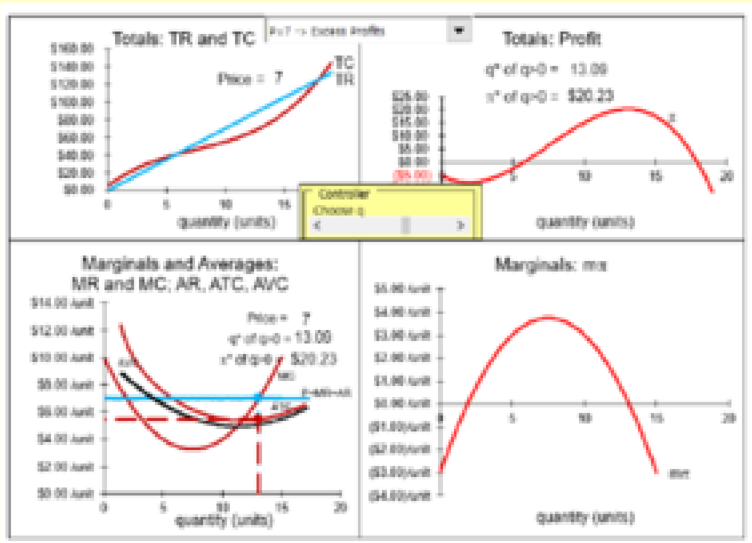
Intermediate Microeconomics

with

Microsoft Excel®

SECOND EDITION



Humberto Barreto

Intermediate Microeconomics with Microsoft Excel®

Humberto Barreto
Professor of Economics and Management
DePauw University

2021

First published in 2009 by Cambridge University Press.

This book was typeset in LaTeX with various packages in TeXstudio. I was helped repeatedly by resources at tex.stackexchange.com. I am awed by this software and its community. I offer a deep bow to those who made these tools freely available and continue to provide support.

ACKNOWLEDGMENTS

This book would not have been possible without the help of many people, but four especially stand out: Frank Howland, Kealoha Widdows, Michele Villinski, and Tami Barreto. Thank you.

I team-taught several courses with Frank and Kay so it is not surprising that their imprint, including examples, phrasing, and pedagogical strategy are embedded in this book. They caught mistakes, gave me ideas, and profoundly influenced my thinking about the best way to teach economics.

Michele sat in on my Intermediate Micro class in the Spring of 2019. She would occasionally give me tips and make suggestions to improve the presentation. I kept a running list and included them in this edition.

Tami copy edited this manuscript, like almost everything else I have written. Her attention to detail and drive for perfection have improved the exposition immensely.

To all of my DePauw and Wabash students.

Really, it has been my pleasure.

Contents

Preface

This is the second edition of a book that was originally published in 2009 by Cambridge University Press. While the core of the book remains the same, this edition refreshes all of the screenshots based on Excel 2019 and updates the data used in real-world applications. It also fixes typos and mistakes. Finally, it includes a new chapter on rational addiction and offers several new optimization problem examples.

The preface of the first edition said:

> In the competitive world of textbooks, different is definitely bad. Authors and publishers, like politicians, stay in the safe middle. Straying too far from the herd is almost a sure way to fail. Fear is strong, but it apparently can be overcome—after all, you are reading a spectacularly unconventional textbook.
>
> The most obvious difference between this book and the usual fare is the use of Microsoft Excel to teach economic theory. This enables students to acquire a great deal of sophisticated, advanced Excel skills while learning economics. No other book does this.
>
> The use of Excel drives other differences. Excel requires concrete, numerical problems instead of the abstract functions and graphs used by other books. Excel's Solver makes possible presentation of numerical methods for solving optimization problems and equilibrium models. No other book does this.
>
> Because numerical solutions are readily available, this book is able to present and explain analytical methods that have been pushed to appendixes or completely ignored in mainstream texts. Every problem is solved twice—once with Excel and once with equations, algebra, and, when needed, calculus. No other book does this.

> Finally, this book is organized differently. It explicitly repeats a single central methodology, the economic approach, so students learn how economists think and how to think like an economist. Other books try to do this, but none brings the economic way of thinking explicitly to the surface, repeating the message in every application.
>
> I wrote this book because I learned Visual Basic and quickly realized that enhancing a spreadsheet with macros made possible a whole new way of teaching economics. When my students loved this approach, I wanted to share it with others.
>
> Because this book is so different, it will probably not challenge the top sellers. It will be the unusual professor who is willing to try something this new. It requires that the professor care enough about students and teaching to invest time and energy in mastering the material. Of course, I think the rate of return is quite high. My hope is that, though few in number, a committed, enthusiastic core of adopters will enable this book to survive.
>
> Thank you for trying this unique entry into the competitive market for micro theory textbooks. I hope you find that the reward was worth the risk.

Well, after more than ten years, I can safely say that I certainly was right that the book would not challenge the top sellers! It strayed far from the herd and went largely unnoticed. When I asked Cambridge University Press to do a second edition, they politely declined.

But, I am not giving up. I believe that teaching economics via Excel is a winner. So, I am ignoring the market, producing my own second edition, and giving it away for free.

I am well aware that this edition will not attract many adopters and that I am engaged in a quixotic fight against foes who are not even aware of my presence. I remain baffled at how badly microeconomics is taught—it is as if computers were never invented. We can and must do better. I will keep this book alive in case someone wants to try a novel, innovative approach to teaching and learning microeconomics.

This edition assumes that many will read it electronically, although you are free to print it out and I am so old school that I certainly would prefer handwriting notes and underlining on paper. Any print shop can do this and, if anyone asks, explain that this is an open access book and you have legal right to print it. You can also print it online at sites such as www.lulu.com/.

I think Adobe Acrobat Reader is a good choice if you decide to read it on screen, but you are, of course, welcome to use your favorite eReader. Here is a list of 15 pdf readers: blog.hubspot.com/marketing/best-free-pdf-reader. One advantage of digital access is that links are highlighted for easy clicking. You should use your pdf reader's commenting capabilities to highlight, search (ctrl-f), and take notes. It should also be easy to look up words you do not know or search for ideas that pique your interest so take full advantage of the electronic tools at your disposal.

I have been teaching economics for a long time now. I am positive that using Excel to learn how economists use models and see the world works for almost all students. You can learn a lot of economics, math, and Excel while working with this book. Do your best and good luck!

Humberto Barreto
hbarreto@depauw.edu
Greencastle, Indiana
November 11, 2021

> The idea for the electronic spreadsheet came to me while I was a student at the Harvard Business School, working on my MBA degree, in the spring of 1978.
>
> Dan Bricklin

User Guide

This book is essentially a manual for how to actively work with and manipulate the material in Excel. This user guide lists minimum requirements, provides instructions for downloading all of the materials and software, offers a few tips before you begin, and describes the organization of the files.

Minimum Requirements

; This book presumes that you have access to and a basic knowledge of Excel. In other words, you can open an Excel file (called a *workbook*), write a formula that adds cells together, make a chart, and save the file. As you will see, however, Excel is much more than a simple adding machine. You will learn how to use Excel in a more advanced way. In addition to analyzing data and learning many new Excel functions, you will solve optimization problems with an *add-in* (a special file that extends the functionality of Excel) called Solver.

The materials in this book will work on any Windows Excel version all the way back to 1997 (version 8). The screenshots are based on Excel 2019, but if you are using an earlier version, it should be easy to figure out what to do.

The workbooks and add-ins are optimized for use with Windows Excel. They can be accessed with a Macintosh computer, but Solver in Mac Excel is temperamental and buggy. Furthermore, Visual Basic (Excel's macro language) on a Mac is limited so not all macros work. The best solution for Mac users is to emulate Windows with software such as Parallels or Boot Camp. For students at an educational institution, accessing Excel from a server (see, for example, VMWare's Horizon software) is an easy solution for Mac users. Desktops.depauw.edu gives my students access to a Windows machine running Excel configured with necessary add-ins.

To ensure that older versions of Excel can open the files, all workbooks have been saved in "compatibility mode" (Excel 97 – 2003 Workbook) with the

.xls filename extension. If you are using Excel 2007 (version 12) or greater, you should save your completed files in the "Excel macro-enabled workbook" format, which carries the .xlsm extension. Do not save your files as an Excel workbook with the .xlsx extension, the macros will not be saved and functionality will be lost.

For non-English versions of Excel, the files will work in the sense that buttons, scroll bars, and macros will function; however, the add-ins and other content will not be translated.

Recently, Microsoft Office has moved online, offering OneDrive and Office 365 cloud access. Regrettably, as of this writing, because of security concerns, online versions of Office do not support Visual Basic, a limitation which renders these options useless for working with macro-enhanced files from within a web browser. You can save a file with macros in your favorite storage area in the cloud, but you will need to download and open it with a desktop Excel version to run the macros. Within a browser, macros cannot be executed.

Downloading and Opening Workbooks

Visit www.depauw.edu/learn/microexcel to download the files that accompany this book. You may download individual files as needed or a compressed archive with all of the files to as many different computers or devices as needed.

Figure 1 shows that, when opening a workbook with macros, Excel will alert you to their presence with a security warning under the Ribbon (and right above the formula bar).

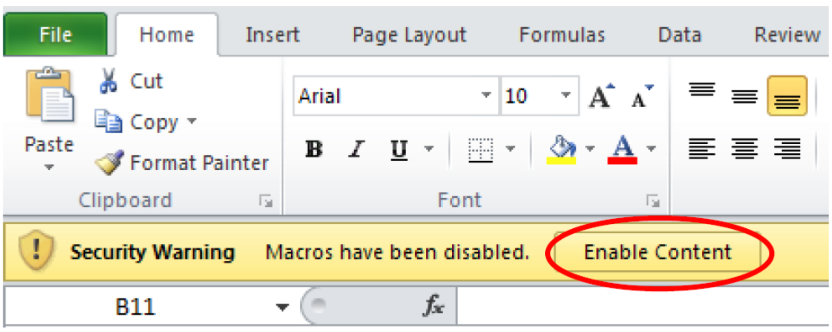


Figure 1: *Enable Content* when opening a Micro Excel workbook.

If you do not see the security warning or have no opportunity to enable content, your security level has been set to block all files with macros. Although malicious code can be harmful, you must dial down the safety measures to allow Excel to utilize fully the information in the workbook. Close the file and change the security setting to allow Excel to open files with macros.

Visit Excel's main support page at support.office.com for more help on setting security and enabling macros.

Tips and Conventions

In this book, a *figure* refers to a variety of graphics, including charts and pictures of portions of a sheet (also known as a screenshot, like Figure 1). A chart or range of cells is often displayed in this printed book as a figure, but you should look at the live version on your computer screen. Thus, in addition to a caption, many figures have a source line indicating their location in the Excel workbook.

The book follows Excel's naming convention for workbooks, sheets, and cells: [workbookname]sheetname!cell address. If the caption of a figure says, [FoodStamp.xls]BudgetConstraint, then you know the figure can be found in the *FoodStamp.xls* workbook in the *BudgetConstraint* sheet. Note that workbook and sheet names in the printed text are italicized to help you locate the proper sheet in a workbook. [RiskReturn.xls]OptimalChoice!B6 refers to cell B6 in the *OptimalChoice* sheet of the *RiskReturn.xls* workbook.

You may need to adjust your display or the objects in Excel. Use the Zoom button to magnify the display. You can also right-click objects such as buttons or scroll bars to select and move them. Once you open a workbook, you can save it to another location or name (by executing File → Save As...) and make whatever changes you wish. This is the same as underlining or writing in a conventional, printed book.

Finally, if something is not working the way you expect, there are many possible causes. It is always a good idea to close Excel completely and reopen it. Even if this does not fix the problem, slowly repeating the steps will help you debug or describe what is happening.

Organization of Files

Figure 2 shows the contents of all materials included in the MicroExcel.zip archive, after downloading it from www.depauw.edu/learn/microexcel.

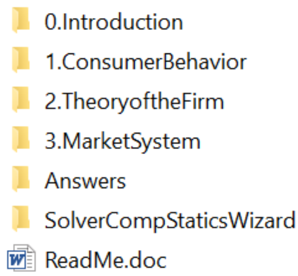


Figure 2: Organization of files.

The Answers folder contains answers to questions posed in Q&A sheets in each Excel workbook. Think of the Q&A material in the Excel workbooks as self-study questions.

There are also Exercises at the end of each chapter. Readers do not have easy access to the answers to the exercise questions. To see these answers, you must be an instructor and register online at www.depauw.edu/learn/microexcel.

The SolverCompStaticsWizard folder contains files that use the Comparative Statics Wizard Excel add-in. When used in conjunction with Excel's own Solver add-in, these files enable numerical comparative statics analysis of optimization problems and equilibrium models.

Active Learning

The most important thing you can do as you read this book is **experiment**. You might find yourself wondering, "What would happen if this cell was 10 instead of 1?" Do not just wonder, change the cell and see what happens! There is deep neuroscience at work here. When you are in control and making up your own questions, you learn best. The beauty of this approach is that everything is alive and you can make points move and lines shift. Take full advantage.

Remember that you can always download the original workbook again if needed. This means you should not worry about changing anything in a workbook. If something goes terribly wrong, simply delete it and download it again.

There are many books devoted to microeconomics. This one is different because it is not meant to be simply read. A great deal of its value lies in the Excel workbooks and additional materials. By reading this book and working in Excel simultaneously, you will become a sophisticated user of Excel and learn a great deal of mathematics and, most importantly, economics.

Download the files from www.depauw.edu/learn/microexcel and get to work!

Spreadsheet History and Resources

For more on the history of the electronic spreadsheet, as told by one of the creators, see bricklin.com/visicalc.htm. This is the source for the epigraph.

I recommend these websites for Excel tips and tricks, workbook and add-in downloads, and Visual Basic code snippets:

- Tushar Mehta: www.tushar-mehta.com/excel/
- Chip Pearson: www.cpearson.com/excel
- Jon Peltier: peltiertech.com/Excel/
- Andy Pope: www.andypope.info

Economics is the science which studies human behavior as a relationship between given ends and scarce means which have alternative uses.

Lionel Robbins

A First Step

Economists see the world through a special pair of glasses. It takes practice and concentration to learn how to see things like an economist. The interpretation of reality that is the hallmark of modern economics has been called the economic way of thinking, the economic approach, and the method of economics. Thinking and seeing the world like an economist is the ultimate goal of this book.

You will learn the economic way of thinking by working through many examples. Here is the first one.

Optimal Allocation of Worker Hours

Suppose that you manage a tech support service for a major software company. You have two types of callers: *Regular* and *Preferred*. Your preferred customers have paid extra money for faster access, which means they expect to spend less time waiting on hold. There are equal numbers of the two types of customers and they call with equal frequency.

Management has given you a fixed number of worker hours per day to answer calls from users needing help. Daily, you have 10 workers, each working 8-hour shifts, and 5 part-time workers (4-hour shifts each); or 100 hours per day in total to support customers calling for help. These 100 hours comprise your *Total Resources*.

When customers call, an automatic message is played asking the caller to input an ID number and the caller is put on hold. The ID number is used to identify the caller as a regular or preferred customer.

Keeping callers on hold creates frustrated, unhappy customers. The callers are already angry since something has gone wrong with the software and they need help. The faster you get support to the caller the better. You keep track of *time waiting* (the amount of time, in seconds, that the typical caller is on hold) and you know that it depends on the number of worker hours available to answer the calls.

To keep things simple, assume typical time waiting = 6000/worker hours allocated. So, say there are 80 worker hours available to answer preferred callers. Dividing 6000 by 80 yields 75, which means the typical hold time is 75 seconds. This leaves 20 worker hours for regular callers, so their hold time is 300 seconds (since 6000/20 = 300). Five minutes is a long time to wait on the phone!

The problem becomes an *economic problem* because you have two types of callers, so you must decide how to allocate your worker hours. When you have to make a decision where you trade-off one thing for another you are doing economics. In this case, the more hours you allocate to one type of caller, the lower that caller's wait time. That's the good news.

The bad news is that the fixed amount of caller-support hours means that more time devoted to one type of caller results, by definition, in fewer hours to the other type and, therefore, higher waiting times for the other type.

So the general structure of the problem is clear: You must decide how to allocate scarce support resources (worker hours) to two competing ends. Figure 3 shows a simplified picture of the problem.

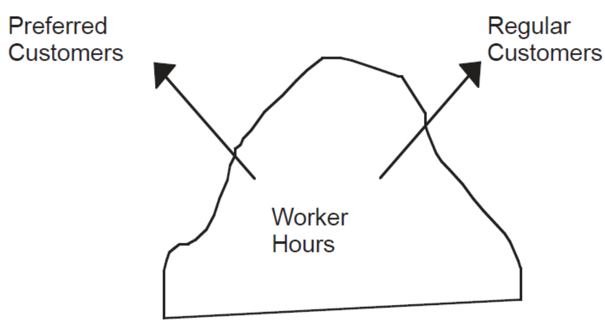


Figure 3: Allocating a scarce resource to two competing ends.

A Complication

It is unclear exactly what preferred customers expect. Do they expect to get help twice as fast or 10 times as fast as regular customers?

To incorporate the fact that the preferred customer merits greater attention, management gives you a *value weight* parameter. The value weight tells you how much more valuable the preferred caller is compared to the regular caller.

We can write the *objective function* as

$$TotalTimeWaiting = \frac{6000}{RegHours} + ValueWeight\frac{6000}{PrefHours}$$

The objective function says that time spent waiting by a preferred caller is multiplied by a factor that reflects how much more we value the preferred customer's time. If $ValueWeight = 1$, then preferred and regular callers are equally valuable. Management has decreed that preferred customers' time is worth twice that of regular customers so $ValueWeight = 2$; you (the call center manager) cannot change this parameter.

So, if you decide to allocate 50 hours each to the regular and preferred customers, then both types of customers will wait 6000/50 = 120 seconds and our objective function will be 120 + 2 x 120 = 360 seconds.

Is there a better allocation, one that yields a smaller total time waiting (adjusted with the value weight), than 50/50? This question, how to allocate 100 worker hours to answering calls from regular and preferred customers in order to minimize value weighted total time waiting, has an answer, called the optimal solution. We have to find it.

Setting Up the Problem

We will solve this problem by first organizing the information into three separate parts. All optimization problems can be set up the same way, with three components: *goal*, *endogenous variables*, and *exogenous variables.*

The goal is synonymous with the objective function. Endogenous, or choice, variables can be controlled by the decision maker. Exogenous variables are given, fixed constants that cannot be changed by the decision maker. The exogenous variables (sometimes called parameters or independent variables) form the environment under which the decision maker acts.

In the tech support time minimization problem, we can organize the information like this:

1. Goal: minimize total time waiting (value weighted)
2. Endogenous variables: worker hours allocated to preferred and regular customers
3. Exogenous variables: total worker hours and value weight

STEP Open the Excel workbook *Introduction.xls*, read the *Intro* sheet, and then go to the *SetUp* sheet to implement the problem in Excel.

This workbook (along with all of the files that accompany this book) is available for download at www.depauw.edu/learn/microexcel. The User Guide has detailed instructions on how to properly configure Excel before downloading and opening these files.

Make sure that you enable macros when you open the file. If the buttons do not work, the most likely suspect is in the security settings.

