tutorials. Now suppose you have the subdirectory `data` in your course folder[2] and in that data directory you have a file called `caschool.dta`.[3]

To import the data in STATA you make use of the `use` command (again, you can make use of the dropdown menus), as follows:

```
use "./data/caschool.dta", clear
```

Note that an option is added using the syntax `, clear`. The comma indicates that an option is expected and the verb `clear` indicates that memory of `STATA` should be cleared from data. That is because, `STATA` can only have one dataset in its memory.

In addition, we would like to know how the dataset looks like, for example what kind of variables it contains. We do this by invoking the command `describe`:

```
describe
```

```
Contains data from ./data/caschool.dta
 Observations:           420
    Variables:            18                  20 Feb 2017 13:10
-------------------------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
-------------------------------------------------------------------------------------------------
observat        float   %9.0g
dist_cod        float   %9.0g
county          str18   %18s
district        str53   %53s
gr_span         str8    %8s
enrl_tot        float   %9.0g
teachers        float   %9.0g
calw_pct        float   %9.0g
meal_pct        float   %9.0g
computer        float   %9.0g
testscr         float   %9.0g
comp_stu        float   %9.0g
expn_stu        float   %9.0g
str             float   %9.0g
```

[2]Often it is as well advisable to make a distinction between *derived* and *raw* data. Raw data is original data and derived data is data you have transformed or worked on. In principle, you do not want to change the original data!

[3]This dataset can be downloaded from Canvas.

```
avginc          float   %9.0g
el_pct          float   %9.0g
read_scr        float   %9.0g
math_scr        float   %9.0g
-------------------------------------------------------------------------------------
Sorted by:
```

This provides information about the amount of observations and variables and the names and types of variables. In this case variables are either a float (a real number) or a string (text). Note as well, that this kind of output is cumbersome and ugly and not fit directly for reporting. Later, we will try to make this look better in an automatic way.

The command `summarize` gives descriptive statistics. Suppose that in this case we are only interested in the variables `testscr` (average testscore by district) and `str` (the student-teacher ratio by district). Then we invoke this by:

```
summarize testscr str
```

```
    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
     testscr |        420    654.1565    19.05335     605.55     706.75
         str |        420    19.64043    1.891812         14       25.8
```

Now we see descriptive statistics for two variables, containing number of observations, the mean, the standard deviation and the minimum and maximum of each variable. For first insight in the relation between class size and class performance we might want to draw a so-called scatterplot. These type of plots relate the values of two variables in a two-dimensional way by giving the values as coordinates. The following syntax will do so.

```
scatter testscr str
```

And this provides the following `STATA` output.

This "cloud" of dots do not yield a clearly visible relation between class performance and class size. However, this can be deceptive. Often it is difficult to discern clear relations from raw data only. Therefore we need to resort to numerical evidence.

5.2.2 Numerical evidence

To assess whether there is a relation between class performance and class size as diplayed in Figure 5.1 we need numerical or statistical evidence. Before we start to engage in regression analysis, we first perform a rather simple analysis, but the underlying mechanisms are identical to that of regression analysis. We first create two groups: namely, districts with "small" (STR < 20) and "large" (STR ≥ 20) class sizes. Then we can adopt three relatively straightforward strategies here:

Figure 5.1: A scatterplot with tests-cores on the y-axis and student-teacher ratio on the x-axis

Table 5.1: Descriptive statistics of small and large classes

Class size	Average score	Standard deviation	n
Small	657.4	19.4	238
Large	650.0	17.9	182

1. Estimation
   - Here, we compare the average test scores in districts with low student-teacher ratios to those with high student-teacher ratios. So, we basically try to assess whether **average** behaviour is different.
2. Hypothesis testing
   - Now, we aim to **test** the "null" hypothesis that the mean test scores in the two types of districts are the same, against the "alternative" hypothesis that they differ.
3. Confidence intervals
   - This strategy estimates an interval for the **difference** in the mean test scores, so small versus large student-teacher ratio districts.

In `STATA` this data analysis would look like:

```
generate large = (str >= 20)
tabulate large, summarize(testscr) mean standard obs
```

```
            |         Summary of testscr
      large |        Mean   Std. dev.         Obs
------------+------------------------------------
          0 |   657.35126   19.358012         238
          1 |   649.97885   17.853364         182
------------+------------------------------------
      Total |   654.15655   19.053348         420
```

The first command *generates* a new variable called "large" and denotes an indicator being 0 if $str < 20$ and 1 if $str \geq 20$. The second command summarizes the testscore variable again, but now only gives the mean, standard deviation and the number of observation and does this by each value of the new variable `large`. Of course, this output is rather ugly and it is better to make a nice table such as Table 5.1.

Now, for all three strategies (estimation, testing, confidence intervals) we want to know something about the difference—usually denoted as Δ. Or, specifically:

1. For estimation: determine the Δ = difference between group means
2. For hypothesis testing: can we *reject* the null-hypothesis that $\Delta = 0$
3. For confidence intervals: can we construct a confidence interval for Δ

5.2.2.1 Estimation

Estimation is rather straightforward as we need to calculate the *difference* between the mean test scores within each group, or:

$$\bar{Y}_{small} - \bar{Y}_{large} = \frac{1}{n_{small}} \sum_{i=1}^{n_{small}} Y_i - \frac{1}{n_{large}} \sum_{i=1}^{n_{large}} Y_i \tag{5.2}$$

$$= 657.4 - 650.0 \tag{5.3}$$

$$= 7.4 \tag{5.4}$$

This basically mean subtracting the average scores of Table 5.1 (later we see how to this automatically in STATA). Now $\Delta = 7.4$ and we then have to ask ourselves whether this is a large difference in a real-world sense? Note that testscores seem to range from 600 to 800 and do not really have a direct meaning for us. A useful trick then is to look at the standard deviation (note that if things are normally distributed, 95% of all probability mass is within the range mean, plus or minus two times the standard deviation). In this case, the difference is about 1/3 of the standard deviation. A different way of looking at this is looking at the percentiles of testscores. In STATA this looks like:

```
tabstat testscr, statistics(p10 p25 p50 p75 p90)
```

```
    Variable |       p10       p25       p50       p75       p90
-------------+--------------------------------------------------
     testscr |   630.375       640    654.45   666.675     679.1
----------------------------------------------------------------
```

where the command `tabstat` asks for a tabulation of certain statistics and `px` gives the x-th percentile of that statistic. Now note that between the 50th and 65th percentile there is only 12 points. So given that, the difference of 7.4 is rather sizable. But whether this difference is big enough to be important for school reform discussions, for parents, or for a school committee is a question we cannot answer with this analysis.

5.2.2.2 Hypothesis testing

We also test the null-hypothesis that the difference $\Delta = 0$. For that we need a so-called difference-in-means test and compute the corresponding t-statistic[4],

$$t = \frac{\bar{Y}_s - \bar{Y}_l}{\sqrt{\frac{s_s^2}{n_s} + \frac{s_l^2}{n_l}}} = \frac{\bar{Y}_s - \bar{Y}_l}{SE(\bar{Y}_s - \bar{Y}_l)} \tag{5.5}$$

where $SE(\bar{Y}_s - \bar{Y}_l)$ is the *standard error* of $(\bar{Y}_s - \bar{Y}_l)$, the subscripts s and l refer to "small" and "large" STR districts, and $s_s^2 = \frac{1}{n_{small}} \sum_{i=1}^{n_s} (Y_i - \bar{Y}_s)^2$

[4] page 94 of syllabus statistics

We can compute this difference-of-means t-statistic by filling this in with the numbers of 5.1:

$$t = \frac{\bar{Y}_s - \bar{Y}_l}{\sqrt{\frac{s_s^2}{n_s} + \frac{s_l^2}{n_l}}} = \frac{657.4 - 650.0}{\sqrt{\frac{19.4^2}{238} + \frac{17.9^2}{182}}} = \frac{7.4}{1.83} = 4.05 \tag{5.6}$$

But then what? Well, recall that we **reject** a null-hypothesis when the critical value is below a certain threshold (usually 5%). In this case that is equivalent with stating that $|t| > 1.96$. So, we reject (at the 5% significance level) the null hypothesis that the two means are the same. We will come back to this procedure in Section 5.3.2.2.

5.2.2.3 Confidence interval

Finally, we can construct a 95% confidence interval for the difference between the means, which is:

$$(\bar{Y}_s - \bar{Y}_l) \pm 1.96 \times SE(\bar{Y}_s - \bar{Y}_l) = 7.4 \pm 1.96 \times 1.84 = (3.7, 11.0) \tag{5.7}$$

So what does this mean again. Well, two things. First, the 95% confidence interval for Δ doesn't include 0 and, second, the hypothesis that $\Delta = 0$ is rejected at the 5% level. We will come back to confidence intervals as well, but for now a confidence interval can be seen as an interval of numbers that will not be rejected as null-hypothesis.

5.2.3 Always be smart (and a bit lazy)

So, why give this rather simple procedure so much attention. That is because all "classical" statistics are constructed by these three elements and all statistical computer will always give, at least, these three. And they are as well very much related with each other. Once you know two of them, you know the third one as well.

Now, although the procedure is rather straightforward, it is also a bit cumbersome and prone to errors. Therefore, it is much easier to let STATA do it:

```
ttest testscr, by(large) unequal
```

```
Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
       0 |     238    657.3513    1.254794    19.35801    654.8793    659.8232
       1 |     182    649.9788    1.323379    17.85336    647.3676    652.5901
---------+--------------------------------------------------------------------
Combined |     420    654.1565    .9297082    19.05335    652.3291     655.984
---------+--------------------------------------------------------------------
```

```
    diff |             7.37241    1.823689                3.787296    10.95752
---------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   4.0426
H0: diff = 0                     Satterthwaite's degrees of freedom =  403.607

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0001          Pr(T > t) = 0.0000
```

So, in this case we want to assess the difference in test score by groups (being small and large classes), where we as well add the option `unequal`, which means that variance (or standard deviations) of both groups are unequal (and they are as Table 5.1 clearly shows).

Now try to find out for yourself that this output gives you the estimation of Δ of Eq. (5.4), the t-value and corresponding test outcome of Eq. (5.5) with the corresponding confidence intervals of Eq. (5.7).

5.3 Univariate regression

The three strategies we adopted in 5.2.2 for assessing the difference between groups directly translate to the case of regression analysis. Here we also look at estimation, hypothesis testing and confidence intervals. But before that we first look at the origin of the name regression in Subsection 5.3.1

5.3.1 Genesis: *regression towards the mean*

The name regression seems to have a negative connotation, as progress is in general seen as good and regress as bad. And actually this is true, the name regression was deliberatily given as to describe a negative process: in full *regression towards the mean*. The concept of regression was actually coined by Sir Francis Galton together with other statistical terms, such as correlation and deviation (Senn 2011). Galton was a notorious statistician who measured everything and else, including the length of french bread and the size of human skulls.

in 1886, Galton started to research the height of adult children with the height of their parents (Galton 1886). The original data can be seen in the scatterplot in Figure 5.2. What Galton expected was that the relation between the height of children and that of their parents was a one-to-one relation. On average children should receive the same height of their parents. So, in fact he expected a 45° line—the red line; a line with slope equal to 1.

However, he found consistently the blue line, a line with positive slope but lower than 1 (the blue line in Figure 5.3). That entails that, *on average* tall parents get tall children but not as tall as themselves. Of course, this goes as well the other way. Short parents get short children but not as short as themselves.

Figure 5.2: Relation heigh fathers and height children

Figure 5.3: Relation heigh fathers and height children

Galton coined this process *regression towards the mean*.[5] In the end we would all converge towards the mean and all look the same. For the Victorian Sir Frances Galton and his contemporaries in an age where income classes were highly separated this was truly a horror. Especially, because his cousin was Charles Darwin who actually claimed that species *diverged*. Of course, in Galton reasoning there is a mistake as this only models genetic influence and not *accidental* differences not influenced by genetics. Note that this analysis says something the average, but not about individual differences.

This regression towards the mean is now seen as a very important characteristics of regression models, and you can easily be fooled by it. It is now stated as:

> ... a concept that refers to the fact that if one sample of a random variable is extreme, the next sampling of the same random variable is likely to be closer to its mean

For instance, suppose that everything went really well for a course and you got a 9 for an examination. That does not mean that the next time you will do equally well (you will still do, but not that well). Or, your favorite footballclub does extreme well in a particular year (Leicester City FC comes to mind who became premier league champion in 2016). That does not mean that the next year it will do equally well, and so forth and so on.

5.3.2 Regression with one regressor

So, linear regression allows us to *estimate*, and make *inferences* about, *population* slope coefficients. Inference means drawing conclusions and population refers to the fact that we do not want to say something about our sample, but instead about our whole population. Ultimately our aim is to estimate the **causal** effect on Y of a unit change in X—but for now, just think of the problem of fitting a straight line to data on two variables, Y and X.

Similar to Subsection 5.2.2 we have three strategies to make inferences:

- We estimation the relation:
  - This now boils down to the question how we should draw a line through the data to estimate the (population) slope using Ordinary Least Squares (OLS—a specific an most common type of regression analysis)
  - And then we have to assess the advantages and disadvantages of OLS
- We could refer to hypothesis testing:
  - Very often this comes down to testing where the slope is zero. Namely, it the slope is zero, then the data does not show a relation between Y and X.
- Using confidence intervals:
  - This is related to constructing a confidence interval for the slope

[5] Modern statisticians actually see this as a form of shrinkage.

Before we look into this we first need some clarification on notation. As mentioned above, we would like to know the population regression line:

$$testscr = \beta_0 + \beta_1 STR, \tag{5.8}$$

where

$$\begin{aligned} \beta_1 &= \text{slope of population regression line} && (5.9)\\ &= \frac{\Delta Testscore}{\Delta STR} && (5.10)\\ &= \text{change in test score for a unit change in STR} && (5.11) \end{aligned}$$

Note the definition here of β_1. It gives the **marginal effect** of a change in STR on $testscr$. So the interpretation of the parameter β_1 is very straightforward. However, we do not know the population value of β_1 and we therefore have to estimate it using data.

In general, the population linear regression *model* is different as we add element u_1.

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \qquad i \ldots n \tag{5.12}$$

Now, X denotes the independent variable or regressor, Y the dependent variable, β_0 the intercept, β_1 the slope, and u_i the regression error. The regression error consists of omitted factors, or possibly measurement error in the measurement of Y. In general, these omitted factors are other factors that influence Y, other than the variable X.

5.3.2.1 Estimating with OLS

To estimate the population linear regression model we apply the ordinary least squares estimator. Again, as Figure 5.4 shows as well, a linear regression line is a straight line through points in a scatterplot. Actually, we want to draw that line such that the distance of all points to that line is minimized. See that in Figure 5.4 the distances between the points and the line are given by the $_i$'s, the regression errors. So, if we somehow can minimize all u_i's we are fine. But those distances could be both positive and negative and they might cancel each other out. Therefore, we first square the regression errors and then minimize (hence the name: ordinary least *squares*). Also, see from Eq. (5.12) that:

$$u_i = Y_i - (\beta_0 + \beta_1 X_i) \longleftrightarrow (u_i)^2 = [Y_i - (\beta_0 + \beta_1 X_i)]^2 \tag{5.13}$$

But how can we estimate β_0 and β_1 from data? For that we will focus on the least squares (ordinary least squares or OLS) estimator of the unknown parameters β_0 and β_1, which solves,

$$\min_{b_0, b_1} \sum_{i=1}^{n} [Y_i - (b_0 + b_1 X_i)]^2 \tag{5.14}$$

Figure 5.4: Drawing a straight line through data in a scatterplot

In fact, the OLS estimators of the slope β_1 and the intercept β_0 are:[6]

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} = \frac{s_{XY}}{s_X^2} \tag{5.15}$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X} \tag{5.16}$$

Although you do need to learn these formula's by heart some insightful comments can be given. First, if a parameter is estimated then it gets a on top. Second, the optimal $\hat{\beta}_1$ is equal to $\frac{s_{XY}}{s_X^2}$ and this is the sampling covariance between X and Y divided by the sampling variance of X. This is not a correlation as the units still depend on X and Y and therefore the slope can be large than 1 or smaller than -1, but it does say something about the relation between X and Y. Third, the constant is governed by the estimated parameter $\hat{\beta}_1$

From here we can predict the values $\hat{Y}_i$ and residuals $\hat{u}_i$ as they are:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i, \qquad i = 1, \ldots, n \tag{5.17}$$

$$\hat{u}_i = Y_i - \hat{Y}_i, \qquad i = 1, \ldots, n \tag{5.18}$$

When we apply this to our data cloud in Figure 5.1 then we get the following optimal population regression line:

where the estimated slope equals $\hat{\beta}_1 = -2.28$, the estimated intercept equals $\hat{\beta}_0 = 698.9$ and the total population regression line can be written as: $\widehat{TestScore} = 698.9 - 2.28 \times STR$. So, how to interpret the estimated slope and intercept now? First, the slope entails that districts with one more student per teacher on average have test scores that are 2.28 points lower (that is, $\frac{\Delta TestScore}{\Delta STR} = -2.28$). Secondly, the intercept (taken literally) means that, according to this estimated line, districts with zero students per teacher would have a (predicted) test score of 698.9. Now, this does not make any sense—it actually extrapolates the line outside the range of the data. In this case we can say that the intercept is not economically meaningful.

Now, how to fill in predictions? One of the districts in the data set is Antelope (CA) for which $STR = 19.33$ and $TestScore = 657.8$ Then the predicted value for the testscore is $\hat{Y}_{Antelope} = 698.9 - 2.28 \times 19.33 = 654.8$ and the resulting residual is $\hat{u}_{Antelope} = 657.8 - 654.8 = 3.0$

In `STATA` both the constant and the slope can be easily retrieved by:

```
regress testscr str, robust
```

```
Linear regression                               Number of obs     =        420
                                                F(1, 418)         =      19.26
```

[6]This result is given but is not all too difficult to prove. However, usually you do need these types of equations in your work.

Figure 5.5: Scatterplot and estimated regression line

```
                                                   Prob > F          =     0.0000
                                                   R-squared         =     0.0512
                                                   Root MSE          =     18.581

------------------------------------------------------------------------------
             |               Robust
     testscr | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         str |  -2.279808   .5194892    -4.39   0.000    -3.300945   -1.258671
       _cons |    698.933   10.36436    67.44   0.000     678.5602    719.3057
------------------------------------------------------------------------------
```

We will discuss the rest of this output later.

5.3.2.2 Hypothesis testing

We can assess the importance of the line as well with hypothesis testing. Again, recall that in in applied econometrics we will only **reject** the **null**-hypothesis, we do not accept an hypothesis based upon one statistical test only. So, we aim to test $H_0 : E(Y) = \mu_{Y,0}$ vs. $H_1 : E(Y) \neq \mu_{Y,0}$, where $\mu_{Y,0}$ is some specified quantity that we are interested in. Typically $\mu_{Y,0} = 0$ as this denotes no relation, but sometimes you could be interested in, e.g., whether $\mu_{Y,0} = 1$ when testing elasticities. Or you could be interested in other quantities.

Testing statistical hypotheses is often very confusing, because of two things. First, you actually test whether the data you have corresponds with the null-hypothesis. Or, in other words:

> What is the probability that your data (D) might be right *given* the null-hypothesis (H_0): $\Pr(D|H_0)$

And that is a strange concept. You first imagine a world H_0 with the data that it *should* provide and then test that imaginary world.

Secondly, there is the notation that often works confusing. First, we have the p-value which equals the probability of drawing a statistic (e.g., $\bar{Y}$) *at least as adverse* to the null (that is: your imaginary world) as the value actually computed with your data, **assuming** again that the null-hypothesis is true—again, you imaginary world. Secondly, there is the significance level of a test which is a *pre-specified* probability of incorrectly rejecting the null, when the null is actually true.

Now, suppose that you want to calculate the p-value based on an estimated coefficient $\hat{\beta}_1$, then you construct the following test:

$$p\text{-value} = \Pr_{H_0}[|\hat{\beta}_1 - \beta_{1,0}| > |\hat{\beta}_1^{act} - \beta_{1,0}| \tag{5.19}$$

where $\hat{\beta}_1^{act}$ is the value of $\hat{\beta}_1$ actually observed, and $\beta_{1,0}$ is the value of β_1 under the null-hypothesis (e.g., $\beta_{1,0} = 0$). Now, this is confusing, but in words it states that if you belief the null-hypothesis (the estimated value should be then $\hat{\beta}_1$), ($\hat{\beta}_1^{act}$) or even more extreme values can be estimated.

To test the null hypothesis H_0 we follow three steps. First, we need to compute the **standard error** of $\hat{\beta}_1$, which is an estimator of $\sigma_{\hat{\beta}_1}$. Using an ordinary least squares estimator, standard errors for coefficients are given by:

$$\sigma_{\hat{\beta}_1} = \sqrt{\frac{1}{n}\frac{\frac{1}{n-2}\sum_{i=1}^{n}(X_i - \bar{X})^2 u_i^2}{\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right]^2}}, \tag{5.20}$$

which is a rather daunting expression.

Second, we need to compute the t-statistic:

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{SE(\hat{\beta}_1)} \tag{5.21}$$

Finally, we need to calculate the p-value. To do so, we need to know the sampling distribution of $\hat{\beta}_1$, which we know is complicated if n is small, but typically you have enough observations to invoke the *Central Limit Theorem*. So, if n is large,

you can use the normal approximation (CLT) as follows

$$
\begin{aligned}
p\text{-value} &= \Pr_{H_0}\left[\left|\hat{\beta}_1 - \beta_{1,0}\right| > \left|\hat{\beta}_1^{act} - \beta_{1,0}\right|\right] && (5.22)\\
&= \Pr_{H_0}\left[\left|\frac{\hat{\beta}_1 - \beta_{1,0}}{SE(\hat{\beta}_1)}\right| > \left|\frac{\hat{\beta}_1^{act} - \beta_{1,0}}{SE(\hat{\beta}_1)}\right|\right] && (5.23)\\
&= \Pr_{H_0}[|t| > |t^{act}|] && (5.24)\\
&\simeq \text{probability under left + right } N(0,1) \text{ tails} && (5.25)
\end{aligned}
$$

where $SE(\hat{\beta}_1)$ again equals the standard error of $\hat{\beta}_1$, denoted with $\sigma_{\hat{\beta}_1}$

So, if you know $\hat{\beta}_1$ and $\sigma_{\hat{\beta}_1}$ you can calculate this. However, computers are much faster, in doing to. For example, suppose we want to test whether $\beta_{1,0} = 0$ using the regression output displayed above which gives $\hat{\beta}_1 = -2.28$ and $\sigma_{\hat{\beta}_1} = 0.52$. That is step 1. Note that `STATA` already calculated the standard error of Eq. (5.20). For the next step we need to compute the t-statistic, which is:

$$t = \frac{2.28 - \beta_{1,0}}{0.52} = \frac{2.28 - 0}{0.52} = -4.39. \tag{5.26}$$

then the p-value can be seen from Figure 5.6:

That is, for large n (and typically we have that), the p-value is the probability that a $N(0,1)$ random variable falls outside $|\hat{\beta}_1^{act} - \beta_{1,0})/\sigma_{\hat{\beta}_1}| = |t|$. That is the blue areas under the normal distribution and they entail a probability *mass*. Now, if both surfaces on the sides are not larger than 2.5%, then we can reject the null-hypothesis against a 5% significance level. Now, the computer output above gives a p-value of 0.000, which a bit strange. The p-value is actually not zero, but a very small number and definitely smaller than 0.05, so we can *reject* the null-hypothesis being $\beta_{1,0} = 0$ at a 5% significance level (and at a 1% and 0.1% significance level as well). Now, if the t-statistic is exactly 1.96 in absolute value, then the p-value is 0.05. So, to repeat the steps, but now using computer output for testing the hypothesis that $\beta_1 = \beta_{1,0}$

1. Get the standard error from computer output
2. Compute the t-statistics as in Eq. (5.21)
3. Get the corresponding p-value. Or, reject the null-hypothesis at the 5% significance level if $|t^{act}| > 1.96$.

Now there is a link between the p-value and the significance level. The significance level is pre-specified. For example, if the pre-specified significance level is 5%, then you reject the null hypothesis if $|t| \geq 1.96$ or equivalently, you reject if $p \leq 0.05$. The p-value is sometimes called the marginal significance level. Often, it is better to communicate the p-value than simply whether a test rejects or not—the p-value contains more information than the "yes/no'' statement about whether the test rejects.

Figure 5.6: Calculating the p-value of a two-sided test when $t^{act} = -4.38$

But recently there has been some debate about using p-values (Amrhein, Greenland, and McShane 2019). Why should you use a 5% significance level, what is so special about that number? It is not better just to report coefficients and standard errors? Figure 5.7 shows a figure from the journal Nature and how scientists across all fields nowadays see p-values and statistical significance. This is not to say that testing does not matter, but more reporting. First of all, p values in themselves do not contain that much information. In the regression output of above you p-values being equal to 0.000 which is not informative. Secondly, the cut-off point of 5% is a bit harsh and could to publications being published only with p-values below 0.05, leading to what is called publication bias.

5.3.2.3 Confidence intervals

The exact definition of confidence intervals is a bit tricky. Namely, a 95% confidence interval for $\hat{\beta}_1$ is an interval that contains the true value of β_1 in 95% of repeated samples. That means that confidence interval do not give a probability (even though we would like to interpret it that way). But you can state that every value within a confidence interval would not be rejected as null-hypothesis, while every value outside the confidence interval would be rejected. Now, if we know both $\hat{\beta}_1$ and $\sigma_{\hat{\beta}_1}$ (again using computer output), then a 95% confidence interval can be very easily constructed. For our regression of output of above this entails

$$\hat{\beta}_1 \pm 1.96 \times \sigma_{\hat{\beta}_1} = -2.28 \pm 1.96 \times 0.52 = [-3.30, -1.26]. \tag{5.27}$$

So, every value between -3.30 and -1.26 will **not** be rejected as null-hypothesis, while every value outside that interval will be rejected. Note that confidence interval is again automatically given by computer output. If one would like a confidence interval against another critical level, say against a 99% critical level, one can use the `level()` option

```
regress testscr str, robust level(99)
```

```
Linear regression                               Number of obs     =        420
                                                F(1, 418)         =      19.26
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0512
                                                Root MSE          =     18.581

------------------------------------------------------------------------------
             |               Robust
     testscr | Coefficient  std. err.      t    P>|t|     [99% conf. interval]
-------------+----------------------------------------------------------------
         str |  -2.279808   .5194892    -4.39   0.000    -3.624061    -.935556
       _cons |    698.933   10.36436    67.44   0.000     672.1137    725.7522
------------------------------------------------------------------------------
```

Figure 5.7: Critical review on the (mis)use of statistical significance

5.3.2.4 Regression with a dummy

Sometimes a regressor is binary, meaning an indicator or dichotomous (0/1) variable. Let's go back again to Subsection 5.2.2, where we created such a binary variable with small and large class sizes ($X = 1$ if class size is small, $X = 0$ if not). Other possible example are gender ($X = 1$ if female, $X = 0$ if male) or being treated or not ($X = 1$ if treated, $X = 0$ if not). We refer to these types of variables as being **dummy** variables—and they are very often used in the social sciences.

Now, suppose we have a population regression model that looks like:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \tag{5.28}$$

Where Y denotes, e.g., test scores, and where X, e.g., denotes a dummy variable for a large class (so, if $STR \geq 20$ then $X_i = 1$; otherwise $X_i = 0$), so there is then only two possibilities:

1. For small class size there should hold that $X_i = 0$ yielding that $Y_i = \beta_0 + u_i$. Namely $\beta_1 \times X_i = \beta_1 \times 0 = 0$. That means automatically that the expectation of Y_i is the constant, being β_0. Another way or writing is that the expectation of model (5.28) *conditional* on the fact that $X_i = 0$ is $\mathbb{E}(Y_i \mid X_i = 0) = \beta_0$.
2. For large classes there should hold that $X_i = 1$ yielding that $X_i = 1$, $Y_i = \beta_0 + \beta_1 + u_i$. This means the expectation of model (5.28) *conditional* on the fact that $X_i = 1$ is $\mathbb{E}(Y_i \mid X_i = 1) = \beta_0 + \beta_1$

So a regression with a dummy as independent variable gives two different *constants*, for each group (small/large classes) one. You can interpret this as a level-effect (only the level changes, not the slope as there is none here). The interpretation of β_1 is in this case rather special and can be denoted as:

$$\beta_1 = \mathbb{E}(Y_i \mid X_i = 1) - \mathbb{E}(Y_i \mid X_i = 0), \tag{5.29}$$

Which is just the population difference in group means.

If we go back to our example with $X_i = 1$ if $STR \geq 20$ and 0 otherwise then we get the following regression output

```
regress testscr large, robust
```

```
Linear regression                               Number of obs     =        420
                                                F(1, 418)         =      16.34
                                                Prob > F          =     0.0001
                                                R-squared         =     0.0369
                                                Root MSE          =     18.721

------------------------------------------------------------------------------
             |               Robust
```

```
     testscr | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       large |   -7.37241   1.823578    -4.04   0.000    -10.95694   -3.787884
       _cons |   657.3513   1.255147   523.72   0.000     654.8841    659.8184
------------------------------------------------------------------------------
```

Now, note that this is the same output ($\Delta = -7.4$, $\sigma_\Delta = 1.82$ and t-statistic is -4.04) as when we did the difference in means test in Subsection 5.2.3. To conclude, this is just another way (and much easier) to do a difference-in-means analysis. And this directly carries over for the situation when we have additional regressors.

5.4 Least squares assumptions for causal inference

As stated at the start of Section 5.2 applied econometrics focuses on finding a **causal** effect. But how do you do know that the $\hat{\beta}_1$ you estimate using the population regression model of Eq. (5.30) is indeed a causal effect. In other words, if you change X_i with one unit, will Y_i then change with β_1 **in reality**?

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \qquad i = 1 \dots n \tag{5.30}$$

Fortunately, there is a small set of assumptions that indeed lead to such a causal interpretation. The so-called three least squares assumptions, being:

1. The conditional distribution of u given X has mean zero, that is, $E(u \mid X = x) = 0$.
   - We also refer to this assumption as the **conditional mean independence** assumption
   - This assumptions implies that $\hat{\beta}_1$ is truly *unbiased*
2. $(X_i, Y_i), i = 1 \dots n$ are i.i.d.
   - This is true if X, Y are collected by simple random sampling
   - This delivers the sampling distribution of $\hat{\beta}_0$ and $\hat{\beta}_1$—again with a relatively large number (say $n > 50$ the sampling distribution can very well be approximated by a normal distribution
3. Large outliers in X and/or Y are rare.
   - Outliers can result in meaningless values of $\hat{\beta}_1$

We will first discuss these three least squares assumptions and then give some other assumptions as well as you frequently encounter them

5.4.1 Least squares assumption 1: conditional mean independence

This first assumptions states that $E(u \mid X = x) = 0$ and is conceptually the most difficult one to grasp. Loosely speaking, it states that the regression error

u is *not* related with the independent variable X. They are not related with each other. Another way of looking at this is displayed in Figure 5.8. Here, whatever the value of student-teacher ratio is, the expectation of the outcome variable (test scores) is always centered around the population regression line. So, on average, you always predict according to *your* model, for each value of X.

Figure 5.8: Condition mean independence assumption

Now, when is this assumption violated. Consider again the population regression model: $TestScore_i = \beta_0 + \beta_1 STR_i + u_i$, where u_i denotes other factors. Now, these other factors can be everything and else. And you should ask yourself whether it is plausible that $E(u|X = x) = 0$ for **all** these other factors?

This assumption lies as well at the heart of experimental settings. Namely, consider a theoretical ideal randomized controlled experiment, where: 1. X is *randomly* assigned to people (students randomly assigned to different size classes or patients randomly assigned to medical treatments). 2. Because X is assigned randomly, all other individual characteristics—the things that make up u—are *independently* distributed of X by definition. 3. Thus it automatically follows that: $E(u \mid X = x) = 0$

Now, both in actual experiments, or with **observational** data, we will need to think hard about whether $E(u|X = x) = 0$ holds. Chapters 6 and 7 provide various examples where this assumption is violated. However, if this assumption is violated it means that you have **biased** inference, which boils down to the fact that your estimated $\hat{\beta}_1$ is not the one that you want and that correct inference based upon this estimated cannot be done.

5.4.2 Least squares assumption 2: independenty and identically distributed

The second least squares assumptions deals with actual sampling of your data, both the dependent (Y) and independent (X) variables. that entail that $(X_i, Y_i), i = 1 \ldots n$ should be *i.i.d.*. This assumptions arises automatically if the entity (individual, district) is sampled by simple *random sampling*. There are possible violations to this assumptions. For example, you sample via your friends on social media (snowballing), or observations are not independent but are correlated, which arises very frequently in the context of temporal correlation or spatial correlation.

The consequence of violating the *i.i.d.* assumption is less severe then violating the conditional mean independence assumption. It leads to wrong *standard errors*, not to biased estimations.

5.4.3 Least squares assumption 3: Large outliers are rare

The third and final least square assumption for causal inference is that large outliers are rare. Large outliers are not well defined and depend on the size of both Y and X, but in general it can be seen as an *extreme* value of X or Y. The problem is that such a large outlier can strongly *influence* the results and in general it can be stated that OLS can be rather sensitive to an outlier. Consider the two population regression lines in 5.9. The flat one (with $\hat{\beta}_1 = 0$) does not take the isolated observation in the upper right corner into account. The one with the positive slope does. Now, clearly the one isolated observation in the upper right corner clearly matter to a large extent and is an important driver for the results of the ordinary least squares estimator.

However, this does not automatically entail that the isolated observation should be deleted. What it does matter is that one should go back to her data and investigate whether the outlier could be a mistake—perhaps a typo made when preparing the data or something that went amiss when converting the data from an `Excel` format to a `STATA` format.

5.5 Other least squares assumptions

Oftentimes, two other least squares assumptions are frequently encountered. However, keep in mind that you do *not* need them for causal inference. They are the assumptions of homoskedasticity and normality.

5.5.1 Homoskedasticity

Homoskedasticity is concerned with the standard errors. Its definition is if $var(u \mid X = x)$ is constant—that is, if the variance of the conditional distribution of u given X does not depend on X—then u is said to be homoskedastic. Otherwise, u is heteroskedastic.

Figure 5.9: Effect of outliers on OLS estimations

Figure 5.10: Homoskedastic standard errors

Consider Figure 5.10. Clearly the variance *around* the population regression line is everywhere the same, regardless the value of student-teacher ratio (X). Recall, that $E(u \mid X = x) = 0$ so u satisfies Least Squares Assumption 1. Now, in addition we also assume that the variance of u does not depend on x. This is the case of homoskedasticity

Figure 5.11: Heteroskedastic standard errors

Now consider Figure 5.11. Now clearly the variance around the population regression line increases in student-teacher ratio (X). So, $E(u \mid X = x) = 0$ is still satisfied, but the variance of u does now depends on x. u is now said to be heteroskedastic.