STEP Answer the three questions in column A (below the exogenous variables). Check yourself by clicking the See Answer buttons.

Finding the Initial Solution

Now that we have set up the problem, we can turn our attention to finding the answer, the optimal solution. There are two ways to solve optimization problems:

- Analytical (algebra and calculus) methods
- Numerical (computer) methods

The analytical method uses pencil and paper to write down equations and manipulate them to find the answer. It was the only way available until computers came along and gave us algorithms for finding solutions. Numerical methods rely on testing many trial solutions very quickly and repetitively, converging to the answer. We will ignore the analytical approach in this example and concentrate on showing how Excel's Solver works.

STEP Click the Data tab (in the Ribbon across the top of the screen), then Solver (in the Analysis group) to bring up the Solver dialog box (as in Figure 4). If Solver is not available, then use the Add-in Manager to install it. Use Excel's Help if you are having trouble or visit support.office.com.

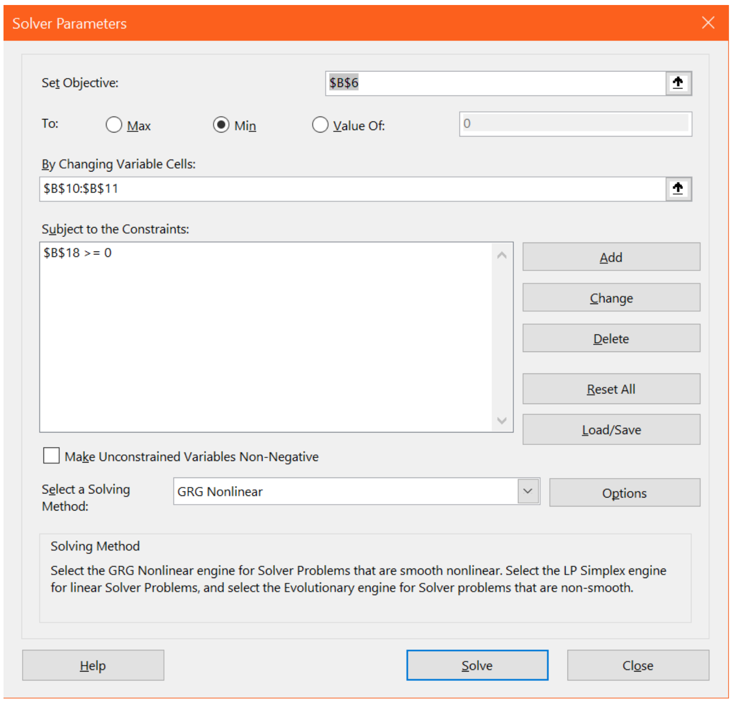


Figure 4: The Solver dialog box.

Note that necessary information is already entered. The objective cell is the (value weighted) total time waiting, the changing variable cells (the endogenous variables) are the worker hours devoted to the regular and preferred customers, and the constraint is that the sum of the worker hours not exceed the 100 hours you have been given.

STEP Click the Solve button to find the solution to the problem. Click the OK button in the Solver Results dialog box to accept Solver's solution and put the optimal solution in the *SetUp* sheet.

Congratulations! You, the call center manager, have just used Solver (a numerical methods approach to optimization) to optimally allocate your scarce resources. We can check Solver's answer for plausibility, noting that it makes sense that preferred callers have more hours allocated to them because they are more valuable. Later, we will see that we can solve this problem using analytical methods and if the two approaches give the same answer, we can be confident that we do indeed have the best solution.

Comparative Statics

We have found the initial solution, but we are usually much more interested in a follow up question: How will the optimal solution change if the environment changes?

Comparative statics is a shorthand way of describing the following procedure: Change an exogenous variable, holding the other parameters constant, and track how the optimal solution changes in response to the shock.

Like finding the initial solution, comparative statics can be done via analytical (algebra and calculus) and numerical (computer) methods. The Comparative Statics Wizard (CSWiz) add-in was used to explore how the optimal allocation of total worker hours would change if worker hours were increased by 10 hours. The CSWiz add-in will be introduced later and you will learn how to do your own comparative statics analyses. For now, we will focus on what it produces.

STEP See the results of the comparative statics analysis by going to the *CS1* sheet.

Cells A1:D15 in the *CS1* sheet were produced by the CSWiz add-in. It is easy to see that increased total worker hours are allocated to regular and preferred customers in a stable pattern. Every additional hour of total resources, holding value weight (the only other exogenous variable in this simple problem) constant, produces an increase of 0.586 hours allocated to preferred customers. The chart below the data (row 16) shows the linear relationship. Usually, economists want to determine the relationship between optimal endogenous and exogenous variables.

Summary: Introducing Optimization

This chapter used an example to show how Excel's Solver can find the optimal solution. It introduced the basics of optimization, including the three parts of every optimization problem:

1. Goal (or objective function),
2. Endogenous variables,
3. Exogenous variables.

As you work with this book, you will learn how to use analytical methods to solve optimization problems. You will also learn how to do comparative statics analysis via analytical and numerical methods.

This introductory example was completely prepared for you. All you had to do was click a few buttons. Future problems will gradually relax the Excel environment, giving you ever more freedom to make decisions and thereby learn what to do. The ultimate goal is for you to be able to set up and solve problems yourself.

Exercises

1. Suppose Management decides that preferred customers are three times as important as regular customers, so that the $ValueWeight = 3$. With 100 workers hours, what is the optimal solution? Describe your procedure and report the optimal values of *PrefHours* and *RegHours*.
2. Compared to the initial solution, when $ValueWeight = 2$, what is the change in the number of hours allocated to preferred customers?
3. The percentage change in *ValueWeight* is 50% (from 2 to 3). What is the percentage change in the number of hours allocated to preferred customers?

References

Each section ends with references and resources for further study. A citation for the epigraph (lead quotation) of the chapter is provided. References may also contain citations documenting sources used, additional information on the history of a concept or person, and suggestions for further reading.

The epigraph to this chapter is found on page 16 of the second edition of *An Essay on the Nature and Significance of Economic Science* by Lionel Robbins. This book was originally published in 1932 and the second edition is available online at www.mises.org/books/robbinsessay2.pdf. Robbins rejects old definitions of economics based on content (the study of business and work) and argues for a definition of economics based on methods used: optimization and comparative statics. Robbins made the definition of economics (in the epigraph to this chapter) famous, but he includes a footnote that cites various precursors who used a similar description of economics.

For more on Robbins, visit www.econlib.org/library/Enc/bios/Robbins.html. Econlib says that Robbins' Essay is "one of the best-written prose pieces in economics."

Nobel laureate Gary Becker's *The Economic Approach to Human Behavior* (first published in 1976) has a classic introductory chapter on the meaning of the economic approach and applies economic analysis to such non-standard topics as discrimination, crime, and marriage. Becker's statement, "what most distinguishes economics as a discipline from other disciplines in the social sciences is not its subject matter but its approach" (p. 5), greatly extends the scope of economics.

Modern economics pays little attention to its own history and how we got to be where we are today. The epigraphs in this book highlight important contributions and individuals (like Robbins and Becker) in the development of modern economic theory. Remember to experiment by clicking and searching items that catch your eye.

In Spring 2012, I videotaped my Intermediate Microeconomics classes at DePauw University. They are about an hour long and are freely available at www.depauw.edu/learn/microexcel/videos.htm. The introduction lecture covers material from this chapter.

Part I

The Theory of Consumer Behavior

Perhaps science does not develop by the accumulation of individual discoveries and inventions.

Thomas S. Kuhn

Overview

The material in this book is organized into three parts. The first part focuses on the Theory of Consumer Behavior and derives the demand curve. The second part derives the supply curve from the Theory of the Firm. Finally, these curves are combined to explain how the Market System functions as a decentralized resource allocation mechanism.

Figure I.1 expands the material in the first part, the Theory of Consumer Behavior, to give a preview of upcoming topics. The Optimal Choice chapter is key because it shows how to solve the consumer's optimization problem, but the chapter that follows is especially critical. It applies comparative statics analysis, changing the price of a good, holding everything else constant, to derive a demand curve. This is the most important concept in the Theory of Consumer Behavior.

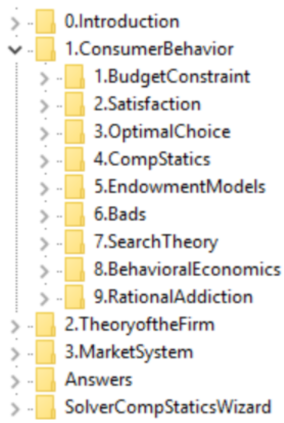


Figure I.1: Content map with focus on consumer behavior.

Focus on the repeated patterns as you work through this material. Economics has a core logic that has been referred to as "the economic way of thinking" or "the economic approach." Learning to see and think like an economist should be your ultimate goal.

References

The epigraph is from the second page of the introductory chapter to Thomas S. Kuhn's classic, *The Structure of Scientific Revolutions* (originally published in 1962). Kuhn argued that progress in science is not generated by incremental puzzle solving (what he called normal science), but that periods of calm are followed by crises that lead to paradigm shifts. The book was as revolutionary as the material it covered, causing debate and controversy in philosophical and scientific circles.

Kuhn would not have been surprised to hear that the derivation of the demand curve did not proceed in an incremental, linear fashion. In fact, the idea of demand for a product depending on the price was known well before we drew graphs of demand curves (in the second half of the 19^{th} century). It was not until economics adopted quantitative and mathematical techniques (what we now call the Marginal Revolution) that the theory of consumer behavior was developed and we could mathematically derive a demand curve.

If we hold money income constant
and allow the price of X to change,
the price ratio line will rotate
about a pivot on the Y axis.

Milton Friedman

Chapter 1

Budget Constraint

The basic idea of the Theory of Consumer Behavior is simple: Given a budget constraint, the consumer buys a combination of goods and services that maximizes satisfaction, which is captured by a utility function. By changing the price of a particular item, ceteris paribus (everything else held constant), we derive a demand curve for that item.

Setting up and solving the consumer's utility maximization problem takes some time. We will proceed slowly and carefully. This chapter focuses on the budget constraint and how it changes when prices or income change.

What can be afforded is obviously a key factor in predicting buying behavior, but it is only part of the story. With the budget constraint alone, we cannot answer the question of how much the consumer wants to buy of each product because we are not incorporating any information about the utility gained by consumption. After we understand the budget constraint, we will model the consumer's likes and dislikes. We can then put the constraint and utility components together and solve the model.

The Budget Constraint in Equation Form

The budget constraint can be expressed mathematically like this:

$$p_1x_1 + p_2x_2 \leq m$$

This equation says that the sum of the amount of money spent on good x_1, which is the price of x_1 times the number of units purchased, or p_1x_1, and the amount spent on good x_2, which is p_2x_2, must be less than or equal to the amount of income, m (for money), the consumer has available.

Obviously, the model would be more realistic if we had many products that the consumer could buy, but the gain in realism is not worth the additional cost in computational complexity. We can easily let x_2 stand for "all other goods."

Another simplification allows us to transform the inequality in the equation to a strict equality. We will assume that no time elapses so there is no saving (not spending all of the income available) or borrowing. In other words, the consumer lives for a nanosecond – buying, consuming, and dying the same instant. Once again, this assumption is not as severe as it first looks. We can incorporate saving and borrowing in this model by defining one good as present consumption and the other as future consumption. We will use this modeling technique in a future application.

Since we know we will always spend all of our income, the budget constraint equation can be written with an equal sign, like this

$$p_1x_1 + p_2x_2 = m$$

Since we will want to draw a graph, we can write in the form of the equation of a line ($y = mx + b$) via a little algebraic manipulation:

$$p_1x_1 + p_2x_2 = m$$

$$p_2x_2 = m - p_1x_1$$

$$x_2 = \frac{m}{p_2} - \frac{p_1}{p_2}x_1$$

The intercept, m/p_2, is interpreted as the maximum amount of p_2 that the consumer can afford. By buying no x_1 and spending all income on x_2, the most the consumer can buy is m/p_2 units of good 2.

The slope, $-p_1/p_2$, also has a convenient interpretation: It states the rate at which the market requires the consumer to give up x_2 in order to acquire x_1. This is easy to see if you remember that the slope of a line is simply the rise (Δx_2) over the run (Δx_1). Then,

$$\frac{\Delta x_2}{\Delta x_1} = -\frac{p_1}{p_2}$$

A Numerical Example of the Budget Constraint

STEP Open the Excel workbook *BudgetConstraint.xls*, read the *Intro* sheet, and then go to the *Properties* sheet to see the budget constraint.

Figure 1.1 shows the organization of the sheet. As you can see, the consumer chooses the amounts of goods 1 and 2 to purchase, given prices and income.

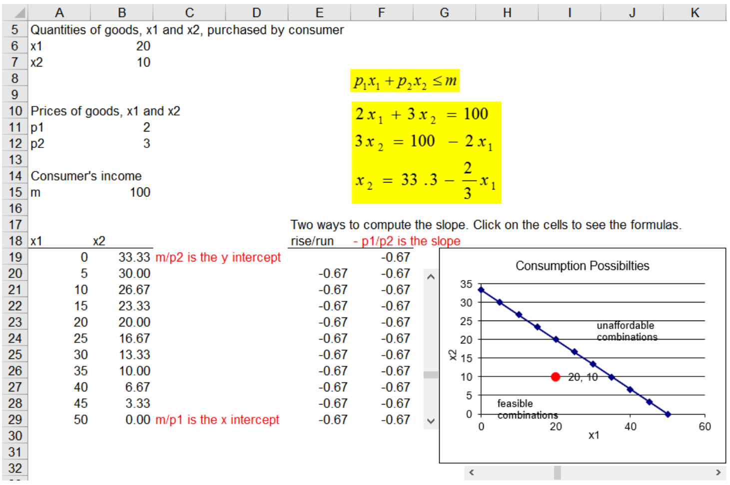


Figure 1.1: The budget line.
Source: BudgetConstraint.xls!Properties

With $p_1 = \$2/\text{unit}$, $p_2 = \$3/\text{unit}$ and $m = \$100$, the equation of the budget line can be computed.

STEP Click on the scroll bars to see the red dot (which represents the consumption bundle), move around in the chart.

By rewriting the budget constraint equation as a line and then graphing it, we have a geometric representation of the consumer's consumption possibilities. All points inside or on the budget line are feasible. Points northeast of the budget line are unaffordable.

By clicking the scroll bars you can easily see that the consumer has many feasible points. The big question is, Which one of these many affordable combinations will be chosen? We cannot answer that question with the budget constraint alone. We need to know how much the consumer likes the two goods. The constraint is simply about feasible options.

Changes in the Budget Line – Pivots and Shifts

STEP Proceed to the *Changes* sheet.

The idea here is that changes in prices cause the budget line to *pivot* or *rotate*, altering the slope, but keeping one of the intercepts the same. Note that changes in income produce a different result, *shifting* the budget line in or out, leaving the slope unchanged.

STEP To see how the budget line pivots, experiment with cell K9 (the price of good 1). Change it from 2 to 5.

The chart changes to reveal a new budget line. The budget line has rotated around the y intercept because if the consumer decided to spend all income on x_2, the amount that could be purchased would remain the same.

If you lower the price of good 1, the budget line swings out. Confirm that this is true.

STEP Changing cell K10 alters the budget line by changing the price of good 2. Once again, change values in the cell to see the effect on the budget line.

STEP Next, click the Reset button to return the sheet to its initial values and work with cell K13. Cut income in half. The effect is dramatically different. Instead of rotating, the budget line has shifted in. The slope remains the same because prices have not changed. Increasing income shifts the budget line out.

This concludes the basics of budget lines. It is worth spending a little time playing with cells K9, K10, and K13 to reinforce understanding of the way budget lines move when there is a change in a price or income. These shocks will be used again when we examine how a consumer's optimal decision changes when prices or income change.

Remember the key lesson: Change in price *rotates* the budget line, but change in income *shifts* it.

Funky Budget Lines

In addition to the standard, linear budget constraint, there are many more complicated scenarios facing consumers. To give you a taste of the possibilities, let us review two examples.

STEP Proceed to the *Rationing* sheet.

In this example, in addition to the usual income constraint, the consumer is allowed a maximum amount of one of the goods. Thus, a second constraint (a vertical line) has been added. When the maximum is above the x_1 intercept (50 units), this second constraint is said to be nonbinding. As you can see from the sheet, when the maximum amount constraint is binding, it lops off a portion of the budget line.

STEP Change cell E13 to see how changing the rationed amount affects the budget constraint.

As we increase the amount of the subsidy, the horizontal line is extended. The downward sloping part has the same slope, but it is pushed outwards,

STEP Proceed to the *Subsidy* sheet.

In this example, in addition to the usual income constraint, the consumer is given a subsidy in the form of a fixed amount of the good.

Food stamps are classic example of subsidies. Suppose the consumer has \$100 of income, but is given \$20 in food stamps (which can only be spent on food), and food (x_1) is priced at \$2/unit. Then the budget constraint has a horizontal segment from 0 to 10 units of food because the most x_2 (other goods) that can be purchased remains at m/p_2 from 0 to 10 units of food (since food stamps cannot be used to buy other goods).

STEP Change cell E13 to see how changing the given amount of food (which is the dollar amount of food stamps divided by the price of food) affects the budget constraint.

Summary: Consumption Possibilities

The budget constraint is a key component of the optimization problem facing the consumer. Graphing the constraint lets us see the consumer's options. Just like a production possibilities frontier tells us what an economy can produce, the budget constraint shows what a consumer can buy. Any combination on or under the constraint is a feasible option. Points beyond the constraint are unattainable.

Changing prices has a different effect on the constraint than changing income. If prices change, the budget line pivots, swings, and rotates (pick your favorite word and remember it) around the intercept. A change in income, however, shifts the line (out or in) and leaves the slope unaffected.

The basic budget constraint is a line, but there are many other scenarios faced by consumers in which the constraint can be kinked or nonlinear. Subsidies (like food stamps) can be incorporated into the basic model. This flexibility is one of the powerful features of the Theory of Consumer Behavior.

The constraint is just one part of the consumer's optimization problem. The desirability of goods and services, also known as tastes and preferences, is another important part. The next chapter explains how we model satisfaction from consuming goods and services.

Exercises

1. Use Excel to create a chart of a budget constraint that is based on the following information: $m = \$100$ and $p_2 = \$3/\text{unit}$, but $p_1 = \$2/\text{unit}$ for the first 20 units and \$1/unit thereafter. Copy your chart and paste it in a Word document.

   STEP Watch a quick, 3-minute video of how to make a chart in Excel by visiting vimeo.com/econexcel/how-to-chart-in-excel.

2. If the good on the y axis is free, what does the budget constraint look like?

3. What combination of shocks could make the new budget line be completely inside and steeper than the initial budget line?

4. What happens to the budget line if all prices and income doubles?

References

The epigraph of this chapter can be found on page 48 of Milton Friedman's revised edition of his *Price Theory* text. The book is essentially his lecture notes from the famous two-quarter price theory course that Friedman delivered for many years at the University of Chicago. It is interesting to see how Micro was taught back then, especially how little emphasis was placed on mathematics. The problems in appendix B are truly thought provoking.