Very often data or model in the social sciences are heteroskedastic. For example, wages are usually heteroskedastic in the amount of education consumed. Figure 5.12 shows the relation between years of education and wages, and the larger the years of education the larger hourly earnings (wages) are—as you would assume would be. But the variance also incrases in years of education. That

Figure 5.12: Wages versus years of education

is because you can easily predict wages when workers enjoyed very few years of schooling—usually those are just above minimum wages—but the spread becomes much wider when years of schooling goes up.

Figure 5.13: Heteroskedasticity in Califonation schools?

Is this now the case for our Californian school dataset. If we look again at the scatterplot between test scores and student-teacher ratios in Figure 5.13, then that is very difficult to see. But then again, does it matter whether you face heteroskedasticity or homoskedasticity.

Note that so far we have (without saying so) assumed that u might be heteroskedastic Recall again the three least squares assumptions:

1. $E(u \mid X = x) = 0$
2. $(X_i, Y_i), i = 1, \ldots, n$, are i.i.d.
3. Large outliers are rare

They do not say anything about homo- or heteroskedasticity and Because we

have not explicitly assumed homoskedastic errors, we have implicitly allowed for heteroskedasticity.

But what if the errors are in fact homoskedastic? Then in fact yYou can prove that OLS has the lowest variance among estimators that are *linear* in Y. The formula for the variance of $\hat{\beta}_1$ and the OLS standard error simplifies: If $var(u_i \mid X_i = x) = \sigma_u^2$, then

$$var(\hat{\beta}_1) = \frac{\sigma_u^2}{n\sigma_X^2} \tag{5.31}$$

which is much simpler than Eq. (5.20). Again note that $var(\hat{\beta}_1)$ is inversely proportional to $var(X)$: more spread in X means more information about $\hat{\beta}_1$—we discussed this earlier but it is clearer from this formula.

But what does this mean for estimation. Note that `STATA` does not automatically apply Eq. (5.20) for its standard errors, but uses the simpler version Eq. (5.31) instead. But if invoke the `, robust` option, as we already did above, `STATA` computes heteroskedasticity-robust standard errors. So if you do not, `STATA` computes homoskedasticity-only standard errors.

The bottom line is that the errors are either homoskedastic or heteroskedastic and if you use heteroskedastic-robust standard errors, you are fine. Namely:

1. If the errors are heteroskedastic and you use the homoskedasticity-only formula for standard errors, your standard errors will be wrong (the homoskedasticity-only estimator of the variance of $\hat{\beta}_1$ is inconsistent if there is heteroskedasticity).
2. The two formulas coincide (when n is large) in the special case of homoskedasticity
3. So, you should **always** use heteroskedasticity-robust standard errors.

5.5.2 Normal distributed regression term

Finally, in many introductionary statistic courses, normal distributed error terms are assumed, which facilitates testing with small samples. So, u should be distributed $N(0, \sigma^2)$. If you have a reasonable amount of observations ($n > 50$), you do not need this assumptions, and especially not for causal inferences

5.6 Measures of fit

A natural question that might arise is how well the population regression line fits or explains the data. For ordinary least squares estimators there are two regression statistics that provide complementary measures of the quality of fit:

1. The regression R^2: This measures the fraction of the variance of Y that is explained by X; it is unitless and ranges between zero (no fit) and one (perfect fit)

2. The standard error of the regression SER: This measures the magnitude of a typical regression residual in the units of Y.

5.6.1 The regression R^2

The regression R^2 is the fraction of the sample variance of Y_i "explained" by the regression. To see this, first note that $Y_i = \hat{Y}_i + \hat{u}_i$ or the observation is equal to the OLS prediction plus the predicted residual. In this notation, the R^2 is the ratio between the sample variance of $\hat{Y}$ and the sample variance of Y. Here we make use of the following equity: Total sum of squares = explained "SS" + Residual "SS", where we can now define R^2 as:

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \overline{Y}\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2}. \tag{5.32}$$

Now if $R^2 = 0$ then that means $ESS = 0$ and if $R^2 = 1$ then that means $ESS = TSS$. So, by definition yields $0 \leq R^2 \leq 1$. There is one additional remark to make and that for an univariate regression model (so with single X on the right side), R^2 equals the square of the correlation coefficient between X and Y.

5.6.2 The Standard Error of the Regression

The standard error of the regression is defined as:

$$SER = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2} \tag{5.33}$$

In comparison with the R^2, the SER is measured in the units of u, which are actually the units of Y. It measures the average "size" of the OLS residual (so the average 'mistake' made by the OLS regression line in absolute terms).

Often the *root mean squared error* ($RMSE$) is used, which is very closely related to the SER:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2}, \tag{5.34}$$

where $RMSE$ only differs from the SER in the *degrees of freedom*.

If we again look at our regression output:

```
regress testscr str, robust
```

```
Linear regression                               Number of obs     =        420
                                                F(1, 418)         =      19.26
```

```
                                                  Prob > F          =    0.0000
                                                  R-squared         =    0.0512
                                                  Root MSE          =    18.581

------------------------------------------------------------------------------
             |               Robust
     testscr | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         str |  -2.279808   .5194892    -4.39   0.000    -3.300945   -1.258671
       _cons |    698.933   10.36436    67.44   0.000     678.5602    719.3057
------------------------------------------------------------------------------
```

then we see that the $R^2 = .05$. So, only 5% of all variation in test scores is explained. This of course makes sense as potential many important variables are not included in the model. However, this does not automatically mean that the impact is biased and especially that the STR is unimportant in a policy sence. Again, we focus on causal inference, not on making a good model for prediction. The $RMSE = 18.6$ indicates that the average error made is 18.6 test score units, which can be seen as sizable. In Chapter 6 we include other, and important, variables and what we then of course will see is that the R^2 increases and the SER decreases.

5.7 Conclusion and discussion

Regression

Chapter 6

Modeling in the Social Sciences

In Chapter 5 we discussed the origins of, working of and assumptions behind univariate regression. That is, a regression model with only one independent variable X on the right hand side.[1] However, and especially in the social science, you almost always see regressions many independent variables. Depending on the field, these variables can be called control variables, confounding factors or moderator variables. But why are these variables included? Is it only to improve model performance or are there other reasons? Section 6.1 deals with this question where after Section 6.2 shows how you can include additional variables in a *multivariate regression model* and especially how you should interpret them. Section 6.3 extends the multivariate regression model and shows how you can actually use this model to estimate a broad range of non-linear economic models. Section 6.4 discusses the use of multiple dummy variables (see again Subsection 5.3.2.4) in a way that economists refer to as *fixed effects*. The last section concludes and provides a futher discussion of the benefits and limitations of multivariate regression models.

6.1 Why more independent variables?

So, why do we include more variables? One possible answer is because it makes a better predictive model. That is, a model that is able to explain the variation in the dependent variable Y better.[2] So, the R^2 increases. But, as argued in

[1] With right hand side we mean on the right side of the equal sign =. It is often abbreviated with RHS.

[2] This is not entirely true. Increasing the R^2 explains **in-sample** variation better, not necessarily **out-of-sample**. The latter is really what matters for prediction and this is the focus of many machine learning techniques. Note that this argument is directly related with the regression towards the mean argument made in Subsection 5.3.1.

Chapter 5 we are not so much interested in prediction, but more in establishing a **causal** relation between X and Y. So, if you change X (and only X) does Y changes and then with how much?

Although economists often claim that they are the only (social-)science which is focused on causality and provides a statistical framework for that, there are other approaches to causality as well. One that is often used in other sciences is the approach of the mathematican Judea Pearl (Pearl 2009). This approach focuses on the use of Directed Acyclical Graphs (DAGs), which is graphical visualisation of causality chains (or, what impacts what). We borrow this approach for the most simple setting as explained in Figure 6.1. Here, we go back to our Californian school district dataset again, where we still are interested in the effect of class size on school performance. So, we suppose that that there is an effect from student teacher ratio on test scores as displayed with an directed arrow in Figure 6.1. We also know that the R^2 of that regression model was rather low (5%), so by default there must be other but yet unknown factors, let us name them for now U, that influence test scores as well (so a directed arrow going from U to test scores).

Figure 6.1: Unrelated omitted variables

Now we are fine with this is as long as U does **not impact** the student teacher ratio. Then, there is still an isolated effect of student teacher ratio on class size and that is exactly what we want to measure. However, if there is directed arrow going from U into STR as depicted by Figure 6.2, then the effect of student teacher ratio is not isolated anymore. Essentially, the effect of student teacher ratio on class size is composed out of two parts:

1) The **causal** effect on student teacher ratio on class size captured by the chain STR $\longrightarrow$ testscore. The one we are after.
2) The impact of the unknown variables on test scores. As we have not modeled them in our regression model, the effect is captured by the chain $U \longrightarrow$ STR $\longrightarrow$ testscore

Economists refer to this phenomenon as **omitted variable bias**, whilst in the statistical world, this is as often called confounding variables or the **the confounding fork** (McElreath 2020) and it, unfortunately, occurs very often.

Figure 6.2: Related omitted variables

So, when **U** is a *common* cause for both student teacher ratio and test scores there is omitted variable bias. If we go back to our population regression model as follows:

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \tag{6.1}$$

then we know that the error u arises because of factors that influence Y but are not included in the regression function; so, there are *always* omitted variables. But do not always lead to bias. For omitted variable bias to occur, the omitted factor, let's call it Z[3], must be:

1. A **determinant** of Y (i.e. Z is part of u)
2. A **determinant** of the regressor X (*at least*, there should hold that $corr(Z, X) \neq 0$)[4]

[3] Z can be both known or unknown, so that is why we change from U to Z

[4] In econometric textbooks, as, e.g, in Stock, Watson, et al. (2003), this condition is weakened to only being correlation (Z and X are correlated). However, if the directed arrow

Thus, both conditions must hold for the omission of Z to result in omitted variable bias.

Now, in our Californian district school dataset we have many more variables. One of them is variable that measures the english language ability (whether the student has English as a second language). Note that in California there are many migrants, especially from Latin-America. Now, you can readily argue that not having English as first language plausibly affects standardized test scores: so, Z is a **determinant** of Y. Moreover, immigrant communities tend to be less affluent and thus have smaller school budgets—and, therefore, higher STR: Z is most likely as well a **determinant** of X.

So, most likely, our original estimation from Chapter 5, $\hat{\beta}_1$, is biased (so not the true causal effect). But can we say something about the direction that bias? Yes, but the argument tend to become very quickly rather complex. In this case, note that districts with more migrant communities tend to have (i) higher class sizes and (ii) lower test scores. So, to the original estimation they added a *negative* effect. Thus, following this reasoning, the "true" effect must be less negative. Now, especially with negative signs this becomes rather complex, so if common sense fails you, there is the following formula:

$$\hat{\beta}_1 \xrightarrow{p} \beta_1 + \frac{\sigma_u}{\sigma_X}\rho_{Xu}, \tag{6.2}$$

where you should focus on the sign of the correlation between X and the regression residual u (all standard errors, σ, are always positive by default). Now, the first least squares assumption states that $\rho_{Xu} = 0$—no correlation between the regressor and the regression residual. But now there is because of omitted variable bias. And because the negative relation between immigrants communities and school performance ρ_{Xu} should be negative. And because the original estimation from Chapter 5 was already negative to begin with the "true" β_1 should be less negative. In conclusion, districts with more English learning students (i) do worse on standardized tests and (ii) have bigger classes (smaller budgets), so ignoring the English learning factor results in overstating the class size effect.

You might wonder whether this is actually going on in the Californian district school data. To see this, Figure 6.3 offers a cross tabulation of test scores by class size and percentage English learners.

Now, the table depicted in Figure 6.3 is complex in its various dimensions. We have our two categories of class size (small and large), together with the difference in test scores, but we now stratify this by four categories of percentage English learners. There are several important observations to make here:

goes from STR into U in Figure 6.2 then that would lead to something else than omitted variables, namely to a difference between a direct (STR $\longrightarrow$ testscore) and an indirect effect (STR $\longrightarrow U \longrightarrow$ testscore).

TABLE 6.1 Differences in Test Scores for California School Districts with Low and
Student–Teacher Ratios, by the Percentage of English Learners in the

	Student–Teacher Ratio < 20		Student–Teacher Ratio ≥ 20		Dif
	Average Test Score	*n*	Average Test Score	*n*	Dif
All districts	657.4	238	650.0	182	
Percentage of English learners					
< 1.9%	664.5	76	665.4	27	−
1.9−8.8%	665.2	64	661.8	44	
8.8−23.0%	654.9	54	649.7	50	
> 23.0%	636.7	44	634.8	61	

Figure 6.3: Cross tabulation of test scores by class size and percentage English learners

1) districts with *fewer* English Learners (so less migrants) have on average *higher* test scores (what we assumed above);
2) districts with *fewer* English Learners (so less migrants) have *smaller* classes (what we assumed above);
3) the effect of class size with comparable percentages English learners is still (mostly negative), but not as much as we compares for all districts together (the Difference-column). This confirms our reasoning that our original estimate was too negative.

No, as already mentioned above omitted variable bias occurs very often. So, how to correct for this such that the bias disappaers. In general, there are strategies:

1. we can run a randomized controlled experiment in which treatment (STR) is randomly assigned: then percentage English learners ($PctEL$) is still a determinant of test score, but by construction $PctEL$ should be uncorrelated with STR. Unfortunately, is very difficult to randomize class size and often this strategy is just not attainable as being too costly or unethical (this accounts for all sciences);
2. we can adopt the cross tabulation approach of above, with finer gradations of STR and $PctEL$. Then by construction, within each group all classes have the same $PctEL$ so we control for $PctEL$. A disadvantages is that one needs many observations, especially when one wants to stratify upon other variables as well;
3. finally, and perhaps the easiest approach, we can use a population regression model in which the omitted variable ($PctEL$) is no longer omitted. We just include $PctEL$ as an additional regressor in a multiple regression model. This is what the next section deals with. Obviously, a disadvantage of this approach is that you need observations for the omitted variable.

6.2 Multivariate regression analysis

So, if we have information about an important omitted variable, as in the case of the size of migrant communities in the example above, then we can use that information in a multivariate population regression model. In the case of two regressors, that would look like"

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i, i = 1, \ldots, n \tag{6.3}$$

where:

- Y is the dependent variable
- X_1, X_2 are the two independent variables (regressors)
- (Y_i, X_{1i}, X_{2i}) denote the ith observation on Y, X_1, and X_2.
- β_0 is the unknown population intercept
- β_1 is the effect on Y of a change in X_1, **holding** X_2 constant
- β_2 is the effect on Y of a change in X_2, **holding** X_1 constant
- u_i is the the regression error (omitted factors)

Now the only element that changes is the interpretation of a parameter, say β_1. In this case, it can still be seen as a 'slope' parameter, although now in 3-dimensional space, but it now states specifically that the other parameter(s) should be held constant. This does facilitate the interpretation of β_1. For example, consider changing X_1 by ΔX_1 while holding X_2 constant. That means that the population regression line before the change looks like:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2, \tag{6.4}$$

whilst the population regression line, after the change, looks like:

$$Y + \Delta Y = \beta_0 + \beta_1(X_1 + \Delta X_1) + \beta_2 X_2 \tag{6.5}$$

And if we take the difference, then the interpretation of β_1 boils down again to the marginal effect:$\Delta Y = \beta_1 \Delta X_1$. Or, $\beta_1 = \frac{\Delta Y}{\Delta X_1}$ when holding X_2 constant and, likewise, $\beta_2 = \frac{\Delta Y}{\Delta X_2}$ when holding X_1 constant. β_0 is now the predicted value of Y when $X_1 = X_2 = 0$

If we do this for the the Californian school district data, then the original population regression line was estimated as:

$$\widehat{TestScore} = 698.9 - 2.28STR \tag{6.6}$$

But if we now include include percent English Learners in the district ($PctEL$) to the model then the population regression 'line' becomes:

$$\widehat{TestScore} = 686.0 - 1.10STR - 0.65PctEL \tag{6.7}$$

Clearly, the effect of student teacher ratio becomes smaller (that is, less negative). And this is what should happen as reasoned above. The `STATA` syntax for a multivariate regression model is now rather straighforward. You basically add another to the regression equation, as below:

```
reg testscr str el_pct, robust
```

```
Linear regression                               Number of obs     =        420
                                                F(2, 417)         =     223.82
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4264
                                                Root MSE          =     14.464

------------------------------------------------------------------------------
             |               Robust
     testscr | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         str |  -1.101296   .4328472    -2.54   0.011     -1.95213   -.2504616
      el_pct |  -.6497768   .0310318   -20.94   0.000     -.710775   -.5887786
       _cons |   686.0322   8.728224    78.60   0.000     668.8754     703.189
------------------------------------------------------------------------------
```

Obviously, the effect of student teacher ration reduces with 50%! The interpretation of the rest of the statistical output, such as measures of fit and test statistics, follows in the subsections below.

6.2.1 Measures of fit for multiple regression

In multivariate regression models, there are four commonly used measures of fit.

1. The standard error of regression or the SER denotes the standard deviation of $\hat{u}_i$ and includes a degrees of freedom correction (degrees of freedom in this case denotes how many variables your have used and typically is denoted with k. The SER is defined as:

$$SER = s_{\hat{u}} = \sqrt{\frac{1}{n-k-1}\sum_{i=1}^{n}\hat{u}_i^2}, \tag{6.8}$$

   where k is the number of variables (including the constant) use in the regression model. Note that in the univariate regression model $k = 2$—the slope coefficient and the constant.
2. The root mean square error (RMSE) which denotes as well the stdandard deviation of $\hat{u}_i$ but now without degrees of freedom. We have seen this before in Eq. (5.34) and does not change.
3. The R^2 which measures the fraction of variance of Y explained by the independent variables. Again, we have seen this one before
4. The adjusted "adjusted R^2" (or $\bar{R}^2$) which is equal to the R^2 with a degrees-of-freedom correction that adjusts for estimation uncertainty. It can be formulated as:

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}\frac{SSR}{TSS}. \tag{6.9}$$

   Note that using this formulation, in a multivariate setting, it always should hold that $\bar{R}^2 < R^2$. Why do we care so much for the amount of variables that we use (k). That is because with each additional variable the R^2 always increases. And it is essential to notice that $k = n$, the $R^2 = 1$, so there is no variation left anymore. But that feels like cheating. You just have a parameter for each observation that you have, but such a model must be meaningless. Therefore, you always want to correct for the number of variables that you use.