Chapter 2

Satisfaction

Preferences

Utility Functions

> [Indifference] curves are negatively sloped, pass through every point in commodity space, never intersect, and are concave from above. The last-mentioned property implies that the marginal rate of substitution of X for Y diminishes as X is substituted for Y so as to maintain the same level of satisfaction.
>
> C. E. Ferguson

2.1 Preferences

The key idea is that every consumer has a set of likes and dislikes, desires, and tastes, called *preferences*. Consumer preferences enable them to compare any two combinations or bundles of goods and services in terms of better/worse or the same. The result of such a comparison has two outcomes:

- Strictly preferred: the consumer likes one bundle better than the other.
- Indifferent: the consumer is equally satisfied with the two bundles.

In terms of algebra, you can think of strictly preferred as greater than (>), indifferent as equal (=).

Since the consumer can compare any two bundles, then by repeated comparison of different bundles the consumer can rank all possible combinations from best to worst (in the consumer's opinion).

Three Axioms

Three fundamental assumptions are made about preferences to ensure internal consistency:

1. Completeness: the consumer can compare any bundles and render a preferred or indifferent judgment.
2. Reflexivity: this identity condition says that the consumer is indifferent when comparing a bundle to itself.
3. Transitivity: this condition defines an orderly relation among bundles so that if bundle A is preferred to bundle B and bundle B is preferred to bundle C then bundle A must be preferred to bundle C.

Completeness and reflexivity are easily accepted. Transitivity, on the other hand, is controversial. As a matter of pure logic, we would expect that a consumer would make consistent comparisons. In practice, however, consumers may make intransitive, or inconsistent, choices.

An example of intransitivity: You claim to like Coke better than Pepsi, Pepsi better than RC, and RC better than Coke. The last claim is inconsistent with the first two. If Coke beats Pepsi and Pepsi beats RC, then Coke must really beat RC!

In mathematics, numbers are transitive with respect to the comparison operators greater than, less than, or equal to. Because 12 is greater than 8 and 8 is greater than 3, clearly 12 is greater than 3.

Sports results, however, are not like math. Outcomes of games can easily yield intransitive results. Michigan might beat Indiana and in its next game Indiana could defeat Iowa, but few people would claim that the two outcomes would guarantee that Michigan will win when it plays Iowa.

When we assume that preferences are transitive, it means that the consumer can rank bundles without any contradictions. It also means that we are able to determine the consumer's choice between two bundles based on answers to previous comparisons.

Displaying Preferences via Indifference Curves

The consumer's preferences can be *revealed* by having her choose between bundles. We can describe a consumer's preferences with an *indifference map*, which is made up of *indifference curves.*

A single indifference curve is the set of combinations that give equal satisfaction. If two points lie on the same indifference curve, this means that the consumer sees these two bundles as tied – neither one is better nor worse than the other.

A single indifference curve and an entire indifference map can be generated by having the consumer choose between alternative bundles of goods. We can demonstrate how this works with a concrete example.

STEP Open the Excel workbook *Preferences.xls*, read the *Intro* sheet, and then go to the *Reveal* sheet to see how preferences can be mapped and the indifference curve revealed.

STEP Begin by clicking the [Ask?] button. For bundle B, enter 4, then a comma (,), then a 3, then click OK.

We are using the coordinate pair notation so 4,3 identifies a combination that has 4 units of the good on the x axis and 3 units of the good on the y axis.

The sheet records the bundles that are being compared in columns A and B and the outcome in column C. The choices are being made by a virtual consumer whose unknown preferences are in the computer. By asking the virtual consumer to make a series of comparisons, we can reveal the hidden preferences in the form of an indifference curve and indifference map.

Notice that Excel plots the point 4,3 on the chart. The green square means the consumer chose bundle B. This means that 3,3 and 4,3 are not on the same indifference curve.

STEP Click the [Ask?] button again. Offer the consumer a choice between 3,3 and 2,3.

This time the consumer chose bundle A and a red triangle was placed on the chart, meaning that the point 3,3 is strictly preferred to the point 2,3.

These two choices illustrate *insatiability*. This means that the consumer cannot be sated (or filled up) so more is always better. The combination 4,3 is preferred to 3,3, which is preferred to 2,3 because good x_2 is held constant at 3 and this consumer is insatiable, preferring more of good x_1 to less.

To reveal the indifference curve of this consumer, we must offer tougher choices, where we give more of one good and less of the other.

STEP Click the [Ask?] button again. This time offer the consumer a choice between 3,3 and 4,2.

The consumer decided that 3,3 is better. This reveals important information about the consumer's preferences. At 3,3, the consumer likes one more unit of x_1 less than the loss of one unit of x_2.

STEP Click the Ask? button several times more to figure out where the consumer's break-even point is in terms of how much x_2 is needed to balance the gain from the additional unit of x_1. Offer 4,2.5 and then try taking away less of good 2, such as 2.7 or 2.9. Once you find the point where the amount of x_2 taken away exactly balances the gain in x_1 of one unit (from 3 to 4), you have located two points on a single indifference curve. If it is difficult to see the points on the chart, use the Zoom control to magnify the screen (say to 200%).

You should find that this consumer is indifferent between the bundles 3,3 and 4,2.9.

STEP Now click the 100 Random ? button.

One hundred pairwise comparisons are made between 3,3 and a random set of alternatives. It is easy to see that the consumer can compare each and every point on the chart to the benchmark bundle of 3,3 and judge each and every point as better, worse, or the same.

STEP Click the Indifference button to display the indifference curve that goes through the benchmark point (3,3), as shown in Figure 2.1. Your version will be similar, but not exactly the same as Figure 2.1 since the 100 dots are chosen randomly.

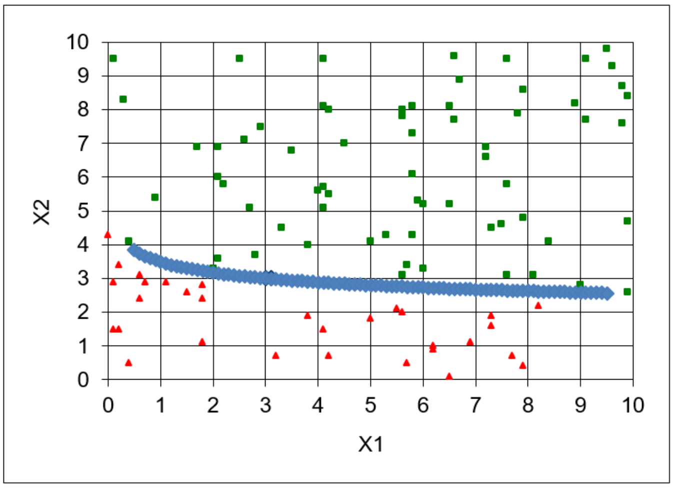


Figure 2.1: Revealing the indifference curve.
Source: Preferences.xls!Reveal

The indifference curve shows the bundles that are the same to this consumer compared to 3,3. All of the bundles for which the consumer is indifferent to the 3,3 bundle lie on the same indifference curve.

The Indifference Map

Every combination of goods has an indifference curve through it. We often display a few representative indifference curves on a chart and this is called an indifference map, as shown in Figure 2.2.

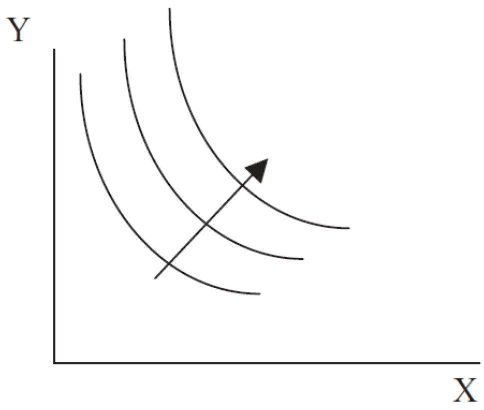


Figure 2.2: An indifference map.

Any point on the curve farthest from the origin, in Figure 2.2, is preferred to any point below it, including the ones on the two lower indifference curves. The arrow indicates that satisfaction increases as you move northeast to higher indifference curves.

There are many (in fact, an infinity) of indifference curves and they are not all depicted when we draw an indifference map. We draw just a few curves. We say that the indifference map is *dense*, which means there is a curve through every point.

STEP Build your own indifference map by copying the *Reveal* sheet and clicking the Reset button, then the Indifference button, and then the Copy Picture button.

This places a picture of the chart under the chart. This is an Excel drawing object, not a chart object, and it has no fill.

STEP Change the benchmark to 4,4 in cell B1 and click the Indifference button to get the indifference curve through the new benchmark point. Click the Copy Picture button.

This copies the chart and pastes the drawing object over the first one. Since it has no fill, it is transparent. You can separate the two pictures if you wish (click and drag), then undo the move so it is on top of the first picture.

STEP Add one more indifference curve to your map by changing the benchmark to 5,5 and clicking the Indifference button, then clicking the Copy Picture button.

You have created an indifference map with three representative indifference curves. Satisfaction increases as you move northeast to higher indifference curves.

Marginal Rate of Substitution

Having elicited a single indifference curve from the virtual consumer in the Excel workbook, we can define and work with a crucial concept in the Theory of Consumer Behavior: the *Marginal Rate of Substitution*, or MRS.

The MRS is a single number that tells us the willingness of a consumer to exchange one good for another from a given bundle. The MRS might be -18 or -0.07. Read carefully and work with Excel so that you learn what these numbers are telling you about the consumer's preferences.

STEP Return to the *Reveal* sheet (with benchmark point 3,3) and click the Copy Picture button to copy and paste an image of the current indifference curve below the graph in the *Reveal* sheet. Now click the New Preferences button to get a new virtual consumer with different preferences and then display the indifference curve for this new consumer (by clicking the Indifference button).

Notice that the indifference curve is not the same as the original one. These are two different consumers with different preferences. You can use the buttons to offer the new consumer bundles that can be compared with the 3,3 benchmark bundle, just like before.

The key idea here is that at 3,3, we can measure each consumer's willingness to trade x_2 in exchange for x_1.

Initially (as shown in Figure 2.1 and in the picture you took), we saw that the consumer was indifferent between 3,3 and 4,2.9. For one more unit of x_1 (from 3 to 4), the consumer is willing to trade 0.1 units of x_2 (from 3 to 2.9). Then the MRS of x_1 for x_2 from 3,3 to 4,2.9 is measured by $\frac{-0.1}{1}$, or -0.1.

With our new virtual consumer, the MRS at 3,3 is a different number. Let's compute it.

STEP Proceed to the MRS sheet. Click the Indifference button. Not only is the indifference curve through 3,3 displayed for this consumer, it also shows some of the bundles that lie on this indifference curve. We can use this information to compute the MRS.

You can compute the MRS at 3,3 by looking at the first bundle after 3,3. How much x_2 is the consumer willing to give up in order to get 0.1 more of x_1? This ratio, $\frac{\Delta x_2}{\Delta x_1}$, (the usual "rise over the run" definition of the slope), is the slope of the indifference curve, which is also the MRS.

The MRS also can be computed as the slope of the indifference curve *at a point* by using derivatives. Instead of computing $\frac{\Delta x_2}{\Delta x_1}$ along an indifference curve from one point to another, one can find the instantaneous rate of change at 3,3. We will do this later.

The crucial concept right now is that the MRS is a number that measures the willingness of a consumer to trade one good for another *at a specific point*. We usually think of it in terms of giving up some of the good on the y axis to get more of the good on the x axis.

Do not fall into the trap of thinking of the MRS as applying to the entire indifference curve. In fact, the MRS is different at each point on the curve. For a typical indifference curve like in Figure 2.1, the MRS gets smaller (in absolute value) as we move down the curve (as it flattens out).

The MRS is *negative* because the indifference curve is sloping downwards: a *decrease* in x_2 is compensated for by an *increase* in x_1. We often drop the minus sign because comparing negative numbers can be confusing. For example, say one consumer has an MRS of -1 at 3,3 while another has an MRS of $-\frac{1}{3}$ at that point. It is true that -1 is a smaller number than $-\frac{1}{3}$, however, we to use the MRS to indicate the steepness of the slope. Thus, to avoid confusion, we make the comparison using the absolute value of the MRS. Figure 2.3 shows that the bigger in absolute value is the MRS, the

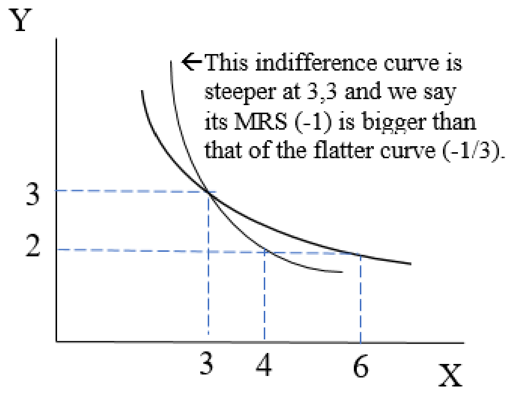


Figure 2.3: Comparing MRS.

more the consumer is willing to trade the good on the y axis for the good on the x axis. Thus, an MRS of -1 at 3,3 means the indifference curve has a steeper slope at that point than if the MRS was $-\frac{1}{3}$. We would say the MRS is bigger at -1 than $-\frac{1}{3}$ even though -1 is a smaller number than $-\frac{1}{3}$ because we look only at the absolute value of the MRS.

Funky Preferences and Their Indifference Curves

We can depict a wide variety of preferences with indifference maps. Here are some examples.

Example 1: Perfect Substitutes — constant slope (MRS)

If the consumer perceives two things as perfectly substitutable, it means they can get the same satisfaction by replacing one with the other.

Consider having one five-dollar bill and five one-dollar bills (as long as we are not talking about several hundred dollars worth of bills). If the consumer does not care about having $10 as a single ten-dollar bill, one five-dollar bill and five one-dollar bills, or ten one-dollar bills, then the indifference curve is a straight line as shown in Figure 2.4. You could argue that there is an indivisibility here and there are actually just 3 points that should not be connected by a line, but the key idea is that the indifference curve is a straight

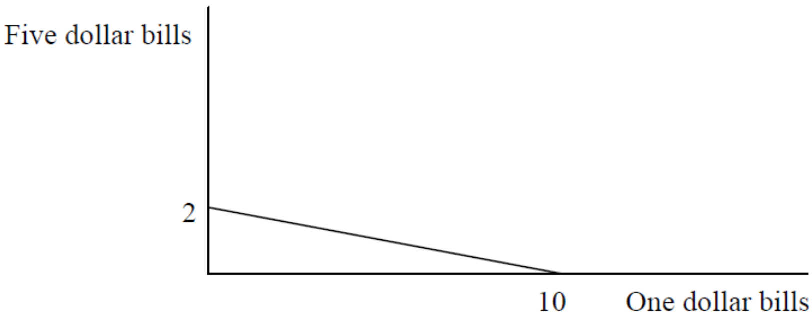


Figure 2.4: Perfect substitutes.

line in the case of perfect substitutes. It has a constant MRS (the slope of the line is $-\frac{1}{5}$), unlike a typical indifference curve where the MRS falls (in absolute value) as you move down the curve.

Example 2: Perfect Complements — L-shaped Indifference Curves

The polar opposite of perfect substitutes are perfect complements. Suppose the goods in questions have to be used in a particular way, with no room for any flexibility at all, like cars and tires. You need four tires for a car to work. With only three tires the car is worthless. Ignoring the spare, having more than four tires does not help you if you still have just one car.

Figure 2.5 illustrates the indifference map for this situation. It says that eight tires with one car gives the same satisfaction as four tires with one car. It also says that eight tires and two cars is preferred to four tires and one car (or eight tires and one car) because the middle L-shaped indifference curve (I_1) is farther from the origin than the lowest indifference curve (I_0).

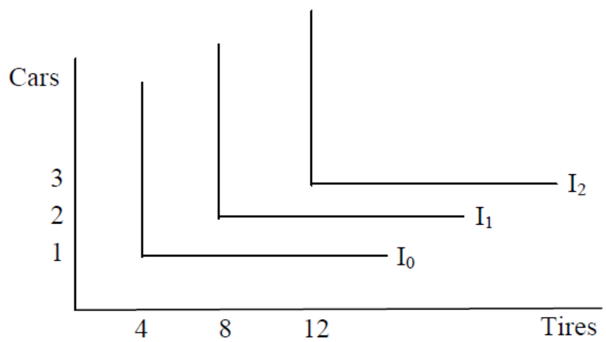


Figure 2.5: Perfect complements.

Notice how the usual indifference curve lies between the two extremes of perfect substitutes (straight lines) and perfect complements (L-shaped). Thus, the typical indifference curve reflects a level of substitutability between goods that is more than perfect complements (one good cannot replace another at all), but less than perfect substitutes (one good can take the place of another with no loss of satisfaction).

Example 3: Bads

What if one of the goods is actually a *bad*, something that lowers satisfaction as you consume more of it, like pollution? Figure 2.6 shows the indifference map in this case.

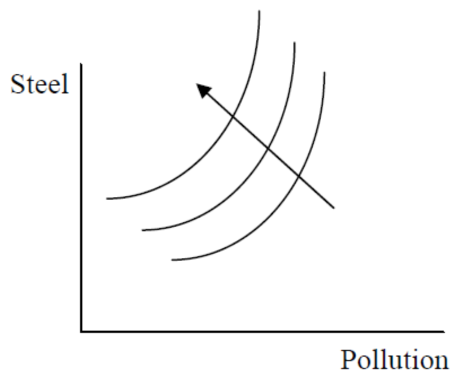


Figure 2.6: Bads.

Along any one of the indifference curves, more steel and more pollution are equally satisfying because pollution is a bad that cancels out the additional good from steel. The arrow indicates that satisfaction increases by moving northwest, to higher indifference curves.

Example 4: Neutral Goods

What if the consumer thinks something is neither good nor bad? Then it is a *neutral good* and the indifference map looks like Figure 2.7.

The horizontal indifference curves for the neutral good on the x axis in Figure 2.7 tell you that the consumer is indifferent if offered more X. The arrow indicates that satisfaction rises as you move north (because Y is a good and having more of it increasing satisfaction).

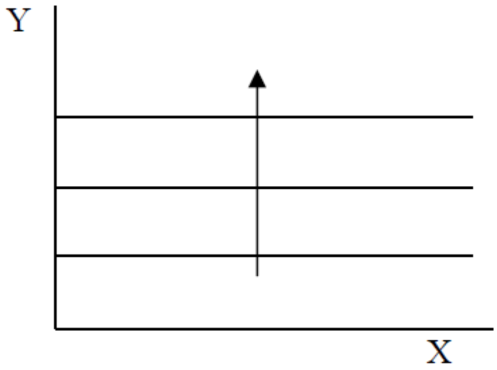


Figure 2.7: X is a neutral good.

These are just a few examples of how a variety of preferences can be depicted with an indifference map. When we want to describe generic, typical preferences that produce downward sloping indifference curves, as in Figure 2.2, economists use the phrase "well-behaved preferences."

Another technical term that is often used in economics is *convexity*, as in convex preferences. This means that midpoints are preferred to extremes. In Figure 2.8, there are two extreme points, A and B, which are connected by a dashed line. Any point on the dashed line, like C, can be described by the equation $zA + (1 - z)B$, where $0 < z < 1$ controls the position of C. This equation is called a convex combination.

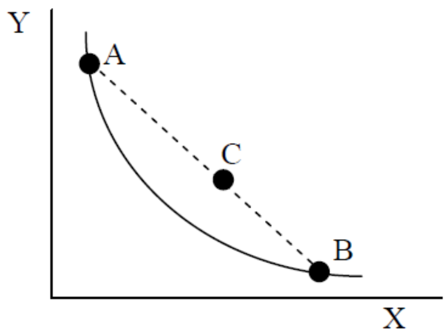


Figure 2.8: Convex preferences.

If preferences are convex, then midpoints like C are strictly preferred to extreme points like A and B. Convexity is used as another way of saying that preferences are well-behaved.

An important property that arises out of well-behaved or convex preferences is that of *diminishing MRS*. As explained earlier, the MRS varies along an indifference curve and applies to a specific point (not to the entire curve). The MRS will start large (in absolute value) at the top left corner, like point A in Figure 2.8, and get smaller as we travel down the indifference curve to point B. This makes common sense. The consumer is readily willing to trade a lot of Y for X (so the MRS is high in absolute value) when he has a lot of Y and little X. When the amounts are reversed, such as point B, a small MRS means he is willing to give up very little Y (since he has little of it) for more X (which he has a lot of already).

Indifference Curves Reflect Preferences

Preferences, a consumer's likes and dislikes, can be elicited or revealed by asking the consumer to pick between pairs of bundles. The indifference curve is that set of bundles that the consumer finds equally satisfying.

The MRS is a single number that measures the willingness of the consumer to exchange one good for another at a particular point. If the MRS is high (in absolute value), the indifference curve is steep at that point and the consumer is willing trade a lot of Y for a little more X.

Standard, well-behaved preferences yield a set of smooth arcs (like Figure 2.2), but there are many other shapes that depict preferences for different kinds of goods and the relationship between goods.

Exercises

1. What is the MRS at any point if X is a neutral good? Explain why.
2. If the good on the y axis was a neutral good and the other good was a regular good, then what would the indifference map look like. Use Word's Drawing Tools to draw a graph of this situation.
3. If preferences are well-behaved, then indifference curves cannot cross. Use Figure 2.9 to help you construct an explanation for why this claim must be true. Note that point C has more X and Y than point A, thus, by insatiability, C must be preferred to A. The key to defending the claim lies in the assumption of transitivity.

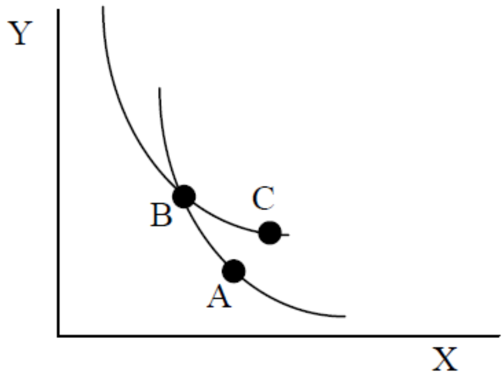


Figure 2.9: An impossible indifference map.

4. Suppose we measure consumer A's and B's MRS at the same point and find that $MRS_A = -6$ and the $MRS_B = -2$. What can we say about the preferences of A and B at this point?

References

The epigraph is from page 26 of C. E. Ferguson's *Microeconomic Theory* (revised edition, 1969), a popular micro text in the 1960s and 1970s. In the preface Ferguson wrote, "This is a textbook; its content is taken from the public domain of economic literature. Conventional topics are treated in conventional ways; and there is no real innovation." Perhaps, but Ferguson adopted a much more mathematical presentation and added content, including general equilibrium theory, that made his book different and important.

[A] cardinal measure of utility is in any case unnecessary; only an ordinal preference, involving "more" or "less" but not "how much," is required for the analysis of consumer's behavior.

Paul A. Samuelson

2.2 Utility Functions

Previously, we showed that a consumer has preferences that can be revealed and mapped. The next step is to identify a particular functional form, called a *utility function*, which faithfully represents the person's preferences. Once you understand how the utility function works, we can combine it with the budget constraint to solve the consumer's optimization problem.

Cardinal and Ordinal Rankings

Jeremy Bentham (1748-1832) was a utilitarian philosopher who believed that, in theory, the amount of utility from consuming a particular amount of a good could be measured. So, for example, as you ate an apple, we could hook you up to some device that would report the number of "utils" of satisfaction received. The word *utils* is in quotation marks because they do not actually exist, but Bentham believed they did and would one day be discovered with an advanced measuring instrument. This last part is not so crazy—an fMRI machine is exactly what he envisioned.

Bentham also believed that utils were a sort of common currency that enabled them to be compared across individuals. He thought society should maximize aggregate or total utility and utilitarianism has come to be associated with the phrase "the greatest happiness for the greatest number." Thus, if I get 12 utils from consuming an apple and you get 6, then I should get the apple. Utilitarianism also implies that if I get more utils from punching you in the face than you lose, I should punch you. This is why utilitarianism is not highly regarded today.

This view of utility treats satisfaction as if we could place it on a cardinal scale. This is the usual number line where 8 is twice as much as 4 and the difference between 33 and 30 is the same as that between 210 and 207.

Near the turn of the 20th century, Vilfredo Pareto (1848-1923, pronounced pa-RAY-toe) created the modern way of thinking about utility. He held that satisfaction could not be placed on a cardinal scale and that you could never compare the utilities of two people. Instead, he argued that utility could be measured only up to an ordinal scale, in which there is higher and lower, but no way to measure the magnitude between two items.

Notice how Pareto's approach matches exactly the way we assumed that a consumer could choose between bundles of goods as preferring one bundle or being indifferent. We never claimed to be able to measure a certain amount of satisfaction from a particular bundle.

For Pareto, and modern economics, the numerical value from a particular utility function for a given combination of goods has no meaning. These values are like the star ranking system for restaurants.

Suppose Critic A uses a 10-point scale, while Critic B uses a 1000-point scale to judge the same restaurants. We would never say that B's worst restaurant, which scored say 114, is better than A's best, a perfect 10. Instead, we compare their rankings. If A and B give the same restaurant the highest ranking (regardless of the score), it is the best restaurant.

Now suppose we are reading a magazine that uses a 5-star rating system. Restaurant X earns 4 stars and Restaurant Y 2 stars. X is better, but can we conclude that X is twice as good as Y? Absolutely not. An ordinal scale is ordered, but the differences between values are not important.

Pareto revolutionized our understanding of utility. He rejected Bentham's cardinal scale because he did not believe that satisfaction could be measured like body temperature or blood pressure. Pareto showed that we could derive demand curves with the less restrictive more-or-less ranking of bundles.

The transition from Bentham's cardinal view of utility to Pareto's ordinal view was not easy. Using the same word, utility, creates confusion (although, to be fair, Pareto tried to create a new word, *ophelimity*, but it never caught on). It bears repeating that, for a modern economist, although a utility function will show numerical values, these should not be interpreted on a cardinal scale, nor should numerical utilities of different people be compared. Since we cannot make interpersonal utility comparisons to add utilities of different people, we cannot give me the apple or let me punch you.

Monotonic Transformation

Once we reveal the consumer's indifference curve and map, we have the consumer's rankings of all possible bundles. Then, all we need to do is use a function that faithfully represents the indifference curves. The utility function is a convenient way to capture the consumer's ordering.

There are many (in fact, an infinity) of functions that could work. All the function has to do is preserve the consumer's preference ranking.

A *monotonic transformation* is a rule applied to a function that changes (transforms) it, but maintains the original order of the outputs of the function for given inputs. Monotonic is a technical term that means always moving in the same direction.

For example, star ratings can be squared and the rankings remain the same. If X is a 4-star and Y a 2-star restaurant, we can square them. X now has 16 stars and Y has 4 stars. X is still higher ranked than Y. In this case, squaring is a monotonic transformation because it has preserved the ordering and X is still higher than Y.

Can we conclude that X is now four times better? Of course not. Remember that the star ranking is an ordinal scale so the distance between items is irrelevant. We say that squaring is a monotonic transformation because it maintains the same ordering and we do not care about the distances between the numeric values. Their only meaning is "higher" and "lower," which indicate better and worse.

It is a fact that the MRS (at any point) remains constant under any monotonic transformation. This is an important property of monotonic transformations that we will illustrate with a concrete example in Excel.

Cobb-Douglas: A Ubiquitous Functional Form

STEP Open the Excel workbook *Utility.xls*, read the *Intro* sheet, and then go to the *CobbDouglas* sheet to see an example of this utility function:

$$u(x_1, x_2) = x_1^c x_2^d$$

In economics, a function created by multiplying variables that are raised to powers is called a *Cobb-Douglas functional form*.

STEP Follow the directions on the sheet (in column K) to rotate the 2D chart so you are looking down at it.

A top-down view of the utility function looks like an indifference map. The utility function itself, in 3D, is a hill or mountain (that keeps growing without ever reaching a top—illustrating the idea of insatiability).

With a utility function, the indifference curves appear as contour lines or level curves. The curves in 2D space are created by taking horizontal slices of the 3D surface. Every point on the indifference curve has the exact same height, which is utility.

STEP The exponents (c and d) in the utility function express "likes and dislikes." Try $c = 4$ then $c = 0.2$ in cell B5.

The higher the c exponent, the more the consumer likes x_1 because each unit of x_1 is raised to a higher power as c increases. Notice that when $c = 4$, the fact that the consumer likes x_1 much more than when $c = 0.2$ is reflected in the shape of the indifference curve. The steeper the indifference curve, the higher the MRS (in absolute value) and the more the consumer likes x_1.

STEP Proceed to the *CobbDouglasLN* sheet, which applies a monotonic transformation of the Cobb-Douglas function. It applies the natural log function to the utility function.

Recall that the natural logarithm of a number x is the exponent on e (the irrational number 2.7128 . . .) that makes the result equal x. You should also remember that there are special rules for working with logs. Two especially common rules are $\ln(x^y) = y \ln x$ and $\ln(xy) = \ln x + \ln y$. We can apply these rules to the Cobb-Douglas function when we take the natural log:

$$u(x_1, x_2) = x_1^c x_2^d$$

$$\ln[u(x_1, x_2)] = \ln[x_1^c x_2^d]$$

$$\ln[u(x_1, x_2)] = c \ln x_1 + d \ln x_2$$

The *CobbDouglasLN* sheet applies the natural log transformation by using Excel's LN() function.

STEP Click on any cell between B12 and Q27 to see the formula. We are computing the natural log of utility, which is x_1 raised to the c power times x_2 raised to the d power.

How does the original utility function compare to its natural log version?

STEP Go back and forth a few times between the two (click on the *CobbDouglas* sheet tab and then the *CobbDouglassLN* sheet tab). It is obvious that the numbers are different.

But did you notice something curious?

STEP Compare the cells with yellow backgrounds in the two sheets to see that these two combinations continue to lie on the same indifference curve, even though the utility values of the two functions are different.

The fact that the cells remain on the same indifference curve after undergoing the natural log transformation demonstrates the meaning of a monotonic transformation. The utility values are different, but the ranking has been preserved. The two utility functions both maintain the same relationship between 1,14 and 2,7 and every other bundle.

So now you know that a Cobb-Douglas utility function can be used to faithfully represent a consumer's preferences (including tweaking the c and d exponents to make the curves steeper or flatter) and that we can use the natural log transformation if we wish. In addition, economists often use the Cobb-Douglas functional form for utility (and production) functions because it has very nice algebraic properties where lots of terms cancel out.

The Cobb-Douglas function is especially easy to work with if you remember the following rules:

Algebra Rules: $\frac{x^a}{x^b} = x^{a-b}$ and $x^{a^b} = x^{ab}$

Calculus Rule: $\frac{dax^b}{dx} = bax^{b-1}dx$

These rules may seem irrelevant right now, but we will see that they make the Cobb-Douglas function much easier to work with than other functions. This goes a long way in explaining the repeated use of the Cobb-Douglas functional form in economics.

Expressing Other Preferences with Utility Functions

STEP Proceed to the *PerfSub* sheet and look around. Scroll down (if needed) and look at the two charts.
Notice how this functional form is producing straight line indifference curves (in the 2D chart). If the consumer treated two goods as perfect substitutes, we would use this functional form instead of Cobb-Douglas. The coefficients (a and b) can be tweaked to make the lines steeper or flatter.

STEP Proceed to the *PerfComp* sheet. This shows how the *min()* functional form produces L-shaped indifference curves.

The *min()* function outputs the smaller of the two terms, ax_1 and bx_2. This means that getting more of one good while holding the amount of the other good constant does not increase utility. This produces an L-shaped indifference curve.

Finally, the *Quasilinear* sheet displays indifference curves that are actually curved, but rather flat.

STEP Go to the *Quasilinear* sheet and click on the different functional form options. These are just a few of the many transformations that can be applied to x_1 and then added to x_2 to produce what is called quasilinear utility. Later, we will see that this functional form has different properties than Cobb-Douglas.

Note that we can represent many different kinds of preferences with utility functions. An important point is that there are many (to be more exact, an infinity) of possible utility functions available to us. We would choose one that faithfully reflects a particular consumer's preferences. We can always apply a monotonic transformation and it will not alter the consumer's preferences.

Computing the MRS for a Utility Function

Now that we have utility functions to represent a consumer's preferences, we are able to compute the MRS from one point to another (like we did in the previous chapter) or by using the instantaneous rate of change, better known as the derivative.

This is not a mathematics book, but economists use math so we need to see exactly how the derivative works. The core idea is convergence: make the change in x (the run) smaller and smaller and the ratio of the rise over the run (the slope) gets closer and closer to its ultimate value. The derivative is a shortcut that gives us the answer without the cumbersome process of making the change smaller and smaller.

But this is way too abstract. We can see it in Excel.

STEP Proceed to the *MRS* sheet to see how the MRS can be computed via a discrete-size change versus an infinitesimally-small change.

The utility function is $x_1 x_2$. This is Cobb-Douglas with exponents (implicitly) equal to 1.

Suppose we are interested in the indifference curve that gives all combinations with a utility of 10. Certainly 5,2 works (since 5 times 2 is 10). It is the red dot in the graph on the *MRS* sheet (and in Figure 2.10).

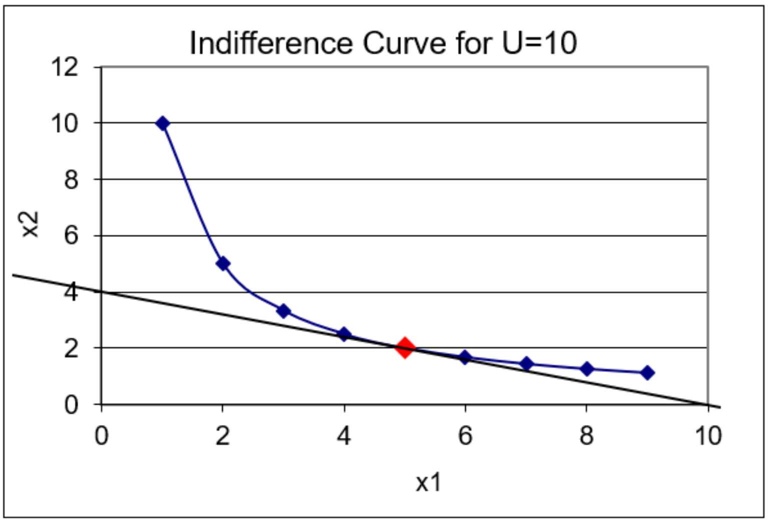


Figure 2.10: Computing the MRS.
Source: Utility.xls!MRS

From the bundle 5,2, if we gave this consumer 1 more unit of x_1, by how much would we have to decrease x_2 to stay on the $U = 10$ indifference curve? A little algebra tells us.

We know that $U = x_1x_2$ and the initial bundle 5,2 yields $U = 10$. We want to maintain U constant with $x_1 = 6$ because we added one unit to x_1, so we have:

$$U = x_1x_2$$
$$10 = 6x_2$$
$$x_2 = \frac{10}{6}$$

We have two bundles that yield $U = 10$ (5,2, and $6, \frac{10}{6}$). We can compute the MRS as the change in x_2 divided by the change in x_1. The delta (or difference) in x_2 is $-\frac{1}{3}$ (because $\frac{10}{6}$ is $\frac{1}{3}$ less than 2) and the delta in x_1 is 1 (6 - 5), so starting *from* the point 5,2, the MRS from $x_1 = 5$ to $x_1 = 6$ is $-\frac{1}{3}$. This is what Excel shows in cell C18.

Another way to compute the MRS uses the calculus approach. Instead of a "large" or discrete-size change in x_1, we take an infinitesimally small change, computing the slope of the indifference curve not from one point to another, but as the slope of the tangent line (as shown in Figure 2.10). We use the derivative to compute the MRS *at* a particular point.

For this simple utility function, holding U constant at 10, we can rewrite the function as x_2 in terms of x_1, then take the derivative.

$$U = x_1x_2$$

$$x_2 = \frac{10}{x_1}$$

$$\frac{dx_2}{dx_1} = -\frac{10}{x_1^2}$$

At $x_1 = 5$, substitute in this value and the MRS at that point is $-\frac{10}{25}$ or -0.4. This is what Excel shows in cell D18. If you need help with derivatives, the next chapter has an appendix that reviews basic calculus.

Computing the MRS this way relies on the ability to write x_2 in terms of x_1. If we have a utility function that cannot be easily rearranged in this way, we will not be able to compute the MRS. There is, however, a more general approach. The procedure involves taking the derivative of the utility function with respect to x_1 (called the marginal utility of x_1) and dividing by the derivative of the utility function with respect to x_2 (called the marginal utility of x_2). Do not forget to include the minus sign when you use this approach. Here is how it works.

With $U = x_1x_2$, the derivatives are simple: $\frac{dU}{dx_1} = x_2$ and $\frac{dU}{dx_2} = x_1$. Thus, we can substitute these into the numerator and denominator of the MRS expression:

$$MRS = -\frac{\frac{dU}{dx_1}}{\frac{dU}{dx_2}} = -\frac{x_2}{x_1}$$

Because we are considering the point 5,2, we evaluate the MRS at that point (which means we plug in those values to our MRS expression), like this:

$$MRS = -\frac{x_2}{x_1}\Bigg|_{\substack{x_1=5\\x_2=2}} = -\frac{2}{5} = -0.4$$

Note that minus the ratio of the marginal utilities gives the same answer as the $\frac{dx_2}{dx_1}$ method. Both are using infinitesimally small changes to compute the instantaneous rate of change of the indifference curve at a particular point.

Also note that the ratio of the marginal utilities approach requires that you divide the marginal utility of x_1 (the good on the x axis) by the marginal utility of x_2 (the good on the y axis). Since we used $\frac{\Delta y}{\Delta x}$ in the discrete-size change approach, it is easy to confuse the numerator and denominator when computing the MRS via the derivative. Remember that $\frac{dU}{dx_1}$ goes in the numerator.

Comparing Δ and d Methods

So far, we know there are two ways to get the MRS: move from one point to another along the indifference curve (discrete change, Δ) or slope of the tangent line at a point (infinitesimally small change, d). We also know that we have two ways of doing the latter (solve for x_2 then take the derivative or compute the ratio of the marginal utilities.)