In our Californian school district example that would amount to the following two outcomes. First for the univariate model:

$$TestScore = 698.9 - 2.28STR \tag{6.10}$$

$$R^2 = .05, SER = 18.6 \tag{6.11}$$

And then for the multivariate model.

$$TestScore = 686.0 - 1.10STR - 0.65PctEL \tag{6.12}$$
$$R^2 = .426, \bar{R}^2 = 0.424, SER = 14.5 \tag{6.13}$$

Note that all measures of fit increases. The $\bar{R}^2$ now indicates that 42% of all variation in test scores are explained. That is a *huge* improvement compared to the 5% explanatory power of the univariate case. That indicates that the $PctEL$ strongly correlates with testscores. But again, we are not so much interested prediction, but want to find the causal impact of class size instead. Another thing to notice here is that the R^2 and the $\bar{R}^2$ are very close. That is because the number of variables is much smaller than the number of observations $k \ll n$, so that the impact of k is not very big.

A list thing is a peculiarity of `STATA`. In the regression output of above `STATA` does not provide the $\bar{R}^2$. That is because of the option `, robust`. Without that option, the regression output would look like, giving both measures of fit.

```
reg testscr str el_pct
```

```
      Source |       SS           df       MS      Number of obs   =       420
-------------+----------------------------------   F(2, 417)       =    155.01
       Model |  64864.3011         2  32432.1506   Prob > F        =    0.0000
    Residual |  87245.2925       417  209.221325   R-squared       =    0.4264
-------------+----------------------------------   Adj R-squared   =    0.4237
       Total |  152109.594       419  363.030056   Root MSE        =    14.464

------------------------------------------------------------------------------
     testscr | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         str |  -1.101296   .3802783    -2.90   0.004    -1.848797   -.3537945
      el_pct |  -.6497768   .0393425   -16.52   0.000    -.7271112   -.5724423
       _cons |   686.0322   7.411312    92.57   0.000     671.4641    700.6004
------------------------------------------------------------------------------
```

Another option is to specifically ask `STATA` to display the $\bar{R}^2$ by invoking the command `display`, then some text (text always goes between strings), and finally the thing you want to see (`e(r2_a)`). Something like:

```
display "adjusted R2 = " e(r2_a)
```

```
adjusted R2 = .42368043
```

6.2.2 The least squares assumptions for multivariate regression

Now, it is easy to add other variables, so that the multivariate regression model now looks like:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + u_i, i = 1, \ldots, n \tag{6.14}$$

Suppose we are interested in β_1. How do we then know whether our estimation $\hat{\beta}_1$ is unbiased? For that we again resort to our least squares assumption, some of them will change a bit and we have to add a fourth one:

1. The first least squares assumptions changes slightly. Now, we state that the conditional distribution of u given all X_i's has mean zero, that is, $E(u|X_1 = x_1, \ldots, X_k = x_k) = 0$. So, β_1 is biased even another variable X_k is correlated with u. So, only variables has to be correlated with u and then all parameters are to a certain extent biased.
2. The second least squares assumption is more or less as before but now in a multivariate fashion, so the whole set of (X_{1i},...,X_{ki},Y_i), i =1,...,n, should be independent and identical distributed ($i.i.d$).
3. The third least squares assumptions states again that large outliers are rare for all variables included, so for all $X_1, \ldots, X_k$, and Y.
4. The fourth assumoption is news and states that there is no perfect multicollinearity. We discuss this further below.

6.2.2.1 Multicollinearity

Multicollinearity comes in two flavours; perfect and imperfect. The former functions as a multivariate least squares assumptions, whilst the latter oftentimes gives the largest problems. We start the discussion with perfect multicollinearity and then continue with the case of imperfect multicollinearity.

6.2.2.1.1 Perfect multicollinearity The official definition of perfect multicollinearity is that there is a **perfect linear combination** amongst your variables. That means that there is not one optimal solution, but instead many (actually, infinitely many) more. Let us illustrate by the following example. Suppose you include STR twice in your regression. Now, `STATA` produces then the following output:

```
reg testscr str str el_pct, robust
```

```
note: str omitted because of collinearity.

Linear regression                                      Number of obs     =        420
                                                       F(2, 417)         =     223.82
                                                       Prob > F          =     0.0000
                                                       R-squared         =     0.4264
                                                       Root MSE          =     14.464
```

```
------------------------------------------------------------------------------
             |               Robust
     testscr | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         str |  -1.101296   .4328472    -2.54   0.011     -1.95213   -.2504616
         str |          0  (omitted)
      el_pct |  -.6497768   .0310318   -20.94   0.000     -.710775   -.5887786
       _cons |   686.0322   8.728224    78.60   0.000     668.8754     703.189
------------------------------------------------------------------------------
```

See that `STATA` drops one of the STR variables. But why is that? See that the impact of twice this variable should be equivalent to:

$$\beta_1 STR = w_1\beta_1 STR + w_2\beta_1 STR = (w_1 + w_2)\beta_1 STR, \tag{6.15}$$

where w_1 and w_2 are weights chosen such that they satisfy the condition that $w_1 + w_2 = 1$. But there is an infinite number of combinations that satisfy this condition! So, there is not an optimal solution and one of these variables should be dropped.

The violation of no perfect multicollearity often occurs when using dummies (see again Subsection 5.3.2.4). Suppose that we regress $TestScore$ on a constant, D, and B, where:$D_i = 1$ if $STR \leq 20$, $= 0$ otherwise ; $B_i = 1$ if $STR > 20$, $= 0$ otherwise. This example is slightly more complex as there is no perfect correlation between B and D. However, the model contains as well a constant and that create a perfect linear combination, namely $B_i + D_i = 1$ and that is the definition of a constant ($\beta_1 \times 1$), so there is perfect multicollinearity in the model. Now, this example is a special case of the so-called dummy variable trap. Suppose you have a set of multiple binary (dummy) variables, which are mutually exclusive and exhaustive—that is, there are multiple categories and every observation falls in one and only one category (Freshmen, Sophomores, Juniors, Seniors, Other). If you include all these dummy variables and a constant, you will have perfect multicollinearity—the dummy variable trap.

There are possible solutions to the dummy variable trap:

1. Omit one of the groups (e.g., the Freshmen), or
2. Omit the intercept

In most cases you omit one of the groups (typically the one with the lowest value). This give the constant then the interpretation of the average value of that left-out category, where the dummy variables are then the relative differences to that left-out category.

Now, perfect multicollinearity usually reflects a mistake in the definitions of the regressors, or an oddity in the data. And, usually this is not a problem, because if you have perfect multicollinearity, your statistical software will let you know—either by crashing or giving an error message or by "dropping" one of the variables

arbitrarily and very often the solution to perfect multicollinearity is to modify your list of regressors such that you no longer have perfect multicollinearity.

6.2.2.1.2 Imperfect multicollinearity Now imperfect and perfect multicollinearity are quite different despite the similarity of the names. Imperfect multicollinearity namely, occurs when two or more regressors are very highly correlated. And if two regressors are very highly correlated, then their scatterplot will pretty much look like a straight line—they are collinear—but unless the correlation is exactly $\pm$ 1, that collinearity is imperfect. What this implies is that one or more of the regression coefficients will be imprecisely estimated. Why is that? That is because of the definition of the coefficient in a multivariate regression model. Namely, the coefficient on X_1 is the effect of X_1 **holding** X_2 **constant**, but if X_1 and X_2 are highly correlated, then there is very little variation in X_1 once X_2 is held constant. That means that the data are pretty much uninformative about what happens when X_1 changes but X_2 doesn't, so the variance of the OLS estimator of the coefficient on X_1 will be large. And this results in large standard errors for one or more of the OLS coefficients. But often this is very hard to detect. Are standard errors high because of imperfect multicollinearity, because the number of observations is very low, or because there is large variation in the data? The answer to this unfortunately boils down to reasoning, but before you start estimating your statistical models it always good to look at scatterplots and correlations between variables.

But what is a high correlation? With a reasonable amount of observations all correlations below 0.9 can be considered fine. In practice, only correlations between variables higher than say 0.95 start to impose problems.

6.2.3 Testing with multivariate regression models

6.2.3.1 Hypothesis tests and confidence intervals for a single coefficient in multiple regression

Recall from Subsection 5.3.2.2 that for hypothesis testing in a classical statistical framework we make use of the fact that $\frac{\hat{\beta}_1 - E(\hat{\beta}_1)}{\sqrt{var(\hat{\beta}_1)}}$ is approximately distributed as $N(0,1)$ according to the Central Limit theorem. Thus hypotheses on β_1 can be tested using the usual t-statistic, and confidence intervals are constructed as $\{\hat{\beta}_1 \pm 1.96SE(\hat{\beta}_1)\}$. And this finding carries over to the multivariate setting where for $\beta_2, \ldots, \beta_k$ we make use of the same framework. One thing to keep in mind is that $\hat{\beta}_1$ and $\hat{\beta}_2$ are generally not independently distributed—so neither are their t-statistics (more on this later).

Now, if we return to our Californian school district data set then we find that for the univariate case holds:

$$TestScore = \underbrace{698.9}_{10.4} - \underbrace{2.28}_{0.52} STR, \tag{6.16}$$

And the population regression "line" for the multivariate case is estimated as:

$$TestScore = \underbrace{686.0}_{8.7} - \underbrace{1.10}_{0.43} STR - \underbrace{0.650}_{0.031} PctEL \tag{6.17}$$

Remember, the coefficient on STR in Eq. (6.17) is the effect on $TestScores$ of a unit change in STR, holding constant the percentage of English Learners in the district. The corresponding 95% confidence interval for coefficient on STR in (2) is $\{-1.10 \pm 1.96 \times 0.43\} = (-1.95, -0.26)$. And the t-statistic testing $\beta_{STR} = 0$ is $t = -1.10/0.43 = -2.54$, so we reject the null-hypothesis at the 5% significance level. More evidence for the strength of the $PctEL$ variable can be seen from the fact that, under the null-hypothesis of $\beta_2 = 0$, the following must hold: t-statistic $= \frac{\hat{\beta_1}}{\sigma_{\hat{\beta_1}}} = \frac{0.65}{0.03} = 21.7$, which is a very high number for a t-statistic.

6.2.3.2 Tests of joint hypotheses

So, testing of single coefficients is just as before. Now in the Californian school district dataset there is as well a variable called $Expn$ denoting the expenditures per pupil. Consider the following population regression model:

$$TestScore_i = \beta 0 + \beta_1 STR_i + \beta_2 Expn_i + \beta_3 PctEL_i + u_i \tag{6.18}$$

The null hypothesis that "school resources don't matter" and the alternative that they do, corresponds to:

- $H_0 : \beta_1 = 0$ and $\beta_2 = 0$ vs
- H_1 : either $\beta_1 \neq 0$ or $\beta_2 \neq 0$ or both

This is a joint hypothesis specifying a value for two or more coefficients, that is, it imposes a restriction on two or more coefficients. In general, a joint hypothesis will involve q restrictions. In the example above, $q = 2$, and the two restrictions are $\beta_1 = 0$ and $\beta_2 = 0$. A "common sense" idea is to reject if either of the individual t-statistics exceeds 1.96 in absolute value. But this "one at a time" test isn't valid: the resulting test rejects too often under the null hypothesis (more than 5%)! That is because the t-statistics themselved are often not independent. Instead, we need a F-statistic, which tests all parts of a joint hypothesis at once. Unfortunately, these types of formulas can become quickly rather comples. Consider the F-test for the special case of the joint hypothesis $\beta_1 = \beta_{1,0}$ and $\beta_2 = \beta_{2,0}$ in a regression with two regressors:

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1,t_2} t_1 t_2}{1 - \hat{\rho}_{t_1 t_2}^2}\right) \tag{6.19}$$

where $\hat{\rho}_{t_1,t_2}$ estimates the correlation between t_1 and t_2. Reject when F is large (typically to be determined from large statistical tables). The F-statistic is large when t_1 and/or t_2 is large and the F-statistic corrects (in just the right way)

for the correlation between t_1 and t_2. The formula for more than two β's is nasty unless you use matrix algebra. There is a nice large-sample approximate distribution, which is the tail probability of the χ^2_q/q distribution beyond the F-statistic actually computed.

Now, `STATA` does this in a much easier way by invoking the `test` command **right** after the regression. So, for example, we want to test the joint hypothesis that the population coefficients on STR and expenditures per pupil ($expn$) are both zero, against the alternative that at least one of the population coefficients is nonzero.

```
reg testscr str expn_stu el_pct, r
test str expn_stu
```

```
Linear regression                               Number of obs     =        420
                                                F(3, 416)         =     147.20
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4366
                                                Root MSE          =     14.353

------------------------------------------------------------------------------
             |               Robust
     testscr | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         str |  -.2863992   .4820728    -0.59   0.553    -1.234002     .661203
    expn_stu |   .0038679   .0015807     2.45   0.015     .0007607    .0069751
      el_pct |  -.6560227   .0317844   -20.64   0.000    -.7185008   -.5935446
       _cons |   649.5779   15.45834    42.02   0.000     619.1917    679.9641
------------------------------------------------------------------------------


 ( 1)  str = 0
 ( 2)  expn_stu = 0

       F(  2,   416) =    5.43
            Prob > F =    0.0047
```

The output shows an F-statistic with $q = 2$ restrictions with outcome 5.43. Do not directly interpret this number, but know that Prob $> F = 0.0047$ gives the probability that under the null-hypothesis this outcome is produced. So the joint null-hypothesis that both types of expenditures are zero (at the same time), can be rejected at a 5% (and a 1%) significance level. Other types of joint tests can easily be constructed as well. For example, when you want to know whether both coefficient add up to 1, then you would state `test str + expn_stu = 1`. The final point to make is the F-test in the regression output itself. Here, that is for example `F(3, 416) = 147.20`. This is joint test that all variables, except the constant, have no impact. So, $\beta_i = 0$ for all i at the **same time**. It not

often that you come across a general regression F-test that does not reject the null-hypothesis. It namely implies that your independent variable do not contain information about the dependent variable.

And with the F-test, we now have discussed all regression outcome components displayed by `STATA`. Most of this information you do not need for your report but we will come back later to this.

6.3 Non-linear specifications

The model we are using is coined the *linear* regression model, and, indeed, one of the underlying assumptions is that the relations between the independent and dependent are linear. Consider the relation aain between test scores and class sizes in the Californian school district data. Using the following code (note now the `twoway` command that 'binds' a scatter plot with a population regression line):

```
graph twoway (lfit testscr str) (scatter testscr str)
```

Which provides the following `STATA` output.

Indeed, there might be evidence that the relation depicted in Figure 6.4—if anything—is linear. But, clearly that is not the case for the relation between test scores and average district income. Namely, the syntax below:

```
graph twoway (lfit testscr avginc) (scatter testscr avginc)
```

provides the following `STATA` output.

Figure 6.5 shows a non-linear relation, where the effect of income tapers off (note the resemble with Figure 1.1)—or, there is a marginal decreasing effect of average district income on average school test scores. Thus, in affluent neighborhood test scores are higher, but increasingly less so. Of course, you can still try to estimate this with a linear population regression line as in Figure 6.5, but this introduces a **bias**. The estimate does not capture that what you want. Namely, it now holds that $E(u \mid X = x) \neq 0$, because for small X, say $X < 10$, the residuals are negative, for medium sized Xs most residuals are positive and for large $X > 40$ all residuals are negative again. So, there is a clear relation between X and u and they fail to be independent. This particular form of bias is coined **specification bias**. There is another issue here and that is that the effect on Y of a change in X depends on the value of X—that is, the *marginal* effect of X is not constant.

To remedy the specification bias, we will use nonlinear regression population regression **functions** of X, or we estimate a regression function that is nonlinear in X. Here, it is important to see that we do so by *transforming* X, so the population regression 'line'. The estimator still remain a linear regression model.

We will analyse below two complementary and often adopted approaches:

Figure 6.4: A linear relation

Figure 6.5: A non-linear relation

1. Using **polynomials** to transform X. That means that the effect is approximated by a quadratic, cubic, or higher-degree polynomial. This approach as well governs to an extent so-called interaction effects which is a special case, where we multiply two different variables.
2. Using **logarithmic** transformations of X, where Y and/or X is transformed by taking its logarithm. Here, the main focus is on the interpretation of the $\hat{\beta}$s, as they change from a unit increase interpretation to a percentages interpretation which often can be found useful.

6.3.1 Polynomials

Our first approach to non-linear specification is applying polynomials of the variables that we suspect has a non-linear impact. If that is the independent variable X, the we can construct the following *linear regression* model by using polynomials:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_i^2 + \ldots + \beta_r X_i^r + u_i \tag{6.20}$$

Note again that this is just the linear regression model—except that the regressors are powers of X! So, in effect we transform the data—actually create new variables X^r—, but the specification in parameters remains linear. Estimation, hypothesis testing, etc. proceeds as in the multiple regression model using OLS. However, the coefficients are now a bit more difficult to interpret. Consider the example of above about the relation between test scores average district income, where $Income_i$ is defined as the average district income in the ith district (thousands of dollars per capita). For a quadratic specification, we specify the linear regression model as below:

$$TestScore_i = \beta_0 + \beta_1 Income_i + \beta_2 (Income_i)^2 + u_i \tag{6.21}$$

For a cubic specification the linear regression model becomes:

$$TestScore_i = \beta_0 + \beta_1 Income_i + \beta_2 (Income_i)^2 + \beta_3 (Income_i)^3 + u_i \tag{6.22}$$

First, we focus on the estimation of the quadratic function. In `STATA` this would look like:

```
reg testscr c.avginc##c.avginc, r
```

```
Linear regression                               Number of obs     =        420
                                                F(2, 417)         =     428.52
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5562
                                                Root MSE          =     12.724

------------------------------------------------------------------------------
                 |               Robust
         testscr | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-----------------+------------------------------------------------------------
```

```
           avginc |   3.850995   .2680941    14.36   0.000      3.32401    4.377979
                  |
c.avginc#c.avginc |  -.0423085   .0047803    -8.85   0.000     -.051705   -.0329119
                  |
            _cons |   607.3017   2.901754   209.29   0.000     601.5978    613.0056
------------------------------------------------------------------------------------
```

Now, it is straightforward to test the null-hypothesis of linearity against the alternative that the regression function is a quadratic. Namely, we only have to consider the t-statistic of the quadratic term. And that is larger than 1.96, so against a 5% significance level we reject the null-hypothesis of linearity.

Note by the way the syntax `c.avginc##c.avginc` which seems a bit strange. However, this particular line of code is very useful for later tabulation, plotting and other manipulations of the output. In this way `STATA` knows that there should be a quadratic effect of the same variable (`avginc`). The syntax `c.` denotes that the variable should be considered as continuous instead of as an integer (try it and behold the horrible output). There are four useful operators that you want to know when working with polynomials and interaction effect:

- `i.` operator: this specifies that the following variable is an integers and should be considered on all its level. This actually create indicator or dummies variables
- `c.` operator: this specifies that the following variable is a continuous variables and should be treated as continuous.
- `#` binary operator that specifies an interaction between two variables
- `##` binary operator that specifies both interaction between two variables and the individual variable effect

Plotting, non-linear population regression lines are a bit tricky. Namely, you want to combine a pylonomial with a linear dimension. One way of doing this is as follows:

```
predict hat1
scatter (testscr avginc) || (line hat1 avginc, sort)
```

where after the regression we **predict** the test scores (and name it something like `hat1`) and then we ask for a line of the prediction for each value of average district income. Note, though, that we have to `sort` the prediction from small to large to get a smooth line. And this provides the nice curved population regression line in the following `STATA` output.

But what is now the marginal effect of average district income. That, now, depends on itself. Namely, $\frac{\partial \text{testscore}}{\partial \text{income}} = \beta_1 + \beta_2 \text{income}$. Another way of seeing this is to compute the effects for different values of X

$$\widehat{TestScore}_i = 607.3 + 3.85 Income_i - 0.0423 (Income_i)^2 \tag{6.23}$$

The predicted change in test scores for a change in income from \$5,000 per capita

Figure 6.6: A non-linear relation

Table 6.1: Effect of X

Change in Income (1000 dollar per capita)	$\Delta \widehat{TestScore}$
from 5 to 6	3.4
from 25 to 26	1.7
from 45 to 46	0.0

to \$6,000 per capita then amounts to:

$$\begin{aligned} \Delta \widehat{TestScore} &= 607.3 + 3.85 \times 6 - 0.0423 \times 6^2 && (6.24) \\ &\quad -(607.3 + 3.85 \times 5 - 0.0423 \times 5^2) && (6.25) \\ &= 3.4 && (6.26) \end{aligned}$$

And if calculate the predicted effects for different values of X, then we get the following table:

Thus, the effect of a change in income is greater at low than high income levels (perhaps, a declining marginal benefit of an increase in school budgets?). But, be careful here! What is the effect of a change from 65 to 66? That is quite negative and already Figure 6.6 shows that a quadratic specification start to decline again the value of about 50; and perhaps that is not the behaviour that you want. So, with polynomials it is essential not to extrapolate outside the range of the data (and still interpret the outcome).

The estimation of a cubic specification is straightforward:

```
reg testscr c.avginc##c.avginc##c.avginc, r
```

```
Linear regression                               Number of obs     =        420
                                                F(3, 416)         =     270.18
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5584
                                                Root MSE          =     12.707

------------------------------------------------------------------------------------------
                           |               Robust
                   testscr | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
---------------------------+--------------------------------------------------------------
                    avginc |   5.018677   .7073504     7.10   0.000      3.62825    6.409103
                           |
         c.avginc#c.avginc |  -.0958052   .0289537    -3.31   0.001     -.152719   -.0388913
                           |
c.avginc#c.avginc#c.avginc |   .0006855   .0003471     1.98   0.049     3.26e-06    .0013677
                           |
                     _cons |    600.079   5.102062   117.61   0.000     590.0499     610.108
```

Where if we now want to test the null- hypothesis of linearity, then we have to have invoke an F-test. Namely, the alternative hypothesis is that the population regression is quadratic and/or cubic, that is, it is a polynomial of degree up to 3, so:

- H_0: Coefficients on $Income^2$ and $Income^3 = 0$
- H_1: at least one of these coefficients is nonzero.

And the outcome below shows that the null-hypothesis that the population regression is linear is rejected at the 5% (and 1%) significance level against the alternative that it is a polynomial of degree up to 3.

```
test avginc#avginc avginc#avginc#avginc
```

```
 ( 1)  c.avginc#c.avginc = 0
 ( 2)  c.avginc#c.avginc#c.avginc = 0

       F(  2,   416) =   37.69
            Prob > F =    0.0000
```

6.3.2 Interaction variables

Using interaction variables is a special case of polynomial effects. Namely, instead of multiply a variable with itself $X \times X = X^2$, you now multiple a variable with another variable. And you want to do this to take into account interactions between independent variables. Assume, for example, that a class size reduction is more effective in some circumstances than in others (which is quite conceivable). Perhaps smaller classes help more if there are many English learners (i.e., large migrant communities), who need more individual attention. That is, $\frac{\partial TestScore}{\partial STR}$ might depend on $PctEL$. More generally, this subsection looks into the fact that the marginal effect of $\frac{\partial Y}{\partial X_1}$ might depend on some other variable X_2.

6.3.2.1 Interactions between two binary variables

First, we look into the simplest (and perhaps most insightful) case of two binary (dummy variables). Consider therefore the following linear regression model:

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + u_i, \tag{6.27}$$

where both D_{1i} and $ D_{2i}$ are now considered to be binary. Now, of course, β_1 is the effect of changing $D_1 = 0$ to $D_1 = 1$. So, in this specification, this effect doesn't depend on the value of D_2. To allow the effect of changing D_1 to depend on D_2, we have to include the interaction term $D_{1i} \times D_{2i}$ as a regressor:

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 (D_{1i} \times D_{2i}) + u_i \tag{6.28}$$

Table 6.2: Interpretation of interaction effects with dummies

	$HiEL = 0$	$HiEL = 1$
$HiSTR = 0$	664.1	$664.1 - 18.2 = 645.9$
$HiSTR = 1$	$664.1 - 1.9 = 662.2$	$664.1 - 1.9 - 18.2 - 3.5 = 640.5$

To interpret now the coefficient β_1 we comparing the two cases for $D_1 = 0$ to $D_1 = 1$"

$$E(Y_i|D_{1i} = 0, D_{2i} = d_2) = \beta_0 + \beta_2 d_2 \quad (6.29)$$

$$E(Y_i|D_{1i} = 1, D_{2i} = d_2) = \beta_0 + \beta_1 + \beta_2 d_2 + \beta_3 d_2 \quad (6.30)$$

If we now subtract them from each other:

$$E(Y_i|D_{1i} = 1, D_{2i} = d2) - E(Y_i|D_{1i} = 0, D_{2i} = d_2) = \beta_1 + \beta_3 d_2 \quad (6.31)$$

then we have the marginal effect of D_1 which now depend d_2. The interpretation of β_3 boils down to being incremental to the effect of D_1, when $D_2 = 1$

Let us go back to our Californian school district example with the following variables to be used: test scores, student teacher ratio, and English learners. Let:

$$HiSTR = 1 \text{ if } STR \geq 20 \text{ and } HiEL = 1 \text{ if } PctEL \geq 10 \quad (6.32)$$

$$HiSTR = 0 \text{ if } STR < 20 \text{ and } HiEL = 0 \text{ if } PctEL < 10 \quad (6.33)$$

$$(6.34)$$

And if we have the estimation results we get the following outcome.

$$\widehat{TestScore} = 664.1 - 18.2HiEL - 1.9HiSTR - 3.5(HiSTR \times HiEL) \quad (6.35)$$

So, how to interpret the various parameters? Perhaps the simple way is to construct the following two-by-two table:

Now, Table 6.2 specifies for each combination (and there are exactly four of them) of $HiSTR$ and $HiEL$ the average expected test score outcome. Clearly, there are different 'marginal' effect of $HiSTR$. Namely, the effect of $HiSTR$ when $HiEL = 0$ is -1.9, whilst the effect of $HiSTR$ when $HiEL = 1$ is $-1.9 - 3.5 = -5.4$. This clearly points out that a class size reduction is estimated to have a bigger effect when the percent of English learners is large. However, when you estimate in `STATA` then you see that this interaction is not statistically significant, because the t-statistic equals $3.5/3.1 = 1.1$

6.3.2.2 Interactions between continuous and binary variables

The second case we consider is between a continuous and a binary variable. First assume the following regression model:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + +u_i, \quad (6.36)$$

where D_i is a binary variable and X is a continuous variable. As specified above, the effect on Y of X (holding D constant) $= \beta_1$, which does not depend on D. To allow the effect of X to depend on D, we can include the interaction term $D_i \times X_i$ as a regressor:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 (D_i \times X_i) + u_i \tag{6.37}$$

What this binary-continuous interaction does is essential create two different population regression lines. Namely, for observations with $D_i = 0$ (the $D = 0$ group or the $D = 0$ regression line) there is:

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \tag{6.38}$$

Whilst for observations with $D_i = 1$ (the $D = 1$ group or the $D = 1$ regression line) the regression line comes down to:

$$\begin{aligned} Y_i &= \beta_0 + \beta_2 + \beta_1 X_i + \beta_3 X_i + u_i & (6.39) \\ &= (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_i + u_i & (6.40) \end{aligned}$$

And these two population regression lines might both differ in the level (the constant) and in the slope of the line. So, there are three possibilities as depicted in Figure 6.7

In the first panel (a), $\beta_3 = 0$, so there is only a level effect. In the second panel (b), both β_2 and β_3 are not 0, so there is both a level and a slope effect. The last panel indicates that $\beta_2 = 0$, meaning that there is only a slope effect. But how to interpreting the coefficients now? Therefore, we take the marginal effect of

$$Y = \beta_0 + \beta_1 X + \beta_2 D + \beta_3 (D \times X) \tag{6.41}$$

which yields:

$$\frac{\partial Y}{\partial X} = \beta_1 + \beta_3 D \tag{6.42}$$

Thus, the effect of X depends on D and β_3 is the increment to the effect of X, when $D = 1$ (a slope effect)

To see this in our Californian school district example we now use the variables test scores, student teacher ratio and the as previously defined dummy variable $HiEL$ as:

$$\widehat{TestScore} = 682.2 - 0.97STR + 5.6HiEL - 1.28(STR \times HiEL) \tag{6.43}$$

Now when $HiEL = 0$ the population regression line amounts to:

$$\widehat{TestScore} = 682.2 - 0.97STR \tag{6.44}$$

And when $HiEL = 1$ the population regression line is:

$$\begin{aligned} \widehat{TestScore} &= 682.2 - 0.97STR + 5.6 - 1.28STR & (6.45) \\ &= 687.8 - 2.25STR & (6.46) \end{aligned}$$

Figure 6.7: Three possible binary-continuous interaction outcomes

Thus we have two regression lines: one for each $HiSTR$ group. And the conclusion is that a class size reduction is estimated to have a larger effect when the percent of English learners (migrant communities) is large.

Hypothesis testing is a before. To test whether the two regression lines have the same slope, the null-hypothesis boils down to the coefficient of $STR \times HiEL$ being zero: the t-statistic of this one become $-1.28/0.97 = -1.32$ and thus we do not reject this test. To test whether the two regression lines have the same intercept, the null-hypothesis becomes the coefficient of $HiEL$ being zero, yielding: $t = -5.6/19.5 = 0.29$, so we do not reject that null-hypothesis either. Interestingly, the null-hypothesis that the two regression lines are the same—population coefficient on $HiEL = 0$ and population coefficient on yields $STR \times HiEL = 0$: $F = 89.94(p - value < .001)$. So, we reject the joint hypothesis but neither individual hypothesis.

6.3.2.3 Interactions between two continuous variables

The last case are interaction between two continuous variables and that is always a difficult case of interpret. Starting again with the model:

$$Y_i = \beta_0 + \beta 1 X_{1i} + \beta_2 X_{2i} + u_i, \tag{6.47}$$

where both X_1, X_2 are continuous and as specified, the effect of X_1 doesn't depend on X_2 and the effect of X_2 doesn't depend on X_1. Now, to allow the effect of X_1 to depend on X_2, we include the interaction term $X_{1i} \times X_{2i}$ as a regressor. Where, to interpret the coefficients, we take the first derivative of X_1 in:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 (X_{1i} \times X_{2i}) + u_i \tag{6.48}$$

which yields:

$$\frac{\partial Y}{\partial X} = \beta_1 + \beta_3 X_2 \tag{6.49}$$

where β_3 should be interpreted as the increment to the effect of X_1 from a unit change in X_2.

6.3.3 Logarithmic transformation

To incorporate non-linear effect, very often logarithmic transformations are used of Y and/or X, where typically we use $\ln(X)$ as the natural logarithm of X. One feature of logarithmic transformations is that they permit modeling relations in percentage terms (like elasticities), rather than linearly. That is because:

$$\ln(x + \Delta x) - \ln(x) = \ln(1 + \frac{\Delta x}{x}) \cong \frac{\Delta x}{x} \tag{6.50}$$

Note that this is an approximation, but from calculus we know that $\frac{d \ln(x)}{dx} = \frac{1}{x}$). And the above approximation works quite well for small numbers. For example, numerically: $\ln(1.01) = .00995 \cong .01$ and $\ln(1.10) = .0953 \cong .10$, where the latter

Table 6.3: Three logarithmic transformation

Case	Population regression model
linear-log	$Y_i = \beta_0 + \beta_1 \ln(X_i) + u_i$
log-linear	$\ln(Y_i) = \beta_0 + \beta_1(X_i) + u_i$
log-log	$\ln(Y_i) = \beta_0 + \beta_1 \ln(X_i) + u_i$

is still rather close. Now remember the following rules for natural logarithms 1. $\ln(a \times b) = \ln(a) + \ln(b)$ 2. $\ln(\frac{a}{b}) = \ln(a) - \ln(b)$ 3. $\ln(a^\alpha) = \alpha \ln(a)$ 4. $\ln(e^X) = X$.

When you encounter a nonlinear model such as the ones adopted in Chapter 2 a strategy that often works is log-linearization. That works as follows

$$Y = AK^\alpha L^{1-\alpha} \rightarrow \ln(Y) = \ln(A) + \alpha \ln(K) + (1-\alpha)\ln(L), \tag{6.51}$$

where you take the natural logarithm on both sides. There are three different cases of logarithmic regression models as specified in Table **??**.

Though statistical testing remains the same, the interpretation of the slope coefficient differs in each case. To derive the interpretationwe want to find the marginal effect of X using the first derivative.

6.3.3.1 Linear-log population regression model

The linear-log population regression model is specified as:

$$Y = \beta_0 + \beta_1 \ln(X) \tag{6.52}$$

Now take the first derivative:

$$\frac{\partial Y}{\partial X} = \frac{\beta_1}{X} \tag{6.53}$$

so

$$\beta_1 = \frac{\partial Y}{\partial X / X} \tag{6.54}$$

In this case that means that β_1 should be interpreted as the absolute change of Y when X changes with $\beta_1/100$ percent. To illustrate this, consider the case where we take natural logarithm od district income, so we define the new regressor as, $\ln(Income)$

The model is now linear in $\ln(Income)$, so the linear-log model can be estimated by OLS, which yields

$$\widehat{TestScore} = 557.8 + 36.42 \times \ln(Income_i) \tag{6.55}$$

so an 1% increase in $Income$ is associated with an increase in test scores of 0.36 points on the test. And again, standard errors, confidence intervals, R^2—all

the usual tools of regression apply here. But the difficulty in plottin the new regression line remains. Consider the following STATA syntax, where we first have to define the new regressor by invoking the generate command.

```
gen lninc = ln(avginc)
reg testscr lninc, r
predict testhat
graph twoway (line testhat avginc, sort) (scatter testscr avginc)
```

This now provides the following STATA output.

When you compare 6.8 with 6.6 then you notice that in the case of logarithm the population remains increasing (but less and less steep). This can be considered as an advantage when you want to estimate decreasing (or increasing) return.