But you may have noticed a potential problem in that the two procedures to get the MRS yield different answers. In the *MRS* sheet and our work above, the discrete change approach tells us that the MRS as measured from $x_1 = 5$ to $x_1 = 6$ is $-\frac{1}{3}$, whereas the derivative method says that the MRS at $x_1 = 5$ is -0.4.

This difference in measured MRS is due to the fact that the two approaches are applying a different size change in x_1 to a curve. As the discrete-size change gets smaller, it approaches the derivative measure of the MRS. You can see this clearly with Excel.

STEP Change the step size in cell B7 to 0.5 and watch how cell C18 changes. Notice that the chart is also slightly different because the point at $x_1 = 6$ is now at 5.5.

You have made the size of the change in x_1 smaller so the point is now closer to the initial value, 5.

STEP Do it again, this time changing the step size in cell B7 to 0.1. The point with $x_1 = 5.1$ is so close to 5 that it is hard to see, but it is there. Do one last change to the step size, setting it at 0.01.

With the step size at 0.01, you cannot see the initial and new points because they are so close together, but they are still a discrete distance apart. Excel displays the point-to-point delta computation in cell C18. It is really close to the derivative measure of the MRS in cell D18 because the derivative is simply the culmination of this process of making the change in x_1 smaller and smaller.

In Figure 2.10, the discrete change approach is computing the rise over the run using two separate points on the curve, while the calculus approach is computing the slope of the tangent line.

STEP Look at the values of the cells in the yellow highlighted row.

The MRS for a given approach are exactly the same. In other words, columns C, H, and M are the same and columns D, I, and N are the same. This shows that the MRS remains unaffected when the utility function is monotonically transformed.

Utility Functions Represent Preferences

Utility functions are equations that represent a consumer's preferences. The idea is that we reveal preferences by having the consumer compare bundles, and then we select a functional form that faithfully reflects the indifference curves of the consumer.

In selecting the functional form, there are many possibilities and economists often use the Cobb-Douglas form. The values of utility produced by inputting amounts of goods are meaningless and any monotonic transformation (because it preserves the preference ordering) will work as a utility function. Monotonic transformations do not affect the MRS.

The MRS is an important concept in consumer theory. It tells us the willingness to trade one good for another and this measure the consumer's likes and dislikes. Willingness to trade a lot of y for a little x produces a high MRS (in absolute value) and this indicates that the consumer values x more than y.

The MRS computed *from* one point to another (Δ), but it can also be computed using the derivative (d) *at* a point. Both are valid and the resulting number for the MRS is interpreted the same way (willingness to trade).

Exercises

The utility function, $U = x - 0.03x^2 + y$, has a quasilinear functional form. Use this function to the answer the questions below. You can see what it looks like by choosing the Polynomial option in the *Quasilinear* sheet.

1. Compute the value of the utility function at bundle A, where $x = 10$ and $y = 1$. Show your work.
2. Working with bundle A, find the MRS as x rises from $x = 10$ to $x = 20$. Show your work.
3. Find the MRS at the point 10,1 (using derivatives). Show your work.
4. Why do the two methods of determining the MRS yield different answers?
5. Which method is better? Why?

References

The epigraph can be found on page 91 of the revised edition of *The Foundations of Economic Analysis*, by Paul Samuelson. This remarkable book, written by one of the greatest economists of the 20$^{\text{th}}$ century, took economics to a new level of mathematical sophistication. Samuelson could not have picked a better opening quote, "Mathematics is a Language," by J. Willard Gibbs.

Chapter 3

Optimal Choice

> *Joseph Louis Lagrange*, the greatest mathematician of the eighteenth century, was born at Turin on January 25, 1736, and died at Paris on April 10, 1813. . . . In appearance he was of medium height, and slightly formed, with pale blue eyes and a colourless complexion. In character he was nervous and timid, he detested controversy, and to avoid it willingly allowed others to take credit for what he had himself done.
>
> W. W. Rouse Ball

3.1 Initial Solution

What you know so far:

1. The *budget constraint* shows the consumer's possible consumption bundles. The standard, linear constraint is $p_1x_1 + p_2x_2 = m$. There are many other situations, such as subsidies and rationing, which give more complicated constraints with kinks and horizontal/vertical segments.

2. The *indifference map* shows the consumer's preferences. The standard situation is a set of convex, downward sloping indifference curves. There are many alternative preferences, such as perfect substitutes and perfect complements. Preferences are captured by utility functions, which accurately reflect the shape of the indifference curves.

Our job is to combine these two parts, one expressing what is affordable and the other what is desirable, to find the combination (or bundle) that maximizes satisfaction (as described by the indifference map or utility function) given the budget constraint. The answer will be in terms of how much the consumer will buy in units of each good.

The optimal solution is depicted by the canonical graph in Figure 3.1. The word *canonical* is used here to mean standard, conventional, or orthodox. In economics, a canonical graph is a core, essential graph that is understood by all economists, such as a supply and demand graph.

It is no exaggeration to say that Figure 3.1 is one of the most fundamental and important graphs in economics. It is the foundation of the Theory of Consumer Behavior and with it we will derive a demand curve.

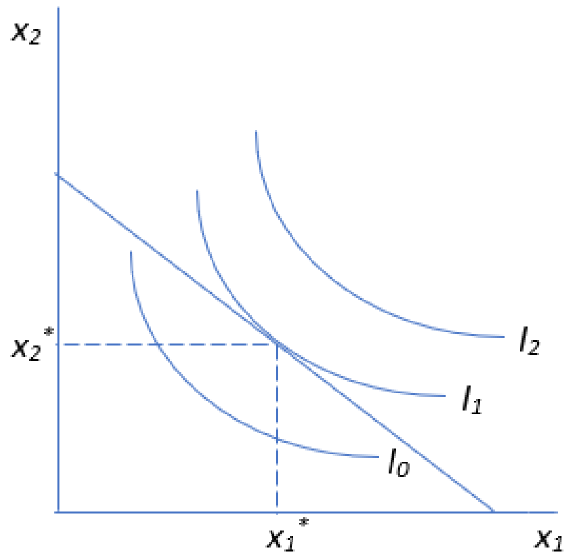


Figure 3.1: The canonical graph of the optimal solution.

One serious intellectual obstacle with Figure 3.1 is that it is highly abstract. Below we work on a concrete problem, with actual numbers, to explain what is going on in this fundamental graph.

Before we dive in, we need to discuss solution strategies. There are two ways to find the optimal solution:

1. Analytical methods using algebra and calculus—this is the conventional, paper and pencil approach that has been used for a long time.

2. Numerical methods using a computer, for example, Excel's Solver—this is a modern solution strategy that uses the computer to do most of the work.

Analytical Approach

Unfortunately, constrained optimization problems are harder to solve than unconstrained problems. The appendix to this chapter offers a short calculus review along with a few common derivative and algebra rules. If the material below makes little sense, go to the appendix and then return here.

Because this is a constrained optimization problem, the analytical approach uses the method developed by Joseph Louis Lagrange. His brilliant idea is based on transforming a constrained optimization problem into an unconstrained problem and then solving by using standard calculus techniques. In the process, a new endogenous variable is created. It can have a meaningful economic interpretation.

Lagrange gave us a recipe to follow that requires four steps:

1. Rewrite the constraint so that it is equal to zero.
2. Form the Lagrangean function.
3. Take partial derivatives with respect to x_1, x_2, and λ.
4. Set the derivatives equal to zero and solve for $x_1{}^*$, $x_2{}^*$, and λ^*.

A Concrete Example

Suppose a consumer has a Cobb-Douglas utility function with exponents both equal to 1 and a budget constraint, $2x_1 + 3x_2 = 100$ (which means the price of good 1 is \$2/unit, the price of good 2 is \$3/unit, and income is \$100).

The problem is to maximize utility subject to (s.t.) the budget constraint. It is written in equation form like this:

$$\max_{x_1,x_2} U(x_1,x_2) = x_1x_2$$
$$\text{s.t. } 100 = 2x_1 + 3x_2$$

This problem is not solved directly. It is first transformed into an unconstrained problem, and then this unconstrained problem is solved. Here is how we apply the recipe developed by Lagrange.

1. Rewrite the constraint so that it is equal to zero.

$$0 = 100 - 2x_1 - 3x_2$$

2. Form the Lagrangean function.

$$\max_{x_1,x_2,\lambda} L = x_1x_2 + \lambda(100 - 2x_1 - 3x_2)$$

Most math books use a fancy script L for the Lagrangean, like this $\mathcal{L}$, but this is difficult to do in Word's Equation Editor (which you will be using) so an extra-large L will work just as well. Also, many books spell Lagrangean with an i, Lagrangian, but both spellings are acceptable.

Note that the Lagrangean function, L, is composed of the original objective function (in this case, the utility function) plus a new variable, the Greek letter lambda, λ, times the rewritten constraint. Called the *Lagrangean multiplier*, λ is a new endogenous variable that is introduced as part of Lagrange's solution strategy.

The next step in Lagrange's recipe can be intimidating. This is not the time to rush through and turn the page. Refer to the appendix at the end of this section if things start to get confusing.

3. Take partial derivatives with respect to x_1, x_2, and λ.

$$\frac{\partial L}{\partial x_1} = x_2 - 2\lambda$$
$$\frac{\partial L}{\partial x_2} = x_1 - 3\lambda$$
$$\frac{\partial L}{\partial \lambda} = 100 - 2x_1 - 3x_2$$

The derivative used here is a partial derivative, denoted by ∂, which is an alternative way of writing a lowercase Greek letter d (which is why the more common symbol for the letter δ is also used). The partial derivative symbol is usually read as the letter d, so the first equation read out loud would be "d L d x one equals x two minus two times lambda." It is also common to read the derivative in the first equation as "partial L partial x one."

The partial derivative is a natural extension of the regular derivative. Consider the function $y = 4x^2$. The derivative of y with respect to x is $\frac{dy}{dx} = 8x$. Suppose, however, that we had a more complicated function, like this: $y = 4zx^2$. This multivariate function says that y depends on two variables, z and x. We can explore the rate of change of this function along the x axis by treating it as a partial function, meaning that we hold the z variable constant. Then the partial derivative of y with respect to x is $\partial y/\partial x = 8zx$. If we hold x constant and vary z, then the partial derivative of y with respect to z is $\partial y/\partial z = 4x^2$.

Applying this logic to the Lagrangean in step 2, when we take the partial derivative with respect to x_1, the first term is x_2 because it is as if we had "$x4$" and took the derivative with respect to x, getting 4.

If we multiply λ through the parenthetical expression in the Lagrangean, we get:

$$\lambda(100 - 2x_1 - 3x_2)$$
$$\lambda 100 - \lambda 2x_1 - \lambda 3x_2 = 0$$

The first and third terms on the left-hand side do not have x_1 so the derivative with respect to x_1 is zero (just like the derivative of a constant is zero). The derivative with respect to x_1 of the middle term produces $-\lambda 2$ which is written by convention as -2λ.

Can you do the other two derivatives in step 3?

4. Set the derivatives equal to zero and solve for x_1*, x_2*, and λ*.

$$\frac{\partial L}{\partial x_1} = x_2 - 2\lambda = 0$$
$$\frac{\partial L}{\partial x_2} = x_1 - 3\lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = 100 - 2x_1 - 3x_2 = 0$$

There are many ways to solve this system of equations, which are known as the first-order conditions. Sometimes, this is the hardest part of the Lagrangean method. Depending on the utility function and constraint, there may not be an analytical solution.

A common strategy involves moving the λ terms in the first two equations to the right-hand side and then dividing the first equation by the second one.

$$x_2 = 2\lambda$$
$$x_1 = 3\lambda$$
$$\frac{x_2}{x_1} = \frac{2\lambda}{3\lambda}$$

The λ terms then cancel out, leaving us with two equations (the one above and the third equation from the original three first-order conditions) and two unknowns (x_1 and x_2).

$$\frac{x_2}{x_1} = \frac{2}{3}$$
$$100 - 2x_1 - 3x_2 = 0$$

The top equation has a nice economic interpretation. It says that, at the optimal solution, the MRS (slope of the indifference curve) must equal the price ratio (slope of the budget constraint).

From the top equation, we can solve for x_2.

$$x_2 = \frac{2}{3}x_1$$

We can then substitute this expression into the bottom equation (the budget constraint) to get the optimal value of x_1.

$$100 - 2x_1 - 3\left[\frac{2}{3}x_1\right] = 0$$
$$100 - 2x_1 - 2x_1 = 0$$
$$100 = 4x_1$$
$$x_1^* = 25$$

Then we substitute x_1* into the expression for x_2 to get x_2*.

$$x_2 = \frac{2}{3}[25]$$

$$x_2{}^* = 16\frac{2}{3}$$

The asterisk is used to represent the optimal solution for a choice variable. This work says that this consumer should buy 25 units of good 1 and $16\frac{2}{3}$ units of good 2 in order to maximize satisfaction given the budget constraint. We can use either equation 1 or 2 from the original first-order conditions to find the optimal value of λ. Either way, we get $\lambda^* = 8\frac{1}{3}$.

For many optimization problems, we would be interested in knowing the numerical value of the maximum by evaluating the objective function (in this case the utility function) at the optimal solution. But recall that utility is measured only up to an ordinal scale and the actual value of utility is irrelevant. We want to maximize utility, but we do not care about its actual maximum value. The fact that utility is ordinal, not cardinal, also explains why the optimal value of lambda is not meaningful. In general, the Lagrangean multiplier tells us how the maximum value of the objective function changes as the constraint is relaxed. With utility as the objective function, this interpretation is not applicable.

Numerical Approach

Instead of calculus (via the method of Lagrange) and pencil and paper, we can use numerical methods to find the optimal solution.

To use the numerical approach, we need to do some preliminary work. We have to set up the problem in Excel, carefully organizing things into a goal, endogenous variables, exogenous variables, and constraint. Once we have everything organized, we can use Excel's Solver to get the solution.

STEP Open the Excel workbook *OptimalChoice.xls*, read the *Intro* sheet, and then go to the *OptimalChoice* sheet to see how the numerical approach can be used to solve the problem we worked on above.

Figure 3.2 reproduces the display you see when you first arrive at the *OptimalChoice* sheet.

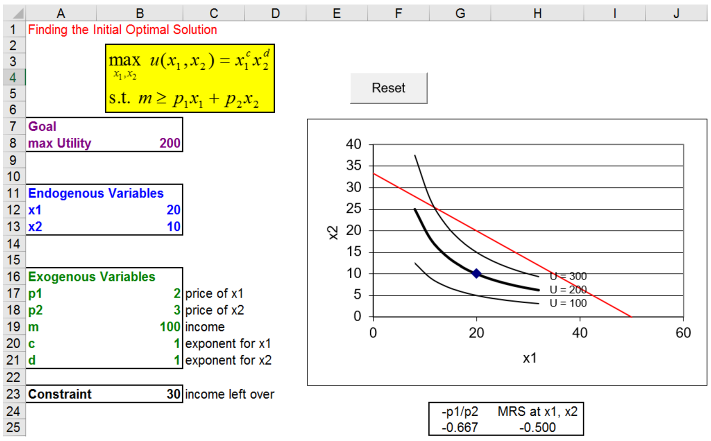


Figure 3.2: The initial display in the *OptimalChoice* sheet.
Source: OptimalChoice.xls!OptimalChoice

Notice how the sheet is organized according to the three components of the optimization problem: goal, endogenous, and exogenous variables. The constraint cell displays how much of the consumer's budget remains available for buying goods. The consumer in Figure 3.2 is not using all of the income available so we know satisfaction cannot be maximized at the point 20,10.

STEP Let's have the consumer buy x_2 with the remaining \$30. At \$3/unit, 10 additional units of x_2 can be purchased. Enter 20 in the x_2 cell (B13) and hit the Enter key. The chart refreshes to display the point 20,20, which is on the budget constraint, and draws three new indifference curves.

Although 20,20 does exhaust the available income, it is not the optimal solution. While you know the answer is 25,16$\frac{2}{3}$, there is another way to tell that the consumer can do better.

STEP Look carefully at the display below the chart. It reveals the MRS does not equal the price ratio. This immediately tells us that something is amiss here.

MRS $> p_1/p_2$ tells us that the slope of the indifference curve at that point is greater than the slope of the budget constraint. The consumer cannot change the slope of the budget constraint, but the MRS can be altered by choosing a different the combination of goods. This consumer needs to lower the MRS (in absolute value) to make the two equal. This can be done by moving down the budget constraint.

If the consumer buys 10 more of good 1 (so 30 units of x_1 total), consumption of x_2 must fall by $6\frac{2}{3}$ units to $13\frac{1}{3}$.

STEP Enter 30 in cell B12 and the formula $= 13 + 1/3$ in B13. Now you are on the other side of the optimal solution. The MRS is less than the price ratio.

You could, of course, continue adjusting the cells manually, but there is a faster way.

STEP Click the Data tab in Excel's Ribbon (on the top of the screen) and click Solver (grouped under the Analyze tab) or execute Tools: Solver in older versions of Excel to bring up the Solver Parameters dialog box (displayed in Figure 3.3).

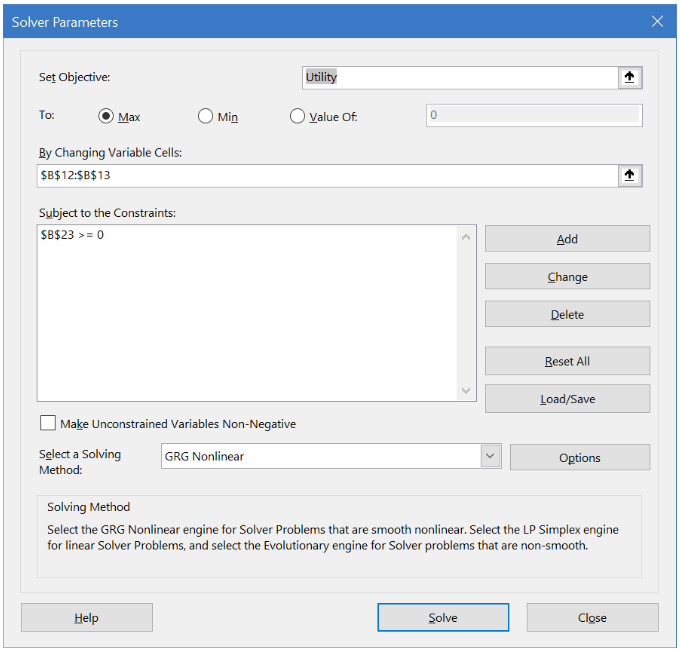


Figure 3.3: Excel's Solver interface.

If you do not have Solver available as a choice, bring up the Add-in Manager dialog box and make sure that Solver is listed and checked. If Solver is not listed, you must install it. Solver is included in a standard installation of Excel. For help, try support.office.com or www.solver.com.

Note how Excel's Solver includes information on the objective function (the target cell), the choice variables (the changing cells), and the budget constraint. These have all been filled in for you, but you will learn how to do this yourself in future work.

STEP Since all of the information has been entered into the Solver Parameters dialog box, simply click the Solve button at the bottom of the dialog box.

Excel's Solver works by trying different combinations of x_1 and x_2 and evaluating the improvement in the target cell, while trying to stay within the constraint. When it cannot improve very much more, it figures it has found the answer and displays a message as shown in Figure 3.4.

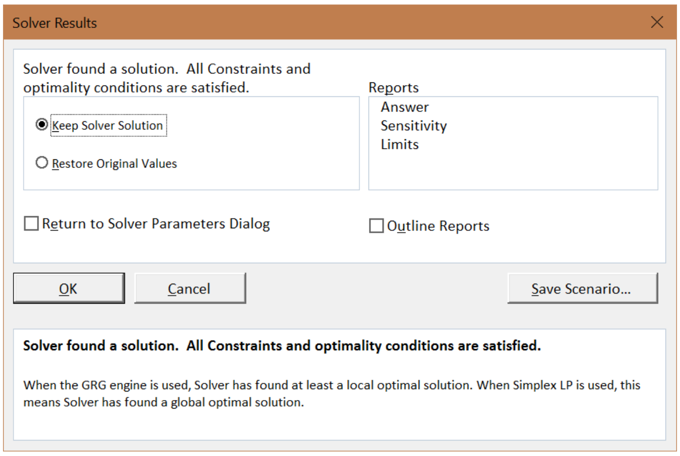


Figure 3.4: Solver reports success.

Although Solver gets the right answer in this problem, we will see in future applications that Solver is not perfect and does not deserve blind trust.

STEP Click the Sensitivity option under Reports and click OK; Excel puts the Solver solution into cells B12 and B13. It also inserts a new sheet into the workbook with the Sensitivity Report.

STEP Click on cells B12 and B13. Notice that Excel did not get exactly 25 and $16\frac{2}{3}$. It got extremely close and you can certainly interpret the result as confirming the analytical solution, but Solver's output requires interpretation and critical thinking by the user. We will focus on the issue of the exactly correct answer later.

STEP Proceed to the *Sensitivity Report* sheet (inserted by Solver) to confirm that this numerical method gives substantially the same absolute value for the Lagrangean multiplier that we found via the Lagrangean method ($8\frac{1}{3}$). We postpone explanation of this because utility's ordinal scale makes interpretation of the Lagrangean multiplier pointless. For now, we simply note that Solver can report optimal lambda and its results agreed with the Lagrangean method.

You might notice that Excel reports a Lagrangean multiplier value of -8.33 (with a few more trailing 3s) yet our analytical work did not produce a negative number. It turns out that we ignore the sign of λ^*. If we set up the Lagrangean as the objective function minus (instead of plus) lambda times the constraint or rewrite the constraint as $0 = 2x_1 + 3x_2 - 100$ (instead of

$0 = 100 - 2x_1 - 3x_2$), we would get a negative value for λ^* in our analytical work. The way we write the constraint or whether we add or subtract the constraint is arbitrary, so we ignore the sign of λ^*.

To be clear, unlike the sign, the magnitude of λ^* can be meaningful, but it is not in this application because utility is not cardinal. We will, however, see examples where the value of λ^* is useful and has an economic interpretation.

Using Analytical and Numerical Methods to Find the Optimal Solution

There are two ways to solve optimization problems:

1. The traditional way uses pencil and paper, derivatives, and algebra. The Lagrangean method is used to solve constrained optimization problems, such as the consumer's choice problem.

2. Advances in computers have led to the creation of numerical methods to solve optimization problems. Excel's Solver is an example of a numerical algorithm that can be used to find optimal solutions.

In the chapters that follow, we will continue to use both analytical and numerical approaches. You will see that neither method is perfect and both have strengths and weaknesses.

Exercises

The utility function, $U = 10x - 0.1x^2 + y$, has a quasilinear functional form. Use this utility function to answer the questions below.

1. Suppose the budget line is $100 = 2x + 3y$. Use the analytical method to find the optimal solution. Show your work.

2. Suppose the consumer considers the bundle 0,33.33, buying no x and spending all income on y. Use the MRS compared to the price ratio logic to explain what the consumer will do and why.

3. This utility function can be written in a more general form with letters instead of numbers, like this: $U = ax - bx^c + dy$. If a increases, what happens to the optimal consumption of x^*? Explain how you arrived at your answer.

References

The epigraph is from page 421 of W. W. Rouse Ball's *A Short Account of the History of Mathematics* (first published in 1888). Of course, there are many books on the history of mathematics, but this classic is fun and easy to read. It mixes stories about people with real mathematical content.

This entire book (and many others) is freely available at books.google.com. You can read it online or download it as a pdf file.

Appendix: Derivatives and Optimization

A *derivative* is a mathematical expression that tells you how y in a function $y = f(x)$ changes given an infinitesimally small change in x. Graphically, it is the slope, or rate of change, of the function at that particular value of x.

Linear functions have a constant slope and, therefore, a constant value for the derivative. For the linear function $y = 6 + 3x$, the derivative of y with respect to x is written $\frac{dy}{dx}$ (pronounced "d y d x") and its value is 3. This tells you that every time the x variable goes up, the y variable goes up threefold. So, if x increases by 1 unit, y will increase by 3 units. This is easy to see in Figure 3.5.

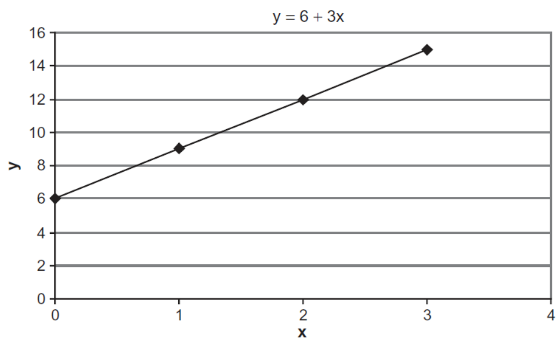


Figure 3.5: A linear function.

Nonlinear functions have a changing slope and, therefore, a derivative that takes on different values at different values of x. Consider the function $y = 4x - x^2$. Figure 3.6 graphs this function. Its derivative is $\frac{dy}{dx} = 4 - 2x$. When evaluated at a specific point, such as $x = 1$, the derivative is the slope of the tangent line at that point.

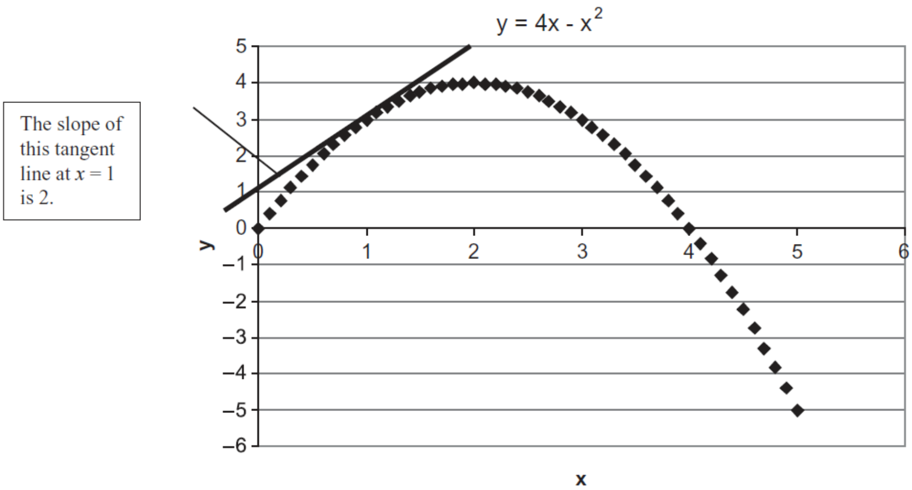


Figure 3.6: A nonlinear function with tangent line at $x = 1$.

Unlike the previous case, this derivative has x in it. This means this function is nonlinear. The slope depends on the value of x. At $x = 1$, the derivative is 2, but at $x = 2$, it is zero $(4 - 2[2])$ and at $x = 3$, it is -2 $(4 - 2[3])$.

In addition, because it is nonlinear, the size of the change in x affects the measured rate of change. For example, the change in y from $x = 1$ to $x = 2$ is 1 (because we move from $y = 3$ to $y = 4$ as we increase x by 1). If we increase x by a smaller amount, say 0.1 (from 1 to 1.1), then $\frac{\Delta y}{\Delta x} = \frac{3.19-3}{1.1-1} = 1.9$. By taking a smaller change in x, we get a different measure of the rate of change.

If we compute the rate of change via the derivative, by evaluating $4 - 2x$ at $x = 1$, we get 2. The derivative computes the rate of change for an infinitesimally small change in x. The smaller the change in x, the closer $\frac{\Delta y}{\Delta x}$ gets to $\frac{dy}{dx}$. You can see this happening as $\frac{\Delta y}{\Delta x}$ went from 1 to 1.9 as Δx fell from 1 to 0.1. If we go even smaller, making $\Delta x = 0.01$ (going from 1 to 1.01), then $\frac{\Delta y}{\Delta x} = \frac{3.0199-3}{1.01-1} = 1.99$.

Optimizing with the Derivative

An optimization problem typically requires you to find the value of an endogenous variable (or variables) that maximizes or minimizes a particular objective function. We can use derivatives to find the optimal solution. This is called an analytical approach.

If we draw tangent lines at each value of x in Figure 3.6, only one would be horizontal (with derivative and slope of zero) and that would be the one at the top. This gives us a solution strategy: to find the maximum, find the value of x with the flat tangent line. This is equivalent to finding the value of x where the derivative is zero.

By solving for the value of x where $\frac{dy}{dx} = 0$, we find the optimal solution. For $y = 4x - x^2$, this is easy. We set the derivative equal to zero and solve for $x*$.

$$\frac{dy}{dx} = 4 - 2x* = 0$$
$$4 = 2x*$$
$$x* = 2$$

The equation that you make when you set the first derivative equal to zero is called the *first-order condition.* The first-order condition is different from the derivative because the derivative by itself is not equal to anything—you can plug in any value of x and the derivative expression will pump out an answer that tells you whether and by how much the function is rising or falling at that point. The first-order condition is a special situation in which you are using the derivative to find a horizontal tangent line to figure out where the function has a flat spot.

A *reduced form* is the answer that you get when the derivative is set equal to zero and solved for the optimal solution. It may be a number or a function of exogenous variables. It cannot have any endogenous variables in the expression. Sometimes, you cannot solve explicitly for $x*$. We say there is no closed form solution in these cases. The solution may exist (and numerical methods may be used to find it), but we cannot express the answer as an equation.

The second derivative is the derivative of the first derivative. It tells you the slope of the slope function. For example, if a function has a constant slope, we saw that its first derivative is a constant value (like 3 in the first example above). Then the second derivative is zero.

Second derivatives are useful in optimization for the following reason: when you find the value of the endogenous variable that makes the first derivative equal to zero, the point that you have located could be either a maximum or a minimum. If you want to be sure which one you have found, you can check the second derivative. For $y = 4x - x^2$, the first derivative is $4 - 2x$ and the second derivative is, therefore, -2. Because the second derivative is negative, we know that our flat spot at $x = 2$ is a maximum and not a minimum.

In this book, we will not use second derivatives to check that our solutions are truly maxima or minima. Our functions will be (mostly) well behaved and we will focus on the economics of the problem, not the mathematics.

In summary, derivatives are used to measure the rate of change of a function based on a vanishingly small change in x. If we set a derivative equal to zero, we are trying to find an optimal solution by finding a value for x where the tangent line is flat. This solution strategy is based on the idea that a point where the tangent line is horizontal must mean that we are at the top of the function (or bottom, if we are minimizing).

Useful Math Facts

This appendix concludes with a short list of common rules for taking derivatives and working with exponents. The idea here is to sharpen your math skills so you can solve optimization problems analytically.

A derivative can be computed by directly applying the definition—i.e., taking the limit of the change in x as it approaches zero and determining the change in y. Fortunately, however, there is an easier way. Differentiation rules have been developed that make it much less tedious to take a derivative. Most calculus books have inside covers that are full of rules. Many students never grasp that these rules are actually shortcuts. Here is a short list, with special emphasis on those used in economics.

The derivative rules are followed by a few algebra rules relating to legal operations on exponents. We will use these rules often to find optimal solutions and reduce complicated expressions to simpler final answers.

Reading these equations is boring and tedious, but may save a lot of time and effort in the future (especially if your math is rusty). You should consider writing out the examples for a different number, say 6. So, instead of x^4, what is the derivative with respect to x for x^6?

Derivative Rules

Let x be the variable and a be a constant.

General Rule	Example of its Application
$\frac{d}{dx}(x) = 1$	
$\frac{d}{dx}(ax) = a$	$\frac{d}{dx}(4x) = 4$
$\frac{d}{dx}(a) = 0$	$\frac{d}{dx}(4) = 0$
$\frac{d}{dx}(x^a) = ax^{a-1}$	$\frac{d}{dx}(x^4) = 4x^3$
$\frac{d}{dx}(a \ln x) = \frac{a}{x}$	$\frac{d}{dx}(4 \ln x) = \frac{4}{x}$

When you take a derivative of a function with respect to a variable, you apply the rules to the different parts of the function. For example, if $y = 4x - x^2$, then you apply the $\frac{d}{dx}(ax) = a$ rule to $4x$, getting 4. You apply the $\frac{d}{dx}(x^a) = ax^{a-1}$ rule to $-x^2$ and get $-2x$. Thus, the derivative of y with respect to x is $\frac{dy}{dx} = 4 - 2x$.

There are other calculus rules, of course, such as the chain rule, but we will explain them when they are needed.

Laws of Exponents

General Rule	Example of its Application
$x^0 = 1$	
$x^{-a} = \frac{1}{x^a}$	$x^{-\frac{1}{2}} = \frac{1}{\sqrt{x}}$
$x^a x^b = x^{a+b}$	$x^2 x^3 = x^5 \Rightarrow 2^2 2^3 = 2^5 = 32$
$\frac{x^a}{x^b} = x^{a-b}$	$\frac{x^5}{x^3} = x^2 \Rightarrow \frac{2^5}{2^3} = 2^2 = 4$
$(xy)^a = x^a y^a$	$(xy)^2 = x^2 y^2 \Rightarrow (2 \cdot 3)^2 = 2^2 3^2 = 36$
$(x^a)^b = x^{ab}$	$(x^2)^3 = x^6 \Rightarrow (2^2)^3 = 2^6 = 64$

> The methods of mathematics apply as soon as spatial or numerical attributes are associated with our phenomena, as soon as objects can be located by points in space and events described by properties capable of indication or measurement in numbers.
>
> R. G. D. Allen

3.2 More Practice and Understanding Solver

We know there are two approaches to solving optimization problems.

1. Analytical methods using algebra and calculus (conventional, paper and pencil, using the Lagrangean method): The idea is to transform the consumer's constrained optimization problem into an unconstrained problem and then solve it using standard unconstrained calculus techniques—i.e., take derivatives, set equal to zero, and solve the system of equations.

2. Numerical methods using a computer (Excel's Solver): Set up the problem in Excel, carefully organizing things into a goal, endogenous variables, exogenous variables, and constraint; then use Excel's Solver. Use the Sensitivity Report in the Solver Results dialog box to get λ^*.

In this chapter, we apply both methods on a new problem.

Quasilinear Utility Practice Problem

A utility function that is composed of a nonlinear function of one good plus a linear function of the other good is called a quasilinear functional form. It is *quasi*, or sort of, linear because one good increases utility in a linear fashion and the other does not.

Below are a general example and a more specific example of quasilinear utility.

$$u(x_1, x_2) = v(x_1) + x_2$$
$$u(x_1, x_2) = (x_1)^c + x_2, \text{ where } c < 1$$

If $c < 1$, then the quasilinear utility function says that utility increases at a decreasing rate as x_1 increases, but utility increases at a constant rate as x_2 increases.

The optimization problem is to maximize this utility function subject to the usual budget constraint. It is written in equation form like this:

$$\max_{x_1,x_2,\lambda} x_1^c + x_2$$
$$\text{s.t. } p_1x_1 + p_2x_2 = m$$

We will solve the general version of this problem, with letters representing exogenous variables instead of numbers, using the Lagrangean method.

1. Rewrite the constraint so that it is equal to zero.

$$0 = m - p_1x_1 - p_2x_2$$

2. Form the Lagrangean function.

$$\max_{x_1,x_2,\lambda} L = x_1^c + x_2 + \lambda(m - p_1x_1 - p_2x_2)$$

Note that the Lagrangean function, L, has the quasilinear utility function plus the Lagrangean multiplier, λ, times the rewritten constraint.

Unlike the concrete problem in the previous chapter, which used numerical values, this is a general problem with letters indicating exogenous variables. General problems, without numerical values for exogenous variables, are harder to solve because we have to keep track of many variables and make sure we understand which ones are endogenous versus exogenous. If the solution can be written as a function of the exogenous variables, however, it is often easy to see how an exogenous variable will affect the optimal solution.

3. Take partial derivatives with respect to x_1, x_2, and λ.

$$\frac{\partial L}{\partial x_1} = cx_1^{c-1} - p_1\lambda$$
$$\frac{\partial L}{\partial x_2} = 1 - p_2\lambda$$
$$\frac{\partial L}{\partial \lambda} = m - p_1x_1 - p_2x_2$$

Remember that the partial derivative treats other variables as constants. Thus, the partial derivative of the quasilinear utility function with respect to x_1 has no x_2 variable in it.

4. Set the derivatives equal to zero and solve for x_1*, x_2*, and λ*.

$$\frac{\partial L}{\partial x_1} = cx_1^{c-1} - p_1\lambda = 0$$
$$\frac{\partial L}{\partial x_2} = 1 - p_2\lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = m - p_1x_1 - p_2x_2 = 0$$

We use the same solution method as before, moving the lambda terms to the right-hand side and then dividing the first equation by the second, which allows us to cancel the lambda terms.

$$cx_1^{c-1} = p_1\lambda$$
$$1 = p_2\lambda$$
$$\frac{cx_1^{c-1}}{1} = \frac{p_1\lambda}{p_2\lambda}$$
$$\frac{cx_1^{c-1}}{1} = \frac{p_1}{p_2}$$

By canceling the lambda terms, we have reduced the three equation, three unknown system to two equations with two unknowns.

$$\frac{cx_1^{c-1}}{1} = \frac{p_1}{p_2}$$
$$m - p_1x_1 - p_2x_2 = 0$$

Remember that not all variables are the same. The endogenous variables, the unknowns, are x_1 and x_2. The other letters are exogenous variables.

From the first equation, we can solve for the optimal quantity of good 1 (see the appendix to the previous section if these steps are confusing).

$$\frac{cx_1^{c-1}}{1} = \frac{p_1}{p_2}$$
$$cx_1^{c-1} = \frac{p_1}{p_2}$$
$$x_1^{c-1} = \frac{p_1}{cp_2}$$
$$x_1^* = \left(\frac{p_1}{cp_2}\right)^{\frac{1}{c-1}}$$

Notice that we used the rule that $(x^a)^b = x^{ab}$. Because we wanted to solve for x_1, we raised both sides to the $\frac{1}{c-1}$ power so that the $c-1$ exponent on x_1 times $\frac{1}{c-1}$ would equal 1.

Usually, when we have the MRS equal to the price ratio, we need to solve for one of the x variables in terms of the other and substitute it into the budget constraint. However, a property of the quasilinear utility function is that the MRS only depends on x_1; thus by solving for x_1, we get the reduced form solution. When solving a problem in general terms, the answer must be expressed as a function of exogenous variables alone (no endogenous variables) and this is called a reduced form.