6.3.3.2 Log-linear population regression model

The second case we consider is the log-linear population regression model, as specified by:

$$\ln(Y) = \beta_0 + \beta_1 X \tag{6.56}$$

To find the interpretation of $\beta 1_1$, we again take the first derivative $\frac{\partial Y}{\partial X}$, but first transform the model like this:

$$Y = exp(\beta_0 + \beta_1 X) \tag{6.57}$$

then take the first derivative:

$$\frac{\partial Y}{\partial X} = \beta_1 exp(\beta_0 + \beta_1 X) = \beta_1 Y \tag{6.58}$$

and collec terms

$$\beta_1 = \frac{\partial Y/Y}{\partial X} \tag{6.59}$$

The interpretation of β_1 now is that one unit change in X causes a β_1 percentage in Y

6.3.3.3 Log-log population regression model

Finally, we have our third case, being the log-log population regression model as specified by:

$$\ln(Y) = \beta_0 + \beta_1 \ln(X) \tag{6.60}$$

To find the interpretation of $\beta 1_1$, we again take the first derivative $\frac{\partial Y}{\partial X}$, but first transform the model like this:

$$Y = exp(\beta_0 + \beta_1 \ln(X)) \tag{6.61}$$

So

$$\frac{\partial Y}{\partial X} = \beta_1 / X exp(\beta_0 + \beta_1 \ln(X)) = \beta_1 Y / X \tag{6.62}$$

Figure 6.8: A non-linear relation

and after collecting terms we end up with an **elasticity**:

$$\beta_1 = \frac{\partial Y/Y}{\partial X/X} \tag{6.63}$$

As an example consider the case when we want to regress ln(test scores) on ln(income). To do so, we first define a new dependent variable, ln(TestScore), and a new regressor, ln(Income) The model is now a linear regression of ln(TestScore) against ln(Income), which can be estimated by OLS as follows

$$ln(\widehat{TestScore}) = 6.336 + 0.0554 \times ln(Income_i), \tag{6.64}$$

where the interpretation is that an 1% increase in $Income$ is associated with an increase of .0554% in $TestScore$ ($Income$ up by a factor of 1.01, $TestScore$ up by a factor of 1.000554)

Suppose that we now want to plot both the log-linear and the log-log specification, then we can use the following syntax:

```
gen lninc = ln(avginc)
gen lntestscr = ln(testscr)
reg lntestscr lninc, r
predict testhat1
reg lntestscr avginc, r
predict testhat2
graph twoway (line testhat1 avginc, sort) (line testhat2 avginc, sort) (scatter lntests
```

which provides the following STATA output.

Note that the y-axis is on a logarithmic scale here, thus the log-linear specification is now a linear line.

6.3.3.4 Summary: logarithmic transformations

We have seen three different cases of logarithmic specification, differing in whether Y and/or X is transformed by taking logarithms. Now, the regression is linear in the new variable(s) $\ln(Y)$ and/or $\ln(X)$, and the coefficients can be estimated by OLS where hypothesis tests and confidence intervals are now implemented and interpreted 'as usual'. Only the interpretation of the coefficients differs from case to case and is directly related to percentage changes (growth) and elasticities. Oftentimes, the choice of specification, however, should be guided by judgment (which interpretation makes the most sense in your application?), tests, and plotting predicted values. Sometimes, though, you have a structural economic model such as Equation (2.7), which defines the type of specification you should use. Finally, see that in economics many models exists with decreasing or increasing return to scale and that these are very closely related with logarithmic specifications.

Figure 6.9: A non-linear relation

6.4 Using fixed effects in panel data

Multivariate regression is a powerfull tool for controlling for the effect of variables for which we have data. But often we do not have data on what we suspect might be important—data, such as individual characteristics like ambition, intelligence, drive or stamina. Or regional of country data, where the type of soil, the ruggedness (hilliness), or population density determine to a large extent the behaviour of people living on it. If we do not have this type of data, then it not always the case that everything is lost. Especially, when we have repeated observations, so observations of the same entity throughout time. This is referred to as panel data and requires one additional subscript t as in X_{it} indicating the observation X on individual made at time t. To understand why this sometimes works, we temporarily change to another dataset and that is the 'fatality' data collected by Levitt and Porter (2001) and deals with the relation between drunk driving and fatal accidents in the States of the US between 1982 and 1988. For this particular example we look at the impact of the 'beer tax', measured as the real tax in dollars on a case of beer, on 'fatality', measured as the number of annual traffic deaths per 10,000 people in the population of each stata. For this we first read the data and manipulate the mortality variable

```
use "./data/fatality.dta", clear
gen fatality = allmort/pop * 10000
```

and then run a simple regression:

```
regress fatality beertax, robust
```

```
Linear regression                               Number of obs     =        336
                                                F(1, 334)         =      47.59
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0934
                                                Root MSE          =     .54374

------------------------------------------------------------------------------
             |               Robust
    fatality | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
     beertax |   .3646054   .0528524     6.90   0.000     .2606399     .468571
       _cons |   1.853308   .0471297    39.32   0.000     1.760599    1.946016
------------------------------------------------------------------------------
```

But these outcomes are very strange. For every dollar increase in tax, number of fatal accidents per 10,000 people increases with 0.36, which is statistically significantly different from 0. What is going on here. Most likely this effect is biased because of omitted variable bias. States in the US differ widely in terms of population density, environment, institutions, religion, poverty, and so on and so forth. And Those state characteristics might influence both the variables beertax and fatality.

Fortunately, for each state we have yearly data. So, that is 7 observations per stata and we can make use of that by using fixed effects, which is a very common technique in the social sciences—especially in economics. We model the use of fixed effects in this example as follows:

$$\text{fatality}_{it} = \beta_0 + \beta_1 \text{beertax}_{it} + \beta_3 S_1 + \ldots + \beta_5 1 S_{48} + u_{it}, \tag{6.65}$$

where S_i denote indicator (dummies) for each state which constitute the fixed effects. In total there are 48 states in this dataset, so we have 48 dummies. Note that these fixed effects only depend on the state variation, not on time variation. So, essentially what these fixed effects capture is all state specific characteristics which are constant over time. And most of the characteristics' examples given above do not vary that much over time, so by using these state fixed effects we can **control** for them. In `STATA` you can estimate this in a straightforward way as `regress fatality beertax i.state, robust`, but this lots of statistical output that you are usually not interested in. Almost just as easy would be is to invoke the `areg` command, where you specifically state that the state variable should be used as dummies but not shown using `absorb(state)`: and then run a simple regression:

```
areg fatality beertax, absorb(state) robust
```

```
Linear regression, absorbing indicators                 Number of obs     =    336
Absorbed variable: state                                No. of categories =     48
                                                        F(1, 287)         =  10.41
                                                        Prob > F          = 0.0014
                                                        R-squared         = 0.9050
                                                        Adj R-squared     = 0.8891
                                                        Root MSE          = 0.1899

------------------------------------------------------------------------------
             |               Robust
    fatality | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
     beertax |  -.6558737   .2032797    -3.23   0.001    -1.055982   -.2557655
       _cons |   2.377075   .1051516    22.61   0.000     2.170109    2.584041
------------------------------------------------------------------------------
```

Now, see what happens with the coefficient of the beer tax variable. It changes sign! So from positive it becomes negative. That is how **disruptive** omitted variable bias can be. Also see that by including all these state fixed effects, the $\bar{R}^2$ now increase enormously to 91%, which does make sense because the states explain the variation in fatality rate to a large extent (e.g., compare Kansas with Connecticut).

This is just a snapshot of the use of fixed effects in panel data, but for now this is enough. But for now, know that the use of fixed effects can go a long way in addressing omitted variable bias.

6.5 Conclusion and discussion

Chapter 7

Specification and Assessment Issues

This concluding chapter deals with an epistemological question, namely how to convey your key results, and an ontological question, how do you know whether what you convey is true (e.g., unbiased). To do so, we first look at specification issues in Section 7.1. Which variables should you include? Thereafter, Section 7.2 discusses how to present your results. And, subsequently, Section 7.3 gives a bit into the question into the validity of your results. When do you know that you really obtained estimated an **unbiased** parameter. The final section concludes.

7.1 Specification of your model

A long-standing but simple question is how to decide which variables to include in a regression model. Unfortunately, the answer to this question is rather complex. A straightforward but naive approach would be to include them all! So, throw every variable in that is in your database. This, however, leads to "causal salad" (McElreath 2020) as displayed in Figure 7.1 and can actually lead to a biased estimator. One reason for this is that if you include a variable that is related to the error term then all other parameters are biased as well.

So, for the final time we return to our Californian school district data and now try to devise a specification that *mimimizes* the changes upon biased estimator. So, our focus is to get an unbiased estimate of the effect on test scores of changing class size, holding constant student and school characteristics but not necessarily holding constant the budget (we do not want to control for budget as this actually governs class sizes).

To do this we need to think about what variables to include and what regressions

Figure 7.1: Causal salad

to run—and we should do this before we actually sit down at the computer. Think beforehand about your model specification and try to avoid throwing everything in (your causal salad).

In practice, and especially economics, most follow the following general approach to variable selection and model specification: 1. First you specify a base or benchmark model. In this case that is the univariate regression of test scores on class size. 2. Then you specify a range of plausible alternative models, which include additional candidate variables. 3. Then you assess whether a candidate variable changes the coefficient of interest ($\hat{\beta}_1$)? You keep focusing on the effect of class size! 4. You assess whether a candidate variable is statistically significantly different from zero; so whether it has an impact of not. - Use judgment, not a mechanical recipe, meaning that variable statistically insignificant different from zero should not automatically be thrown out. - In all cases, do not just try to maximize $\bar{R}^2$. You focus on a causal effect, not on prediction

Considering the last point, tt is easy to fall into the trap of maximizing the $\bar{R}^2$—but this loses sight of our real objective, an unbiased estimator of the class size effect. Recall that a high $\bar{R}^2$ means that the regressors explain the variation in Y. It does **not** mean

- that you have eliminated omitted variable bias;
- that you have an unbiased estimator of a causal effect (β_1);
- that the included variables are statistically significant.

So, in this case, what variables would you want—ideally—to include to estimate the effect on test scores of STR using school district data?

There is a whole set of potential relevant variables in the California class size

data set, being:

- student-teacher ratio (STR)—the variable we focus on
- percent English learners in the district ($PctEL$)—as a proxy for large migrant communities
- school expenditures per pupil—largely correlated with student-teacher ratio
- name of the district (so we could look up average rainfall, for example)
- percent eligible for subsidized/free lunch—proxies district income
- percent on public income assistance—proxies district income
- average district income—a measure for district affuency

So, which of these variables would you want to include?

Looking at Figure 7.2, all three percentage variables (English learners, subsidized lunch, and income assistance) behave in a similar manner. But interestingly, the strongest relation is between subsidized lunch and test scores and that is at least the variable that we would like to include.

7.2 Presentation of results

So, we have a number of regressions (also called specifications) and we want to report them. Often, it is awkward and difficult to read regressions written out in equation form, so instead it is conventional to report them in a table. Note that reading regression estimates from computer output is even more difficult. On top of that it is ugly and contains way too much information. Try to avoid statistical computer output as much as possible—at least in your thesis. Now, regression tables should include a couple of elements:

- The estimated regression coefficients.
- The standard errors or the t-statistics. Having both of them is too much. Do not report p-value, because often they are not informaties (being $p = 0.000$).
- Some measures of fit (usually just the $\bar{R}^2$ would do).
- The number of observations.
- Some relevant F-statistics, if any. Usually they are not included.
- Any other pertinent information but typically there is none.

You can find most of this information in the final estimation Table 7.3 as presented in Stock, Watson, et al. (2003).

So, here the variable of interest (student-teacher ratio) is the first variable on top. And the table keeps focusing on that one. Moreover, specification (3) and (5) seems to be preferred as they have the highest $\bar{R}^2$, although that is perhaps of lesser importance. What we can infer from this is that the estimate for student-teacher ratio remains robust around -1 and is significantly different from 0. Does this now mean that this effect is **unbiased**? Most likely not, but that is something that the next section will discuss.

Figure 7.2: Test scores versus various independent variables

TABLE 7.1 **Results of Regressions of Test Scores on the Student–Teacher Ratio an Characteristic Control Variables Using California Elementary School D**

Dependent variable: average test score in the district.

Regressor	**(1)**	**(2)**	**(3)**	
Student–teacher ratio (X_1)	−2.28** (0.52)	−1.10* (0.43)	−1.00** (0.27)	
Percent English learners (X_2)		−0.650** (0.031)	−0.122** (0.033)	
Percent eligible for subsidized lunch (X_3)			−0.547** (0.024)	
Percent on public income assistance (X_4)				
Intercept	698.9** (10.4)	686.0** (8.7)	700.2** (5.6)	6
Summary Statistics				
SER	18.58	14.46	9.08	
$\overline{R}^2$	0.049	0.424	0.773	
n	420	420	420	

These regressions were estimated using the data on K-8 school districts in California, described in Appe are given in parentheses under coefficients. The individual coefficient is statistically significant at the *5 cance level using a two-sided test.

Figure 7.3: Various specifications of test score models

7.3 Potential sources of bias

No to include we would like to answer the question whether there is a systematic way to assess regression studies? We already have seen that multivariate regression models have some key virtues:

1. They provide an estimate of the marginal effect of the variable of interest X on Y.
2. They resolve the problem of omitted variable bias, if an omitted variable can be measured and included.
3. They can handle nonlinear relations (effects that vary with the X's) and therefore resolve the problem of misspecification bias.

Still, OLS might yield a **biased** estimator of the true causal effect. In other words, it might not yield valid inferences. That what you want to measure is not what you actually measure. In general there is two ways to assess statistical studies: threats to internal and threats to external validity.

- **Internal validity**: the statistical inferences about causal effects are valid for the population being studied.
- **External validity**: the statistical inferences can be generalized from the population and setting studied to other populations and setting.

7.3.1 Threats to external validity

So, above we came to a (tentative) conclsusion about the impact of class size on test scores. But we have done so in the context of Californian school districts in the year 2005. Can we extend this finding and generalize class size results from California school districts to other population, for example to that of Massachusetts or Mexico in 2005? And can we do so for differences in institutional settings as there are different legal requirements concerning special education, different treatment of bilingual education, and differences in teacher characteristics across regions and countries.

We therefore always to be careful to transfer our finding to that of other settings. Note that this as well a special case of omitted variable bias but now outside the scope of our study (our population).

7.3.2 Threats to internal validity

In applied econometrics, the following five threats to the internal validity of regression studies are usually given (in statistics there is a different framework for this, but in most cases they come down to the same thing)

1. Omitted variable bias
2. Wrong functional form
3. Errors-in-variables bias or measurement error
4. Sample selection bias
5. Simultaneous causality bias

All of these imply that $E(u_i|X_{1i}, \ldots, X_{ki}) \neq 0$, in which case the OLS estimates are therefore **biased**.

7.3.2.1 Omitted variable bias

Omitted variable bias arises if an omitted variable is both a determinant of Y and a determinant of at least one included regressor. We first discussed omitted variable bias in regression with a single X, but Omitted variable bias will arise when there are multiple X's as well, if the omitted variable satisfies the two conditions above. Fortunately, there are potential solutions to omitted variable bias

- If the variable can be measured, include it as an additional regressor in multiple regression;
- Possibly, use panel data in which each entity (individual) is observed more than once;
- If the variable cannot be measured, use instrumental variables regression (for later courses);
- Run a randomized controlled experiment.

7.3.2.2 Wrong functional form

This threat to internal validity arises if the functional form is incorrect. For example, if an interaction term is incorrectly omitted, then inferences on causal effects will be biased. There is a potential solution to functional form misspecification and that is to use the appropriate nonlinear specifications in X (logarithms, interactions, etc.). Sometimes this is not possible and then one has to resort to direct non-linear estimation techniques.

7.3.2.3 Errors-in-variables bias or measurement error

The third threat is measurement error or sometimes know as Errors-in-variables bias. So far we have assumed that X is measured without error. In reality, (economic) data is often measured with error have measurement error. Especially surveys are prone to measurement error. For example recollection errors that arise with questions as "which month did you start your current job?". Or ambiguous questions problems as "what was your income last year?" What is meant with latter: monthly or yearly income, gross or net income? Also respondents sometimes have an incentive not to answer honestly (intentionally false response problems) with questions as "What is the current value of your financial assets?" or "How often do you drink and drive?". There are potential solutions to errors-in-variables bias, such as

- Obtain better data, but that is bit easy
- Develop a specific model of the measurement error process. This is only possible if a lot is known about the nature of the measurement error—for example a subsample of the data are cross-checked using administrative records and the discrepancies are analyzed and modeled.

- Instrumental variables regression.

7.3.2.4 Sample selection bias

So far we have assumed simple random sampling of the population. In some cases, simple random sampling is thwarted because the sample, in effect, **selects itself**. Now, then sample selection bias arises when a selection process both influences the availability of data and if that process is related to the dependent variable. To illustrate this, I will adopt a hypothetical example given by McElreath (2020). Here we want to look at the relation between trustworthy science and newsworthy science. This example is motivated by the fact that newsworthy science (clickbait in the social media) oftentimes turns out not to be true. To given a reason why this might, we first simulate an artificial database of 400 observations of both newsworthy and trustworthy. Both variable are constructed such that they are $i.i.d.$ and standard normally distributed. So, there is no relation whatsoever and, indeed, Figure 7.4 shows a rather random cloud plot.

But what if editors on social media have a decision rule: scientific output should be either thrustworthy or newsworthy, and preferably both. So, as a rule of thumb the select only the top 10% scientific outcomes , so the ones that score in the top 10% when both scores are added up (trustworthy + newsworthy). If we now depict the selected ones in grey in Figure 7.5, then clearly suddenly a negative relation emerges between newsworthiness and thrustworthiness. And that negative relation is caused by the selection (external) editors make. So, if there is a selection somewhere in the process, estimates of what you want to estimates can quickly become biased.