To get x_2, we simply substitute x_1 into the budget constraint and solve for x_2.

$$m - p_1\left[\left(\frac{p_1}{cp_2}\right)^{\frac{1}{c-1}}\right] - p_2x_2 = 0$$
$$x_2^* = \frac{m}{p_2} - \frac{p_1}{p_2}\left(\frac{p_1}{cp_2}\right)^{\frac{1}{c-1}}$$

It is a bit messy, but it is the answer. We have an expression for the optimal amount of x_2 that is a function of exogenous variables alone.

To get the optimal value of lambda, we can use the second first-order condition, which simply says that $\lambda^* = \frac{1}{p_2}$. If you use the first condition, substituting in the value for optimal x_1, it will take a little work, but you will get the same result.

Practice with the MRS = $\frac{p_1}{p_2}$ Logic

Economists stress marginal thinking. The idea is that, from any position, you can move and see how things change. If there is improvement, continue moving. The optimal solution is on a flat spot, where improvement is impossible.

When we move the lambda terms over to the right-hand side and divide the first equation by the second equation, we get a crucial statement of the fact that improvement is impossible and we are optimizing.

The familiar MRS equals the price ratio expression, along with the third first-order condition, which says that the consumer must be on the budget line (exhausting all income), is a mathematical way of describing marginal thinking.

The MRS condition tells us that if the MRS is not equal to the price ratio, there are two possibilities, depicted in Figure 3.7.

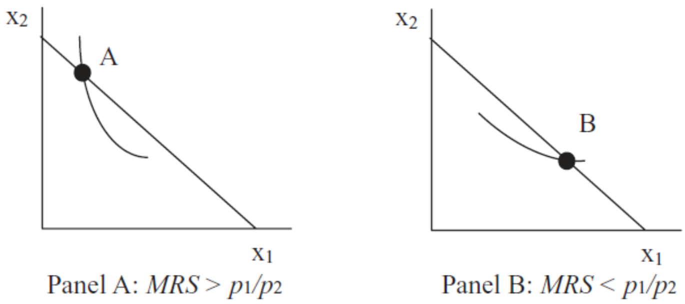


Figure 3.7: MRS does not equal the price ratio.

In Panel A, the slope of the indifference curve at point A is greater than the slope of the budget line (in absolute value). This consumer should crawl down the budget line, reaching higher indifference curves, until the MRS equals the price ratio. At this point, the slope of the indifference curve will exactly equal the slope of the budget line and the consumer's indifference curve will just touch the budget line. The consumer cannot possibly get to a higher indifference curve and stay on the budget constraint. This is the best possible solution.

In Panel B, the story is the same, but reversed. The slope of the indifference curve at point B is less than the slope of the budget line. This consumer should crawl up the budget line, reaching higher indifference curves, until the MRS equals the price ratio. At this point, the slope of the indifference curve will exactly equal the slope of the budget line and the consumer's indifference curve will just touch the budget line.

Numerical Approach to Quasilinear Practice Problem

STEP Open the Excel workbook *OptimalChoicePractice.xls*, read the *Intro* sheet, and then go to the *QuasilinearChoice* sheet to see how the numerical approach can be used to solve this problem.

It is easy to see that the consumer cannot afford the bundle 5,20 given the prices and income on the sheet. If she buys five units of x_1, what's the maximum x_2 she can buy?

STEP Enter this amount in cell B12. Does the chart and cell B21 confirm that you got it right?

If you entered 13 in B12, then the chart updates and shows that the consumer is now on the budget line. In addition, the constraint cell, B21, is now zero.

Without running Solver or doing any calculations at all, is she maximizing at 5,13?

The answer is that she is not. It's hard to see on the chart whether the indifference curve is cutting the budget line, but the information below the chart shows that the MRS is not equal to the price ratio. That tells you that the indifference curve is, in fact, not tangent to the budget line so the consumer is not optimizing. Because the MRS is greater than the price ratio (in absolute value) we also know that the consumer should buy more x_1 and less x_2, moving down the budget line until the marginal condition is satisfied. Let's find the optimal solution.

STEP Run Solver. Select the Sensitivity Report to get λ*.

How does Excel's answer compare to our analytical answer? Recall that we found:

$$x_1^* = \left(\frac{p_1}{cp_2}\right)^{\frac{1}{c-1}}$$

$$x_2^* = \frac{m}{p_2} - \frac{p_1}{p_2}\left(\frac{p_1}{cp_2}\right)^{\frac{1}{c-1}}$$

STEP Create formulas in Excel to compute these two solutions (using cells C11 and C12 would make sense). This requires some care with the parentheses. Here is the formula for good 1: =(p1_/(c_*p2_))^(1/(c_-1)).

You should discover that Excel's Solver is quite close to the exactly correct solution, 6.25, 12.75. We conclude that the two methods, analytical and numerical, substantially agree.

It is true, however, that Solver is ever so slightly off the computed analytical result. In general, there are two reasons for minuscule disagreement between the two methods.

1. Excel cannot display the algebraic result to an infinite number of decimal places. If the solution is a repeating decimal or irrational number, Excel cannot handle it. Even if the number can be expressed as a decimal—for example, one-half is 0.5—precision error may occur during the computation of the final answer. This is not the source of the discrepancy in this case.

2. Excel's Solver often misses the exactly correct answer by small amounts. Solver has a convergence criterion (that you can set via the Options button in the Solver Parameters dialog box) that determines when it stops hunting for a better answer. Figure 3.8 offers a graphical representation of Solver's algorithm in a one-variable case.

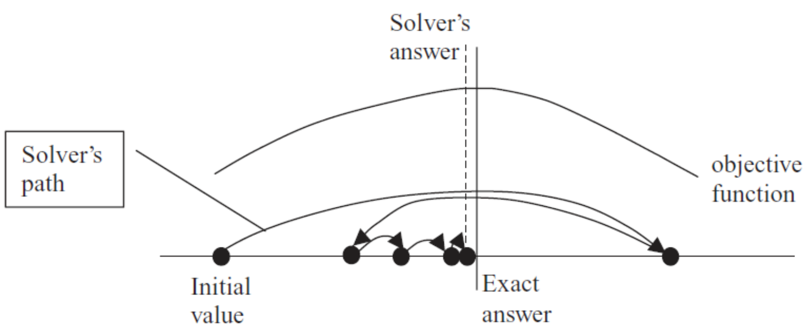


Figure 3.8: Solver in action.

The *stylized graph* (which means it represents an idea without using actual data) in Figure 3.8 shows that Solver works by trying different values and seeing how much improvement occurs. The path of the choice variable (on the x axis) is determined by Solver's internal optimization algorithm. By default, it uses Newton's method (a steepest descent algorithm), but you can choose an alternative by clicking the Options button in the Solver dialog box.

When Solver takes a step that improves the value of the objective function by very little, determined by the convergence criterion (adjustable via the Options button), it stops searching and announces success. In Figure 3.8, Solver is missing the optimal solution by a little bit because, if we zoomed in, the objective function would be almost flat at the top. Solver cannot distinguish additional improvement.

When we say that the analytical method agrees with Solver, we do not mean that the two methods exactly agree, but simply that they correspond, in a practical sense. If Solver is off the exact answer in the 15th decimal place, that is agreement, for all practical purposes.

Furthermore, it is easy to conclude that Solver must give an exact answer because it displays so many decimal places. This is incorrect. Solver's display is an example of *false precision*. It is not true that the many digits provide useful information. The exact answer is 6.25 and 12.75. What you are seeing is Solver noise. You must learn to interpret Solver's results as inexact and not report all of the decimal places.

There is another way in which Solver can fail us and it is much more serious than incorrectly interpreting the results.

Solver Behaving Badly

STEP Start from $x_1 = 1, x_2 = 20$ to see a demonstration that Solver is not perfect. After setting cells B11 and B12 to 1 and 20, respectively, run Solver. What happens?

A *miserable result* (an actual, technical term in the numerical methods literature) occurs when an algorithm reports that it cannot find the answer or displays an obviously erroneous solution. Figure 3.9 displays an example of a miserable result. Solver is clearly announcing that it cannot find an answer.

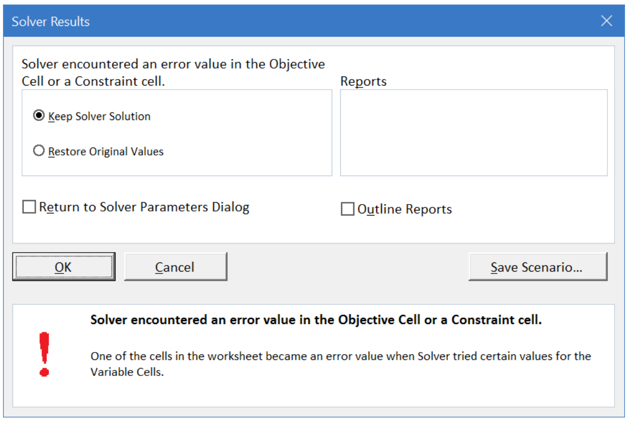


Figure 3.9: A miserable result.

If you look carefully at the spreadsheet (click cancel or OK if needed to return to the sheet), you will see that Solver blew up when it tried a negative value for x_1. The objective function cell, B7, is displaying the error #NUM! because Excel cannot take the square root of a negative number.

To be clear, when we start from 1,20, Excel tries to move left and crosses over the y axis into negative x territory. Since the utility function is $x_1^{0.5}$, it tries to take the square root of a negative number, producing an error, and crashing the algorithm.

When Solver fails, there are three basic strategies to fix the problem:

1. Try different initial values (in the changing cells). If you know roughly where the solution lies, start near it. Always avoid starting from zero or a blank cell.

2. Add more structure to the problem. Include non-negativity constraints on the endogenous variables, if appropriate. In the case of consumer theory, if you know the buyer cannot buy negative amounts, add this information.

3. Completely reorganize the problem. Instead of directly optimizing, you can put Solver to work on equations that must be met. In this problem, you know that MRS $= \frac{p_1}{p_2}$ is required. You could create a cell that is the difference between the MRS and the price ratio and have Solver find the values of the choice variable that force this cell to equal zero.

Let's try the second strategy.

STEP Reset the initial values to 1 and 20, then launch Solver (click the Data tab and click Solver) and click the Add button (at the top of the stacked buttons on the right).

Solver responds by popping up the Add Constraint dialog box.

STEP Select both of the endogenous variables in the Cell Reference field, select >=, and enter 0 in the Constraint field so that the dialog box looks like Figure 3.10. Click OK.

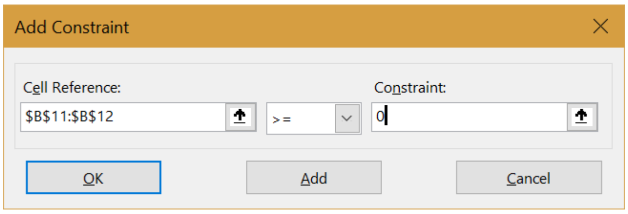


Figure 3.10: A miserable result.

You are returned to the main Solver Parameters dialog box, but you have added the constraint that cells B11 and B12 must be non-negative.

You might notice that you could have have simply clicked the *Make Unconstrained Variables Non-Negative* option, but adding the constraint shows how to work with constraints.

STEP Once back at the main Solver Parameters dialog box, click Solve.

This time, Solver succeeds. Adding the non-negativity constraint prevented Solver from trying negative x_1 values and producing an error.

Perfect Complements Practice Problem

Recall that L-shaped indifference curves represent perfect complements, which are reflected via the following mathematical function:

$$u(x_1, x_2) = min ax_1, bx_2$$

Suppose $a = b = 1$ and the budget line is $50 = 2x_1 + 10x_2$.

First, We want to solve this problem analytically.

The Lagrangean method cannot be applied because the function is not differentiable at the corner of the L. The Lagrangean method, however, is not the only analytical method available. Figure 3.11 shows that when a = b = 1, the optimal solution must lie on a ray from the origin with slope +1.

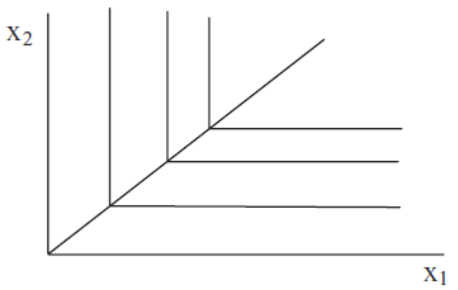


Figure 3.11: The optimal solution line with perfect complements.

The optimal solution has to be on the corner of the L-shaped indifference curves because a non-corner point (on either the vertical or horizontal part of the indifference curve) implies the consumer is spending money on more of one of the goods without getting any additional satisfaction. Thus, we know that the optimal solution must lie on the line $x_2 = x_1$.

We can combine this optimal solution equation with the budget constraint to find the optimal solution. The two equation, two unknown system can be solved easily by substitution.

$$\left.\begin{array}{l} x_2 = x_1 \\ 50 = 2x_1 + 10x_2 \end{array}\right\} \Rightarrow 50 = 2x_1 + 10\,[x_1] \Rightarrow 50 = 12x_1 \Rightarrow x_1^* = 4\frac{1}{6}.$$

Of course, we know $x_2 = x_1$ so optimal x_2 is also $4\frac{1}{6}$. Can Excel do this problem and do we get the same answer? Let's find out.

STEP Proceed to the *PerfectComplements* sheet to see how we set up the spreadsheet in Excel. Click on cell B7 to see the utility function.

STEP Run Solver and get a Sensitivity Report. Solver can be used to generate a value for the Lagrangean multiplier (via the Sensitivity Report) even though we could not use the Lagrangean method in our analytical work.

As with the previous problem (with quasilinear utility), we find that Solver and the analytical approach substantially agree. The answer is a repeating decimal, so Excel cannot get the exact answer, $4\frac{1}{6}$, but it is really close.

Previously, we saw that Solver could crash and give a miserable result. Now, let's learn that Solver can really misbehave.

STEP Starting from $x_1 = 1, x_2 = 1$, run Solver. What happens?

You are seeing an example of a *disastrous result* which occurs when an algorithm reports that it has found the answer, but it is wrong. There is no obvious error and the user may well accept the answer as true.

Solver reports a successful outcome, but the answer it gives is 1,1 and we know the right answer is $4\frac{1}{6}$ for both goods.

Disastrous results include an element of interpretation. In this case, we might notice that 1,1 is way inside the budget constraint and, therefore, the algorithm has failed. A truly disastrous result occurs when there is no way to independently test or verify the algorithm's wrong answer.

Miserable and disastrous results are well defined, technical terms in the mathematical literature on numerical methods. Disastrous results are much more dangerous than miserable results. The latter are frustrating because the computer cannot provide an answer, but disastrous results lead the user to believe an answer that is actually wrong. In the world of numerical optimization, they are a fact of life. Numerical methods are not perfect. You should never completely trust any optimization algorithm.

Understanding Solver—Be Skeptical

This chapter enabled practice solving the consumer's constrained optimization problem with two different utility functions, a quasilinear function and perfect complements. In both cases, we found that Excel's Solver agreed, practically speaking, with the analytical method.

The ability to solve optimization problems with two independent methods means we can be really sure we have found an optimal solution when they give the same answers.

In addition, we explored how Solver actually works. It evaluates the objective function for different values of the choice variables. It continues searching for a better solution until it cannot improve much (an amount determined by the convergence criterion).

Solver can fail by reporting that it cannot find a solution (called a miserable result) or—even worse—by reporting an incorrect answer with no obvious error (which is a disastrous result).

It is easy to believe that a result displayed by a computer is guaranteed to be correct. Do not be careless and trusting—numerical methods can and do fail, sometimes spectacularly.

This point deserves careful repetition. You run Solver and it happily announces that a solution has been found and offers up a 15 or 16 digit number for your inspection. The problem, however, is that the solution is *way off*. Not in the millionth or even tenth decimal place, but completely, totally wrong. How this might happen takes us too far afield into the land of numerical optimization, but suffice it to say that you should always ask yourself if the answer makes common sense.

Solver really is a powerful way to solve optimization problems, but it is not perfect. You need to always remember this. After running Solver, format the results with an eye toward ease of understanding and think about the result itself. Do not mindlessly accept a Solver result. Stay alert even if Solver claims to have hit pay dirt—it may be a disastrous result!

More explanation of Solver is available in the *SolverInstructions.doc* file in the *SolverCompStaticsWizard* folder.

Exercises

1. In the quasilinear example in this chapter, use the first equation in the first-order conditions to find λ^*. Show your work.

2. Use analytical methods to find the optimal solution for the same perfect complements problem as presented in this chapter, except that $a = 4$ and $b = 1$. Show your work.

3. Draw a graph (using Word's Drawing Tools) of the optimal solution for the previous question.

4. Use Excel's Solver to confirm that you have the correct answer. Take a picture of the cells that contain your goal, endogenous variables, and exogenous variables.

References

As economics became more mathematical, a new course was born, Math Econ. The course needed books and R. G. D. Allen's *Mathematical Analysis for Economists* (first published in 1938) became a classic textbook. As E. Schneider, a reviewer, said, "This book fills a long-felt want. At last we possess a book which presents the mathematical apparatus necessary to a serious study of economics in a form suited to the needs of the economist." See *The Economic Journal*, Vol. 48, No. 191 (September, 1938), p. 515. The epigraph is from page 2 of *Mathematical Analysis for Economists*, as Allen discusses how and why mathematics can be applied to the study of economics.

Tastes are the unchallengeable axioms of a man's behavior; he may properly (usefully) be criticized for inefficiency in satisfying his desires, but the desires themselves are *data*.

George Stigler and Gary Becker

3.3 Food Stamps

This chapter applies the consumer choice model to a real-world example. We will see that the model can be used to explain why someone would illegally sell food stamps. We also tackle an important policy question: If cash dominates food stamps, why not just help low-income people by giving them cash?

A Short History of Food Assistance in the United States

The primary responsibility for ensuring poor people (including children) in the United States have enough to eat lies with the Department of Agriculture (USDA). They run a program that enables low-income people to spend government-provided benefits on eligible food in stores.

The USDA's web page, (www.fns.usda.gov/snap/short-history-snap), is the source of the information below. The Data and Research tab on the USDA's website has usage and cost data—there are around 40 million participants and the program spends roughly $70 billion per year. This is one of the largest transfer programs in the fight against poverty. It offers critical support for low-income households.

The first Food Stamp Program, in 1939, was very different from today's version. Originally, "the program operated by permitting people on relief to buy orange stamps equal to their normal food expenditures. For every $1 worth of orange stamps purchased, 50 cents worth of blue stamps were received. Orange stamps could be used to buy any food. Blue stamps could only be used to buy food determined by the Department to be surplus."

Important changes were made in the 1960s and, in 1977, the purchase requirement was eliminated. Households below the poverty line who met other criteria (such as work or study requirements) were eligible to receive food stamps. Figure 3.12 shows that these stamps were like paper currency; they were rectangular, but only about half the size of a dollar bill. There were different dollar denominations in a booklet. When buying food at the supermarket, the consumer tore out the stamp and paid for the food. They would pay for any non-food items with cash or a check.

Figure 3.12: Old US food stamps.
Source: Public domain file photo.