This process occurs more often than you might think. Consider the two following examples:

1. **Aircraft noise externality**. Here the question is to what extent people "value" aircraft noise (that is in a negative sense)? To aim for an answer we adopt the following empirical strategy; we collect housing prices close to Schiphol airport (say Zwanenburg) and compare them with identical houses further away (say Schagen). We have data for individual housing prices (including characteristics) since 1985. As an estimator we assess then the average mean difference between the Zwanenburg and Schagen location. Now the question is whether there is sample selection bias. And indeed there is and that is caused by the fact that humans react on their own situation based upon their preferences. In this case, they react by means of moving residence. So, those who have strong negative preference regarding aircraft noise are the first to move out Zwanenburg (if possible). So the population in both locations is not identical, instead they sorted spatially.

2. **Returns to education**. The question here is rather straightforward and involved the monetary returns to an additional year of education. As empirical strategy we collect data of all employed workers in the

Figure 7.4: Random observations of newsworthy and trustworthy

Figure 7.5: Negative relation amongst the selected points

Netherlands (actually this data exists and is called micro-data), including worker characteristics, years of education, and hourly wages. Our approach is here to regress $\ln(Earnings)$ on $YearsEducation$ and a large set of other characteristics. Now, ignore issues of omitted variable bias and measurement error, then the question is: is there sample selection bias? And, indeed, there is again, as you only sample those people who are employed and not the unemployed (they have no current wage). And this leads to a different population than you wanted in the first place.

In there there are some potential solutions to sample selection bias and most of them deal with data issues. For example, you might want to collect the sample in a way that avoids sample selection. For example you might want to focus on those people who moved between Schagen and Zwanenburg or you include the unemployed as well in the returns to education example.

7.3.2.5 Simultaneous causality bias in equations

Finally, our last threat to causality is simultaneous or reverse causality bias. This means that the causal effect might go either way as in the following system

- Causal effect on Y of X: $Y_i = \beta_0 + \beta_1 X_i + u_i$
- Causal effect on X of Y: $X_i = \gamma_0 + \gamma_1 Y_i + v_i$

Where a large u_i means a large Y_i, which implies large X_i (if $\gamma_1 > 0$) and therefore, by definition, $corr(X_i, u_i) \neq 0$. Thus, $\hat{\beta}_1$ is biased and inconsistent. In our Californian school district example it might as well be that a district with particularly bad test scores given the STR (negative u_i) receives extra resources, thereby lowering its STR; so STR_i and u_i are then correlated

There are some potential solutions to simultaneous causality bias

The first and always the best one is to conduct a randomized controlled experiment. Because,if X_i is chosen at random by the experimenter, there is no feedback from the outcome variable to Y_i (assuming perfect compliance). Secondly, you can develop and estimate a complete model of both directions of causality. This is the idea behind many large macro models (e.g. those of the Federal Reserve Bank in the US). This is difficult in practice. Finally, you can use instrumental variables regression again to estimate the causal effect of interest (effect of X on Y, ignoring effect of Y on X). But that is not for this course.

7.4 Concluding remarks

Chapter 8

In conclusion

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky. 2019. "Regressive Sin Taxes, with an Application to the Optimal Soda Tax." *The Quarterly Journal of Economics* 134 (3): 1557–1626.

Amrhein, Valentin, Sander Greenland, and Blake McShane. 2019. "Scientists Rise up Against Statistical Significance." Nature Publishing Group.

Andreyeva, Tatiana, Michael W Long, and Kelly D Brownell. 2010. "The Impact of Food Prices on Consumption: A Systematic Review of Research on the Price Elasticity of Demand for Food." *American Journal of Public Health* 100 (2): 216–22.

Atkinson, Anthony B. 2009. "Economics as a Moral Science." *Economica* 76: 791–804.

Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *The Quarterly Journal of Economics* 126 (2): 805–55.

Bernheim, B Douglas, and Dmitry Taubinsky. 2018. "Behavioral Public Economics." *Handbook of Behavioral Economics: Applications and Foundations 1* 1: 381–516.

Brons, Martijn, Eric Pels, Peter Nijkamp, and Piet Rietveld. 2002. "Price Elasticities of Demand for Passenger Air Travel: A Meta-Analysis." *Journal of Air Transport Management* 8 (3): 165–75.

Busse, Meghan R, Devin G Pope, Jaren C Pope, and Jorge Silva-Risso. 2015. "The Psychological Effect of Weather on Car Purchases." *The Quarterly Journal of Economics* 130 (1): 371–414.

Cadario, Romain, and Pierre Chandon. 2020. "Which Healthy Eating Nudges Work Best? A Meta-Analysis of Field Experiments." *Marketing Science* 39 (3): 465–86.

Chetty, Raj. 2015. "Behavioral Economics and Public Policy: A Pragmatic Perspective." *American Economic Review* 105 (5): 1–33.

Chorus, Caspar G. 2015. "Models of Moral Decision Making: Literature Review and Research Agenda for Discrete Choice Analysis." *Journal of Choice Modelling* 16: 69–85.

Chorus, Caspar G, Baiba Pudāne, Niek Mouter, and Danny Campbell. 2018. "Taboo Trade-Off Aversion: A Discrete Choice Model and Empirical Analysis." *Journal of Choice Modelling* 27: 37–49.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2021. "The Production Function for Housing: Evidence from France." *Journal of Political Economy* 129 (10): 2766–816.

Epple, Dennis, Brett Gordon, and Holger Sieg. 2010. "A New Approach to Estimating the Production Function for Housing." *American Economic Review* 100 (3): 905–24.

Friedman, Milton. 1953. "The Methodology of Positive Economics."

Galton, Francis. 1886. "Regression Towards Mediocrity in Hereditary Stature." *The Journal of the Anthropological Institute of Great Britain and Ireland* 15: 246–63.

Godfray, H Charles J, Paul Aveyard, Tara Garnett, Jim W Hall, Timothy J Key, Jamie Lorimer, Ray T Pierrehumbert, Peter Scarborough, Marco Springmann, and Susan A Jebb. 2018. "Meat Consumption, Health, and the Environment." *Science* 361 (6399): eaam5324.

Green, Richard K, Stephen Malpezzi, and Stephen K Mayo. 2005. "Metropolitan-Specific Estimates of the Price Elasticity of Supply of Housing, and Their Sources." *American Economic Review* 95 (2): 334–39.

High, Jack. 1985. "Is Economics Independent of Ethics." *Reason Papers* 10 (1): 3–16.

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Vol. 25. Harvard university press.

Infante, Gerardo, Guilhem Lecouteux, and Robert Sugden. 2016. "Preference Purification and the Inner Rational Agent: A Critique of the Conventional Wisdom of Behavioural Welfare Economics." *Journal of Economic Methodology* 23 (1): 1–25.

Jachimowicz, Jon M, Shannon Duncan, Elke U Weber, and Eric J Johnson. 2019. "When and Why Defaults Influence Decisions: A Meta-Analysis of Default Effects." *Behavioural Public Policy* 3 (2): 159–86.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Macmillan.

Klamer, Arjo. 2017. *Doing the Right Thing: A Value Based Economy*. Ubiquity Press.

Laibson, David, and John A List. 2015. "Principles of (Behavioral) Economics." *American Economic Review* 105 (5): 385–90.

Layard, Richard, Guy Mayraz, and Stephen Nickell. 2008. "The Marginal Utility of Income." *Journal of Public Economics* 92 (8-9): 1846–57.

Lazear, Edward P. 2000. "Economic Imperialism." *The Quarterly Journal of Economics* 115 (1): 99–146.

Leonard, Thomas C. 2008. "Richard h. Thaler, Cass r. Sunstein, Nudge: Improving Decisions about Health, Wealth, and Happiness." Springer.

Levitt, Steven D, and Jack Porter. 2001. "How Dangerous Are Drinking Drivers?" *Journal of Political Economy* 109 (6): 1198–1237.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. Chapman; Hall/CRC.

Medema, Steven G. 2007. "The Hesitant Hand: Mill, Sidgwick, and the Evolution of the Theory of Market Failure." *History of Political Economy* 39 (3): 331–58.

Mouter, Niek, Paul Koster, and Thijs Dekker. 2021. "Contrasting the Recommendations of Participatory Value Evaluation and Cost-Benefit Analysis in the Context of Urban Mobility Investments." *Transportation Research Part A: Policy and Practice* 144: 54–73.

Pearl, Judea. 2009. *Causality.* Cambridge university press.

Pillai, Unni. 2015. "Drivers of Cost Reduction in Solar Photovoltaics." *Energy Economics* 50: 286–93.

Popper, Karl. 2005. *The Logic of Scientific Discovery.* Routledge.

Posner, Eric A, and Cass R Sunstein. 2017. "Moral Commitments in Cost-Benefit Analysis." *Va. L. Rev.* 103: 1809.

Reurink, Arjan. 2018. "Financial Fraud: A Literature Review." *Journal of Economic Surveys* 32 (5): 1292–1325.

Roth, Alvin E. 2018. "Marketplaces, Markets, and Market Design." *American Economic Review* 108 (7): 1609–58.

Sandel, Michael J. 2013. "Market Reasoning as Moral Reasoning: Why Economists Should Re-Engage with Political Philosophy." *Journal of Economic Perspectives* 27 (4): 121–40.

Sen, Amartya. 2000. "The Discipline of Cost-Benefit Analysis." *The Journal of Legal Studies* 29 (S2): 931–52.

Senn, Stephen. 2011. "Francis Galton and Regression to the Mean." *Significance* 8 (3): 124–26.

Simon, Herbert A. 1957. "Models of Man; Social and Rational."

Singh, Prabhpal et al. 2020. "Moral Realism and Expert Disagreement." *Trames* 24 (3): 441–57.

Spash, Clive L. 1997. "Ethics and Environmental Attitudes with Implications for Economic Valuation." *Journal of Environmental Management* 50 (4): 403–16.

Stanovich, Keith E, and Richard F West. 2000. "Individual Differences in Reasoning: Implications for the Rationality Debate?" *Behavioral and Brain Sciences* 23 (5): 645–65.

Stock, James H, Mark W Watson, et al. 2003. *Introduction to Econometrics.* Vol. 104. Addison Wesley Boston.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5 (4): 297–323.

Van Drunen, Michiel, Pieter Van Beukering, and Harry Aiking. 2010. "The True Price of Meat." In *Meat the Truth, Essays on Livestock Production, Sustainability and Climate Change*, 87–104. Nicolaas G. Pierson Foundation.

Van Staveren, Irene. 2007. "Beyond Utilitarianism and Deontology: Ethics in Economics." *Review of Political Economy* 19 (1): 21–35.

Varian, Hal R. 2003. *Intermediate Microeconomics: A Modern Approach.* Elsevier Brasil.

Verhoef, Erik, Peter Nijkamp, and Piet Rietveld. 1996. "Second-Best Congestion Pricing: The Case of an Untolled Alternative." *Journal of Urban Economics*

40 (3): 279–302.

Visschers, Vivianne HM, Nadine Wickli, and Michael Siegrist. 2016. "Sorting Out Food Waste Behaviour: A Survey on the Motivators and Barriers of Self-Reported Amounts of Food Waste in Households." *Journal of Environmental Psychology* 45: 66–78.

Wempe, Ben, and Jeff Frooman. 2018. "Reframing the Moral Limits of Markets Debate: Social Domains, Values, Allocation Methods." *Journal of Business Ethics* 153 (1): 1–15.

White, Mark D. 2009. "In Defense of Deontology and Kant: A Reply to van Staveren." *Review of Political Economy* 21 (2): 299–307.

Appendix A

Derivation of the optimal cost function

The Lagrangian function is written as:

$$H = rK + \omega L + \lambda(Q - TK^{\alpha_K}L^{\alpha_L}) \tag{A.1}$$

The three first-order conditions are given by:

$$\frac{\partial H}{\partial \lambda} = Q - TK^{\alpha_K}L^{\alpha_L} = 0 \tag{A.2}$$

$$\frac{\partial H}{\partial K} = r - \lambda\frac{\alpha_K}{K}TK^{\alpha_K}L^{\alpha_L} = 0 \tag{A.3}$$

$$\frac{\partial H}{\partial L} = \omega - \lambda\frac{\alpha_L}{L}TK^{\alpha_K}L^{\alpha_L} = 0 \tag{A.4}$$

Solving the second first-order condition for the Lagrangian multiplier gives:

$$\lambda = \frac{r}{\frac{\alpha_K}{K}TK^{\alpha_K}L^{\alpha_L}} = r\alpha_K^{-1}T^{-1}K^{1-\alpha_K}L^{-\alpha_L} \tag{A.5}$$

Substituting this intermediate solution in the third first-order condition gives:

$$\omega - \frac{r}{\frac{\alpha_K}{K}TK^{\alpha_K}L^{\alpha_L}}\frac{\alpha_L}{L}TK^{\alpha_K}L^{\alpha_L} \longleftrightarrow \omega = r\frac{\frac{\alpha_L}{L}}{\frac{\alpha_K}{K}} \tag{A.6}$$

Solving this equation for labour L gives:

$$L = \frac{r}{\omega}\frac{\alpha_L}{\frac{\alpha_K}{K}} = K\frac{r}{\omega}\frac{\alpha_L}{\alpha_K} \tag{A.7}$$

Substituting this intermediate solution in the first first-order condition gives:

$$Q - TK^{\alpha_K}\left(K\frac{r}{\omega}\frac{\alpha_L}{\alpha_K}\right)^{\alpha_L} = Q - TK^{\alpha_K+\alpha_L}\left(\frac{r}{\omega}\frac{\alpha_L}{\alpha_K}\right)^{\alpha_L} = 0 \tag{A.8}$$

Solving this equation for capital gives:

$$K^{\alpha_K+\alpha_L}\left(\frac{r}{\omega}\frac{\alpha_L}{\alpha_K}\right)^{\alpha_L} = \frac{Q}{T} \tag{A.9}$$

$$K^{\alpha_K+\alpha_L} = \frac{Q}{T}\left(\frac{r}{\omega}\frac{\alpha_L}{\alpha_K}\right)^{-\alpha_L} \tag{A.10}$$

Which leads to the following expression for optimal K^*:

$$K^* = \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{r}{\omega}\frac{\alpha_L}{\alpha_K}\right)^{\frac{-\alpha_L}{\alpha_K+\alpha_L}} \tag{A.11}$$

$$= \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{r}{\omega}\right)^{\frac{-\alpha_L}{\alpha_K+\alpha_L}}\left(\frac{\alpha_L}{\alpha_K}\right)^{\frac{-\alpha_L}{\alpha_K+\alpha_L}} \tag{A.12}$$

$$= \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{\omega}{r}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}}\left(\frac{\alpha_L}{\alpha_K}\right)^{\frac{-\alpha_L}{\alpha_K+\alpha_L}} \tag{A.13}$$

$$= \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{\omega}{r}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}}\left(\frac{\alpha_K}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} \tag{A.14}$$

Substituting this in the intermediate solution for labour gives:

$$L^* = K^*\frac{r}{\omega}\frac{\alpha_L}{\alpha_K} = \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{r}{\omega}\frac{\alpha_L}{\alpha_K}\right)^{\frac{-\alpha_L}{\alpha_K+\alpha_L}}\frac{r}{\omega}\frac{\alpha_L}{\alpha_K} = \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{r}{\omega}\frac{\alpha_L}{\alpha_K}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} \tag{A.15}$$

The last step is to simplify the exponent:

$$L^* = \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{r}{\omega}\frac{\alpha_L}{\alpha_K}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} \tag{A.16}$$

$$= \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{r}{\omega}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}}\left(\frac{\alpha_L}{\alpha_K}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} \tag{A.17}$$

Using the inputs, we can derive the optimal cost function which expresses the costs of the firm as a function of target quantities and production parameters:

$$C(K^*, L^*) = rK^* + \omega L^* \tag{A.18}$$

$$= r\left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{\omega}{r}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}}\left(\frac{\alpha_K}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} + \omega\left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}}\left(\frac{r}{\omega}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}}\left(\frac{\alpha_L}{\alpha_K}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} \tag{A.19}$$

Now note that:

$$r\left(\frac{\omega}{r}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} = r\omega^{\frac{\alpha_L}{\alpha_K+\alpha_L}} r^{-\frac{\alpha_L}{\alpha_K+\alpha_L}} = \omega^{\frac{\alpha_L}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K+\alpha_L}{\alpha_K+\alpha_L}-\frac{\alpha_L}{\alpha_K+\alpha_L}} = \omega^{\frac{\alpha_L}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K}{\alpha_K+\alpha_L}} \tag{A.20}$$

$$\omega\left(\frac{r}{\omega}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} = \omega\omega^{-\frac{\alpha_K}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K}{\alpha_K+\alpha_L}} = \omega^{\frac{\alpha_K+\alpha_L}{\alpha_K+\alpha_L}-\frac{\alpha_K}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K}{\alpha_K+\alpha_L}} = \omega^{\frac{\alpha_L}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K}{\alpha_K+\alpha_L}} \tag{A.21}$$

$$C(K^*, L^*) = rK^* + \omega L^* \tag{A.22}$$

$$= \left(\frac{Q}{T}\right)^{\frac{1}{\alpha_K+\alpha_L}} \omega^{\frac{\alpha_L}{\alpha_K+\alpha_L}} r^{\frac{\alpha_K}{\alpha_K+\alpha_L}} \left(\left(\frac{\alpha_K}{\alpha_L}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}} + \left(\frac{\alpha_L}{\alpha_K}\right)^{\frac{\alpha_L}{\alpha_K+\alpha_L}}\right) \tag{A.23}$$

Appendix B

Taxation in the presence of behavioural error

The policy goal is to optimize economic surplus subject to the constraint that marginal decision benefits are equal to the consumer price and the supply price is equal to the marginal decision costs. Define the Lagrangian function as:

$$H = \int_0^Q p_e(Q')dQ' - \int_0^Q s_e(Q')dQ' + \lambda(p_d(Q) - p^* - \tau) \tag{B.1}$$

The derivatives of the Lagrangian function are given by:

$$\frac{\partial H}{\partial Q} = p_e(Q) - s_e(Q) + \lambda \frac{\partial p_d(Q)}{\partial Q} = 0 \tag{B.2}$$

$$\frac{\partial H}{\partial \lambda} = p_d(Q) - p^* - \tau \tag{B.3}$$

$$\frac{\partial H}{\partial \tau} = -\lambda = 0 \tag{B.4}$$

Now note that in the regulated equilibrium: $p_d(Q^{*R}) = b_c p_e(Q^{*R})$ and $p^* = s_d(Q^{*R}) = b_p s_e(Q^{*R})$. Combining the conditions and substituting gives:

$$\frac{1}{b_c}(p^* + \tau) - \frac{1}{b_p}p^* = 0 \tag{B.5}$$

Solving for the tax results in:

$$\frac{1}{b_c}\tau = \frac{1}{b_p}p^* - \frac{1}{b_c}p^*. \tag{B.6}$$

This results in:

$$\tau = p^* \left(\frac{b_c}{b_p} - 1 \right) = p^* \left(\frac{b_c - b_p}{b_p} \right). \tag{B.7}$$

Now add the external costs $\frac{\partial e(Q)}{\partial Q}$ to the economic surplus function. Define the Lagrangian function as:

$$H = \int_0^Q p_e(Q')dQ' - \int_0^Q s_e(Q')dQ' - e(Q) + \lambda(p_d(Q) - p^* - \tau) \tag{B.8}$$

The derivatives of the Lagrangian function are now given by:

$$\frac{\partial H}{\partial Q} = p_e(Q) - s_e(Q) - \frac{\partial e(Q)}{\partial Q} + \lambda \frac{\partial p_d(Q)}{\partial Q} = 0 \tag{B.9}$$

$$\frac{\partial H}{\partial \lambda} = p_d(Q) - p^* - \tau \tag{B.10}$$

$$\frac{\partial H}{\partial \tau} = -\lambda = 0 \tag{B.11}$$

Now note that $p_d(Q) = b_c p_e(Q)$ and $p^* = s_d(Q) = b_p s_e(Q)$. Substituting gives:

$$\frac{1}{b_c}(p^* + \tau) - \frac{1}{b_p} p^* - \frac{\partial e(Q)}{\partial Q} = 0 \tag{B.12}$$

Solving for the tax gives

$$\tau = p^{R*} \left(\frac{b_c - b_p}{b_p} \right) + b_c \frac{\partial e(Q)}{\partial Q}. \tag{B.13}$$

Appendix C

Background calculations for Chapter 4

The derivations in Appendices C.1 and C.2 follow the methodological approach of Verhoef, Nijkamp, and Rietveld (1996) for deriving implicit tax expressions. For the regulated and unregulated case, it is assumed that an equilibrium exists.

C.1 Deriving the tax that optimizes adjusted social surplus

For the case without externalities, the adjusted social surplus AES is given by Eq. (4.5). The Lagrangian function is given by:

$$\mathcal{L} = AES + \lambda(p(Q) - p - \tau) \tag{C.1}$$

and the first-order conditions are given by:

$$\frac{\partial \mathcal{L}}{\partial Q} = \bar{\alpha} p(Q) - \bar{\delta} s(Q) + \lambda \frac{\partial p(Q)}{\partial Q} = 0 \tag{C.2}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = p(Q) - p - \tau = 0 \tag{C.3}$$

$$\frac{\partial \mathcal{L}}{\partial \tau} = -\lambda = 0 \tag{C.4}$$

Substituting Eqs. (C.4) and (C.3) in Eq. (C.3) and using $s(Q) = p^{*R}$ gives an implicit equilibrium expression for the tax:

$$\tau = p^{*R} \frac{\bar{\delta} - \bar{\alpha}}{\bar{\alpha}}. \tag{C.5}$$

C.2 Pricing of a consumption externality

The adjusted social surplus AES with externalities is given by Eq. (4.9). The Lagrangian is given by:

$$\mathcal{L} = AES + \lambda(p(Q) - p - \tau) \tag{C.6}$$

and the first-order conditions are given by:

$$\frac{\partial \mathcal{L}}{\partial Q} = \bar{\alpha} p(Q) - \bar{\delta} s(Q) - \bar{\eta} \frac{\partial e(Q)}{\partial Q} = \lambda \frac{\partial p(Q)}{\partial Q} = 0 \tag{C.7}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = p(Q) - p - \tau = 0 \tag{C.8}$$

$$\frac{\partial \mathcal{L}}{\partial \tau} = -\lambda = 0 \tag{C.9}$$

Substituting Eqs. (C.9) and (C.8) in Eq. (C.8) and rearranging gives the equilibrium expression for the first-best tax:

$$\tau = p^{*R} \frac{\bar{\delta} - \bar{\alpha}}{\bar{\alpha}} + \frac{\bar{\eta}}{\bar{\alpha}} \frac{\partial e(Q)}{\partial Q}. \tag{C.10}$$

Appendix D

Reviewing probability and statistics

This appendix very briefly reviews the statistical knowledge you need for applied econometrics. We assume you had an introductory course in statistics already and will go over this material very quickly.

D.1 Reviewing probability

D.1.1 Probability

As a definition of probability we use the concept of *empirical probability* which is the proportion of time that something (a specific outcome or event X) occurs in the long-run of total events. Usually it is give by:

$$\text{probability} = p = \frac{\text{Number of times specific event } X \text{ happens}}{\text{Total amount of events that can happen}} \tag{D.1}$$

Now probabilities are *defined* by a set of definations (axioms). These are:

1) Probabilities, p, are always between 0 and 1. So, $0 \leq p \leq 1$
2) If something does not happen, then $p = 0$
3) If something always happens, then $p = 1$
4) Probabilities for the total amount of events always add up to 1. So, if the probability that something happens is p, then the probabilities that it will not happen is $1 - p$ (see that $p + 1 - p = 1$)

D.1.2 Population & random variables

In general we see a population as the the group or collection of all possible entities of interest (school districts, inhabitants of the Netherlands, homeowners) and we will think of populations as infinitely large (∞). From this population we then *sample* specific observations. This sample contains then a random variable Y, which denotes a characteristics of the entity (district average test score, prices of houses, prices of meat). An important feature is that the sample characteristics are unknown, that is before measurement (y), after measurement the sample is know and is called data.

So, a random variable (also called a stochastic variable) is a mathematical formalization of something that depends on *random* outcomes. Unfortunately, randomness is not clearly defined and depends on specific scientific philosophical schools. The scientific philosophical school we implicitly assume in this course—and, in fact, in most statistical courses—is that of frequentist statistics. Here we assume that all things we measure are *intrinsically* random. In fact, this is an *ontological* argument—in other words, what are our beliefs in the state of the world. Because all things we measure are random, every time we measure something our measurements are (slightly) different. However, the more we measure, the more precise we *know* something. But there is still randomness.

In general, there are two types of random variables. First, there are *discrete* random variables, where outcomes can be counted, such as $0, 1, 2, 3, \ldots$ and *continuous* random variables, where outcomes can be any real number.[1]

D.1.3 Distribution functions

Random variables are governed by *distribution functions* which are mathematical function that provides the probabilities of occurrence of all different possible outcomes of the random variable *experiment*: e.g. for a discrete distribution, $f(x) = \Pr(Y = y)\ \forall y$. Or, in other words, the distribution function maps discrete outcomes to probabilities. For continuous distribution function, this is not possible as there an infinite number of possible outcomes, so that means that for each y must yield $\Pr(Y = y) = 0$. Therefore, with continuous distribution, often the *cumulative distribution function* is used, which is defined as $F(x) = \Pr(Y \leq y)$. This is why we always use the surface of areas under the *normal* distribution function.

Distribution functions have characteristics of which the most important are: - The mean, also known as the expected value (or expectation) of Y. It is usually denoted as $E(Y) = \mu_Y$ and can as well be interpreted as the long-run average value of Y over repeated realizations of Y: $\frac{1}{n}\sum_{i=1}^{n} y_i$

- The variance, which is denoted as $E(Y - \mu_Y)^2$. Usually it is associated with

[1] There is slightly more to this as fractions such as $\frac{1}{2}$ can in fact be counted as well, and continuous outcomes can be as well complex numbers. But for now we typically see integer numbers as discrete, and real numbers as continuous.

σ_Y^2 and provides a measure of the squared spread of the distribution. If we take the square root then we have the standard deviation ($= \sqrt{\text{variance}} = \sigma_Y$). For a symmetrical normal distribution, it is useful to know that the mean plus or minus 1 time the standard deviation governs about 2/3 of all probability while the mean plus or minus 2 times the standard deviation governs about 95% of all probability associated with that random variable.

Now in statistics we are usually related in relations between random variables, and luckily most entities in real life are related. To capture the relation we need two concept, joint distributions and covariance. If we assume that that random variables X and Z have a joint distribution then the covariance between X and Z is:

$$cov(X, Z) = E[(X - \mu_X)(Z - \mu_Z)] = \sigma_{XZ} \tag{D.2}$$

Note that this covariance is a measure of the *linear* association between X and Z and that its units are units of X times units of Z. $cov(X, Z) > 0$ means a positive relation between X and Z, and finally if X and Z are independently distributed, then $cov(X, Z) = 0$. Note that the covariance of a random variable with itself is just its variance:

$$cov(X, X) = E[(X - \mu_X)(X - \mu_X)] = E[(X - \mu_X)^2] = \sigma_X^2 \tag{D.3}$$

However, the covariance is still measured in the units of X and Z. To correct for that, we often use the correlation coefficient, defined by:

$$corr(X, Z) = \frac{cov(X, Z)}{\sqrt{var(X)var(Z)}} = \frac{\sigma_{XZ}}{\sigma_X \sigma_Z} = r_{XZ} \tag{D.4}$$

where $-1 \leq corr(X, Z) \leq 1$, a $corr(X, Z) = 1$ means perfect positive linear association, a $corr(X, Z) = -1$ means perfect negative linear association, and a $corr(X, Z) = 0$ denotes no linear association.

It is very important to notice that a correlation coefficient measures **linear** association. So, $corr(X, Z) = 0$ does not mean that there is no relation, there is only no linear correlation. This is illustrated by figure D.1. In panel (a) there is clearly a positive relation, and panel (b) shows a negatve relation, but what about panel (d)? Here, the correlation coefficient is 0, just as in panel (c), but obviously there is a clear *non-linear* relation.

D.1.4 Conditional distributions and conditional means

An important notion is applied statistics (and in applied econometrics) is that of the condition distributions, that is the distribution of Y, given value(s) of some other random variable, X. For example, in our California school example, we might want to know something about the distribution of test scores, **given** that $STR < 20$. Therefore, we use the concept of conditional mean, which is defined as the mean of a conditional distribution $= E(Y|X = x)$. Note here

Figure D.1: The correlation coefficient and the relation between observed x and y

the | symbol—it means given that a random variable X is measured with x. As an example: $E(Testscores|STR < 20)$ which denotes the mean of test scores among districts with small class sizes. We also denote this with the *conditional* mean.

Now, if we want to know the difference in means, then we can denote that with

$$\Delta = E(Testscores|STR < 20) - E(Testscores|STR \geq 20), \qquad \text{(D.5)}$$

which is a very important concept in applied economics as it resembles two groups of which one received *treatment* and the other one not. Other examples of the use of conditional means: difference in wages among gender (glass ceiling) and mortality rate differences between those who are treated and those who are. Now if $E(X|Z)$ is constant, then $corr(X, Z) = 0$. We then say that X and Z are independent.

D.2 Sampling in frequentist statistics

So, we mentioned above that we sample from the population which is assumed to be infinitely large. Now, how does this sampling then carry over to statistics. For that we need a statistical framework based on *random sampling*. First, choose an individual, i, (or district, firm, etc.) at random from the population. Now, prior to sample selection, the value of what we want to know Y_i is random because the individual selected is **random**. Once the individual is selected and the value of Y is observed, then Y is just a number—not random anymore but data. And then we say it has the value y. Hence the notation $\Pr(Y = y)$.

If we sample multiple entities, the we construct a data set that looks like $(y_1, y_2, \ldots, y_n)$, where y_i = value of y for the ith individual (district, entity) sampled. Again the lower case here denotes a realisation—the dataset. Now, we want to know what the distribution of the random variables $Y_1, \ldots, Y_n$ is under simple random sampling. Note that because entities (say individuals) #1 and #2 are selected at random, the value of Y_1 has no information content for Y_2. Thus: Y_1 and Y_2 are independently distributed. And if Y_1 and Y_2 come from the same distribution, that is, Y_1, Y_2 are identically distributed, then we say that, under simple random sampling, Y_1 and Y_2 are independently and identically distributed (i.i.d.). More generally, under simple random sampling, Y_i, $i = 1, \ldots, n$, are *i.i.d*—this term always come back in all sorts of statistics.

This simple framework already allows rigorous statistical inferences about, e.g., *the mean* $\bar{Y}$ of population distributions using a sample of data from that population. The next subsection does this because the mean is not only an important, but mainly because the results immediately can be transferred to the regression context.

D.2.1 The sampling distribution of $\bar{Y}$

Now because $\bar{Y}$ is formed by a sample of $\{Y_i\}'s$ it is as well a random variable, and its properties are determined by the *sampling distribution* of $\bar{Y}$. Again, we assume that the elements in the sample are drawn at random, that thus the values of $(Y_1, \ldots, Y_n)$ are random, and that thus functions of $(Y_1, \ldots, Y_n)$, such as $\bar{Y}$, are random: had a different sample been drawn, they would have taken on a different value. Finally, the distribution of $\bar{Y}$ over different possible samples of size n is called the sampling distribution of $\bar{Y}$, which underpins all of *frequentists* statistics.

D.2.2 Example: simple binomial random variables

So how this work. Let's take the easiest statistical example: coin flipping, where the coin is this case is notoriously biased. Suppose the random variable Y takes on 0 (head) or 1 (tails) with the following probability distribution, $\Pr[Y = 0] = 0.22$, $\Pr(Y = 1) = 0.78$. Then the mean and variance are given by:

$$\mu_Y = p \times 1 + (1-p) \times 0 = p = 0.78 \tag{D.6}$$

$$\sigma_Y^2 = E[Y - \mu_Y]^2 = p(1-p) \tag{D.7}$$

$$= 0.78 \times 0.22 = 0.17 \tag{D.8}$$

But this is only one throw ($n = 1$). We would like to have multiple observations to derive at our sampling distribution of $\bar{Y}$, which we assume to depend on the number of throws, n.

Consider therefore first the case of $n = 2$. The sampling distribution of $\bar{Y}$ is,

$$\Pr(\bar{Y} = 0) = 0.22^2 = 0.05 \tag{D.9}$$

$$\Pr(\bar{Y} = 1/2) = 2 \times 0.22 \times 0.78 = 0.34 \tag{D.10}$$

$$\Pr(\bar{Y} = 1) = 0.78^2 = 0.61. \tag{D.11}$$

but this start to become boring as n increases. Therefore, we turn to `STATA`. Let's first check for $n = 2$.

```
set obs 10000
generate Y = rbinomial(2,0.78)/2
hist(Y), fraction
```

The first line of this code sets the number of experiments (called `obs`). So, I throw a coin twice, for 10000 times in a row. The second line generates the outcomes, which in this case are no heads (0), head once (1), or two heads (2). To arrive at probabilities I divide by 2 again. Finally, the third line gives a history of fraction (not counts). And this provides the following `STATA` output.

But what if I do this a 100 times, so $n = 100$?

Figure D.2: Sampling distribution when you throw a coin two times

```
clear
set obs 10000
generate Y = rbinomial(100,0.78)/100
hist(Y), fraction
```

Note that fhe first line now clears STATA's memory as I actually create a new dataset. The histogram can be seen now in Figure D.3.

But isn't this strange. Apart from some omitted bars (which is a fluke of STATA), we can now observe a couple of things. First, the average of the distribution of Figure D.3 is very close to 0.78, which is the actual probability that our biased coin provides tails. But, more importantly the distribution starts to look like a symmetric normal distribution. And we started with a binomial distribution!

This is the result of two amazing statistical theorems:

1. **The law of large numbers**: the average of the results obtained from a large number of trials should be close to the expected value and tends to become closer to the expected value as more trials are performed. That is, if there a no biases in the experiment itself. It also means that with more experiments the precision become better, or the variance decreases. In general this implies that:
   - $\bar{Y}$ is an *unbiased* estimator of μ_Y (that is, $E(\bar{Y}) = \mu_Y$)
   - $\text{var}(\bar{Y})$ is *inversely proportional* to n
   - Thus the standard error associated with $\bar{Y}$ is $\sqrt{\frac{\sigma_Y^2}{n}}$ (that means that with larger samples there is less uncertainty but see the square-root law)
2. **The Central Limit Theorem**: when independent random variables are summed up, their properly normalized sum tends toward a normal distribution even if the original variables themselves are not normally distributed. So $\bar{Y}$ is approximately distributed $N(\mu_Y, \frac{\sigma_Y^2}{n})$
   - When working with standardized variables then $\bar{Y} = \frac{\bar{Y}-\mu_Y}{\sigma_Y/\sqrt{n}}$ is approximately distributed as $N(0,1)$
   - The larger is n, the better is the approximation. And this already holds for $n \geq 50$. So with a reasonable amount of observations, the mean of *i.i.d.* variables is normally distributed

Figure D.3: Sampling distribution when you throw a coin 100 times