In 2008, it was renamed the Supplemental Nutrition Assistance Program (SNAP) to avoid stigma. It could be embarrassing to pay with food stamps since everyone in line immediately knew that you were receiving government assistance. Today, both names, food stamps and SNAP, are used.

SNAP has always been battered by politics, with benefits expanding and contracting depending on the rhetoric of the day. There are the usual arguments over administrative costs, but cheating on the part of recipients has been an especially contentious issue. In 2002, all states were required to use Electronic Benefits Transfer (EBT) cards. This was supposed to stop the illegal sale of food stamps (and reduce stigma), but fraud remains a focus of critics.

We can model and analyze food stamps with the Theory of Consumer Behavior. We will focus on how food stamps can be incorporated into the consumer's optimization problem and why selling food stamps is so difficult to stop.

Food Stamp Theory

Recall from the Budget Constraint chapter that food stamps are a subsidy that produces a budget constraint with a horizontal segment, as shown in Figure 3.13. We use the x_1 variable on the x axis to represent units of food. The x_2 variable on the y axis captures all other goods lumped together. We get the flat part of the constraint because food stamps can be used to buy only food.

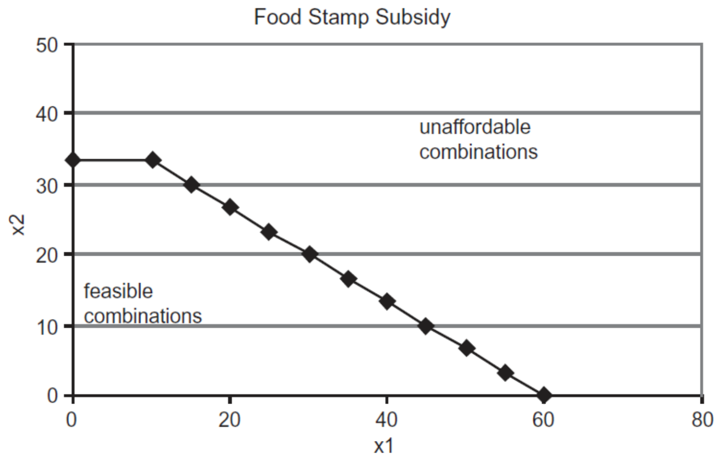


Figure 3.13: The budget constraint with food stamps.
Source: FoodStamps.xls!BudgetConstraint

STEP Open the Excel workbook *FoodStamp.xls* and read the *Intro* sheet. Proceed to the *BudgetConstraint* sheet. Change cell E13 from 10 to 20.

Notice that the horizontal segment, which is the monetary value of the food stamps divided by the price of food, gets longer. Also notice that the chart on the right, showing the budget constraint if the food stamp amount was treated as cash, has no horizontal segment. In the chart on the right, the value of the food stamp subsidy is computed (*xbar* times price of food) and then added to income as if it were cash; hence the name, cash-equivalent subsidy.

It should be quite clear that the cash-equivalent subsidy provides consumption possibilities that are unattainable above the horizontal segment of the food stamp budget constraint. The most other goods the food stamp recipient can buy is $33\frac{1}{3}$ units, while the cash-equivalent consumer can buy 40 units of x_2.

STEP Proceed to the *Inframarginal* sheet. It combines a food stamp budget constraint with a Cobb-Douglas utility function.

The word *inframarginal* (or submarginal) means below the edge or margin. The edge in this case is the kink in the budget constraint.

This consumer is *inframarginal* because his optimal solution is on the downward sloping part of the budget line, below the kink. He will use up his food stamp allotment on food and then spend some of his cash income to get additional food. The sheet reveals that he buys 35 units of food (valued at $70, as shown in cell B15), 20 of which he obtains with food stamps and the remaining 15 he buys with cash.

We can easily see that he is optimizing because the "MRS equals the price ratio" condition is met. This is reflected in the graph where the highest attainable indifference curve is just touching the budget constraint.

STEP Click on cell B25 to see the formula for the budget constraint.

This formula is using an IF statement to implement the constraint in Excel. Expressed as an equation, the budget line looks like this:

$$\begin{aligned} &\text{if } x_1 \leq \bar{x}, \quad x_2 = m/p_2 \\ &\text{if } x_1 > \bar{x}, \quad x_2 = m/p_2 - p_1/p_2\,(x_1 - \bar{x}) \end{aligned}$$

The first equation says that if the consumer buys an amount of food that is less than or equal to *xbar*, that frees up his whole cash income to spend on good 2. This is the horizontal line component.

Things are more complicated if the consumer wants more than *xbar* of food. The second equation says that the consumer will have to use cash to buy amounts of x_1 greater than *xbar* and it computes the amount of x_2 that can be purchased as a function of x_1.

This constraint (rewritten to equal zero) has been entered in a single cell with an IF statement:

=IF(x1_<x1bar,m/p2_-x2_,m/p2_-(p1_/p2_)*(x1_-x1bar)-x2_)

The underscore (_) character is used in the variable names to distinguish them from cell addresses—e.g., p2_ is not cell P2.

From Excel's Help on the IF function:

> Returns one value if a condition you specify evaluates to TRUE and another value if it evaluates to FALSE.
>
> Use IF to conduct conditional tests on values and formulas.
>
> Syntax: IF(logical test,value if true,value if false)

Applying this information to the formula in cell B25, we can see that it has three parts, separated by commas. The first part says that if $x_1 < x1bar$ (that is the condition being evaluated), then the consumer can buy m/p_2 amount of x_2 (this second part produces the horizontal line in the budget constraint), else (the third part is what happens if x_1 is not less than *x1bar*) the consumer can buy x_2 along the downward sloping part of the budget line.

This problem shows that Excel can be used to handle complicated examples in the Theory of Consumer Behavior. This food stamp problem has a kinked budget constraint, but using Excel's IF statement allows us to implement the constraint in the workbook and use Solver to find the optimal solution.

This problem also can be solved via analytical methods, but it is cumbersome and difficult to deal with the kinked budget constraint. We will use the easier numerical approach to conduct our analysis.

STEP Proceed to the *Distorted* sheet.

This sheet is exactly the same as the *Inframarginal* sheet with one crucial exception: the preferences, in cells B21 and B22, are different. The consumer in the *Distorted* sheet prefers other goods more and food less than the consumer in the *Inframarginal* sheet.

The change in exponents in the Cobb-Douglas utility function has affected the indifference map. The curves are much flatter in the *Distorted* sheet compared with the *Inframarginal* sheet.

The *Distorted* sheet opens with the optimal values for food and other goods from the *Inframarginal* sheet. It is obvious that the MRS does not equal the price ratio and the indifference curve is cutting the budget constraint at the current bundle of x_1 and x_2. This consumer is not optimizing at this point.

Corner Solution

STEP Run Solver on the *Distorted* sheet.

Solver announces it has found the optimal solution, yet the MRS still does not equal the price ratio. Is this really the optimal solution? Yes, it is the optimal solution. We have encountered what is called a *corner solution* (or boundary optimum). In this case, the equimarginal condition, MRS $= \frac{p_1}{p_2}$, does not hold because the optimal solution is found at one of the end points (or corners) of the constraint.

STEP To see what is happening here, copy the optimal solution from the *Inframarginal* sheet (copy cells B13 and B14) and paste in the *Distorted* sheet (select cells B13 and B14 and then paste).

The graph and MRS is immediately updated and you can see that the distorted consumer would not select the inframarginal consumer's bundle. Which way should this consumer move—up or down the budget line? The graph makes clear that up is the right way to go, but you should notice that the marginal condition, MRS $< \frac{p_1}{p_2}$, tells you the same thing.

STEP Click the [Crawl Up the Budget Line] button. Click a few more times and pay attention to the chart and the MRS in cell H26. Also keep an eye on utility in cell B9. Each click lowers the amount of x_1 by one unit and increases the amount of x_2 by $\frac{2}{3}$.

By moving up the budget line, this consumer is improving her satisfaction and closing the gap between the MRS and the price ratio.

Do not be misled by the display – the indifference curves are not shifting. Remember that the indifference map is dense, meaning that every point has an indifference curve through it. We cannot draw in all of the indifference curves because the graph would then be solid black. The consumer is simply moving from one indifference curve to another one that was not previously displayed.

STEP Keep clicking the [Crawl Up the Budget Line] button. Eventually, you will hit the kink in the budget line and you will not be able to move northwest any longer. Instead, you will be on the horizontal segment and as you move strictly west, utility falls. Notice that the price ratio is now showing zero.

On the flat part of the budget line, when the amount of food purchased is less than or equal to how much food can be bought with food stamps alone, it makes sense that additional food is free, in terms of spending cash on food. The consumer simply has to use the available food stamps to acquire food and this does not reduce cash income.

Once you are on the flat part of the budget line, you should see that the graph and marginal condition point you to choosing more food.

STEP Click on the Crawl Down the Budget Line button repeatedly to move east and, eventually, down the budget line. Use the two buttons to crawl up and down until you find the bundle that maximizes utility.

You should end your travels at the kink – and MRS does not equal the price ratio there! This happens because the complicated constraint is producing a corner solution.

The distorted consumer wishes she could continue crawling up the downward sloping line, consuming less than the food stamp allotment of food and more of other goods, but she cannot do this. She cannot use food stamps to buy other goods. Thus, her best, or optimal, solution is at the kink.

In a corner solution, we accept that the "MRS equals the price ratio" condition is not met. We really are maximizing even though the MRS does not equal the price ratio. We have found the best we can do given the constraints on our choices.

Another way to explain what is happening is that we always want to minimize $|MRS - \frac{p_1}{p_2}|$. With an interior solution, we can make this difference zero, but with a corner solution, we cannot because a constraint is preventing us from reaching MRS $= \frac{p_1}{p_2}$. However, a corner solution does give us the lowest $|MRS - \frac{p_1}{p_2}|$ value and we are doing the best we can at this solution.

Corner solutions are an important concept and we will see them again in future work. They arise whenever we are prevented from continuing to improve by going in a particular direction.

Cash Instead of Food Stamps

STEP Proceed to the *Cash* sheet. Notice that cell B24 computes the cash value of the food stamps and that the chart has a linear budget constraint with no kink. Click cell B25 to see that the constraint is the familiar income minus expenditures, with income equal to the sum of income plus the cash value of the food stamps.

The idea here is that instead of giving food stamps, we provide low-income people the cash-equivalent value. They are no longer constrained to buy food alone, but can purchase any goods with the cash received. The cash subsidy shifts the budget line out, with no kink or horizontal segment like we saw with the food stamp program.

The sheet opens with the inframarginal consumer's optimal solution. It is the same as before, when she was given food stamps. Cash or food stamps are the same to this consumer.

STEP Click on the Set to Distorted button to quickly apply the preferences for the distorted consumer. Run Solver.

With cash, the distorted consumer chooses an optimal bundle that is different from the one chosen under the Food Stamp Program. She finds an interior (as opposed to a corner) solution in the far northwest corner, which means she has opted for little food and more of other goods.

Figure 3.14 summarizes our work to this point. If you compare the inframarginal consumer, by looking top left and then bottom left, in Figure 3.14, you can easily see that there is no change in his behavior: $40 in food stamps versus $40 in cash are the same to this consumer.

On the other hand, comparing the top right and bottom right panels in Figure 3,14 reveals that the distorted consumer chooses less food and more other goods when given cash. This is why we say her choices are *distorted* by the food stamp program. If she had cash, she would make different choices. The distortion results in a decrease in satisfaction for this consumer.

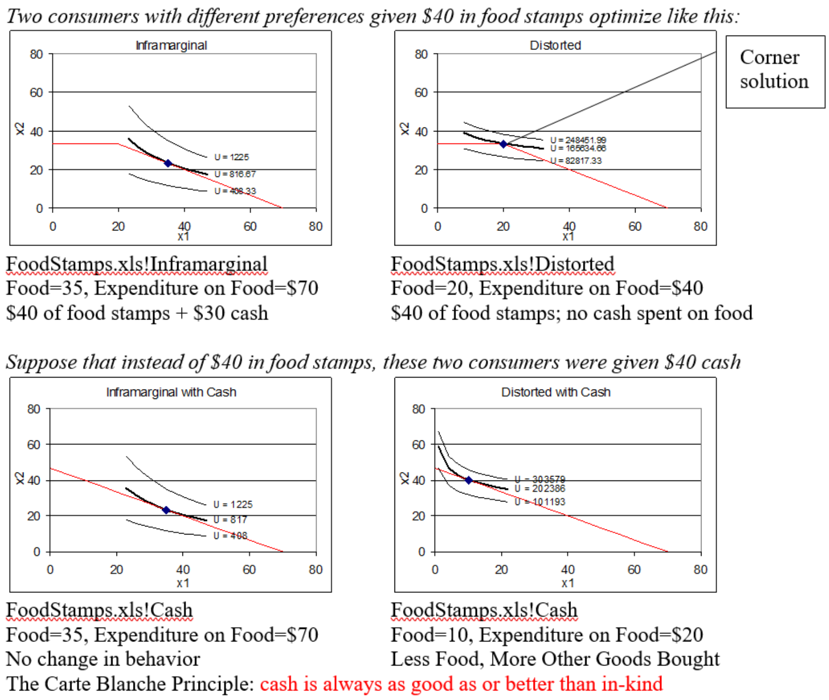


Figure 3.14: Comparing food stamps versus cash-equivalent.

The Carte Blanche Principle and Deadweight Loss

Carte blanche, a term of obvious French origin (literally, "blank document"), means unconditional authority or freedom to act in any way you wish.

In economics, the *Carte Blanche Principle* means that cash is always as good as or better than in-kind. Cash allows the consumer to buy anything, while in-kind transfers, such as food stamps, restrict the set of choices.

Figure 3.14 shows the Carte Blanche Principle in action. Cash dominates food stamps. If you are an inframarginal consumer, the cash and food stamps are the same. This consumer is going to buy more food than can be purchased with the allotment of food stamps anyway so if you gave him the cash equivalent value, he would spend the cash on food.

If you are a distorted consumer, however, you are better off if you are given cash because cash can be used to buy the other goods that you prefer over food. With food stamps, when you maximize utility and do the best you can, you end up at a lower level of utility than if you had the cash-equivalent.

In economics, *deadweight loss* is a measure of inefficiency. It is a number that tells you how much a given solution differs from the best solution. In this application, deadweight loss is the difference in utility due to using food stamps instead of cash.

We could try to compute, for each consumer, the maximum utility with cash minus the maximum utility with food stamps. For the inframarginals, this number would be zero, but it would be positive for the distorted consumers.

Unfortunately, this approach would be exceedingly difficult to actually carry out. Even if we managed to do it, remember, we cannot simply add the utility values for different people. Utility is ordinal, ranking only by higher or lower, with no meaningful information about distance or magnitude. Thus, we can never add the utilities of different people.

Theory tells us deadweight loss exists, but the inability to make interpersonal utility comparisons means we are severely limited in how we can measure the sum of deadweight losses of two or more people. As a first pass, we can try to figure out how many distorteds and inframarginals there are. After all, if there are only a few distorted consumers, then we would know that food stamps were not affecting the decisions of too many people.

A Food Stamp Experiment

The empirical work described below comes from Whitmore's "What are Food Stamps Worth?" available at arks.princeton.edu/ark:/88435/dsp01z603qx42c.

Whitmore describes two controlled experiments carried out by the USDA in the early 1990s. In the San Diego experiment, around 1,000 people who were receiving food stamps were randomly selected to participate in the experiment. Half were randomly assigned to the control group and given food stamps as usual, while the other half, the treatment group, were given cash-equivalent aid (checks).

Of the roughly 500 people given checks, about 100 were distorted—they bought less food compared to what they bought when they were given food stamps.

But what were these distorted consumers buying instead of food? This is a crucial question. Most economists are willing to let individuals choose what to buy because the Theory of Consumer Behavior is built on rational, optimizing decision making. The fundamental world view of economic theory is that individuals know best how to spend their money.

Others, however, argue that low-income consumers make poor decisions if left free to choose what to buy. They think distortion is a good thing because they want aid recipients to buy food. Whitmore (p. 3) says this:

> To some, this distortion is the best part of the food stamp program: the government can ensure that needy families get enough to eat and that they don't spend the money on other things. To others, this distortion represents a waste of resources—it is inefficient to give in-kind transfers instead of cash.

At its most extreme, the issue can be stated this way: Taxpayers will support buying food for the poor, but not drugs, alcohol, and other wasteful consumption. But exactly how distorted consumers would spend cash is an empirical question and Whitmore has the data to answer it.

Researchers in the San Diego experiment kept careful food diaries. When Whitmore compared the purchases of the distorted treatment group to the food stamp control group, she found a marked decrease in a few specific items, like juice and soda, for distorteds. So, surprisingly,

> Even though spending on food declines for the treatment group, the food diary data from San Diego provide no firm evidence that cashing-out food stamps leads to declines in nutritional intake, and suggest that it may actually reduce extreme over-consumption of calories, an important contributing factor to obesity. (Whitmore, p. 35)

The picture that many have of the indigent as drug addicts or exceptionally poor decision makers is unsupported by Whitmore's data. It is true that if forced to spend a subsidy on food, low-income households will spend more on food, but that does not imply that this is better. By definition, low-income people are struggling with paying for, not just food, but a whole host of necessities, including shelter, clothing, transportation, and utility bills. A cash-equivalent subsidy means they can buy food if that is the greatest need or make other important purchases.

The Illegal Sale of Food Stamps

The Theory of Consumer Behavior can be used to explain what most people find puzzling when they first hear about it—there is an active, illegal market in food stamps. Whitmore (p. 4) estimated that food stamps sold for 61 cents on the dollar. The theory can also explain why it has proven incredibly difficult to stop the illegal sale of food stamps.

STEP Proceed to the *Selling* sheet.

Observe that the budget constraint has been modified yet again. The segment below the food stamp allotment (*x1bar*) is no longer horizontal. We have enabled the consumer to sell food stamps and move up the budget constraint.

The slope of this portion of the budget constraint is $ER * p_1/p_2$,where ER is the exchange rate of food stamps for cash. With ER initially set at 0.6 (in cell B24), a seller of food stamps would get 60 cents for every dollar of food stamps sold. The slope of the budget line is 60% of the p_1/p_2 ratio or 1.2.

Notice that cell B16 has been added and it reports the income generated by the sale of food stamps. It shows zero because the opening position is at the kink (20, 33.33) so this distorted consumer isn't selling any food stamps.

STEP Change cell B13 to 10 and watch how the cells and the chart change.

B16 now reports that the consumer is making \$12 from the sale of food stamps. They "sold" ten units of food, valued at \$20 in cash, but only 60% of that in food stamps. With $p_2 = 3$, she can buy four more units of x_2.

STEP Set cell B14 to 37.33 to move the consumer to the budget line.

But is this is the optimal solution? In fact, comparing cell G27 to H26 tells you that it is not. The consumer is selling too many food stamps at this point.

STEP Run Solver. You should get a result like Figure 3.15, which shows the consumer choosing just under 15 units of food and adding \$6.29 of food stamp income (explaining how they managed to buy more than $33\frac{1}{3}$ units of x_2). Notice also that, once again, the MRS (-0.4) equals the slope of the budget constraint (-0.4) on the relevant part of the budget line.

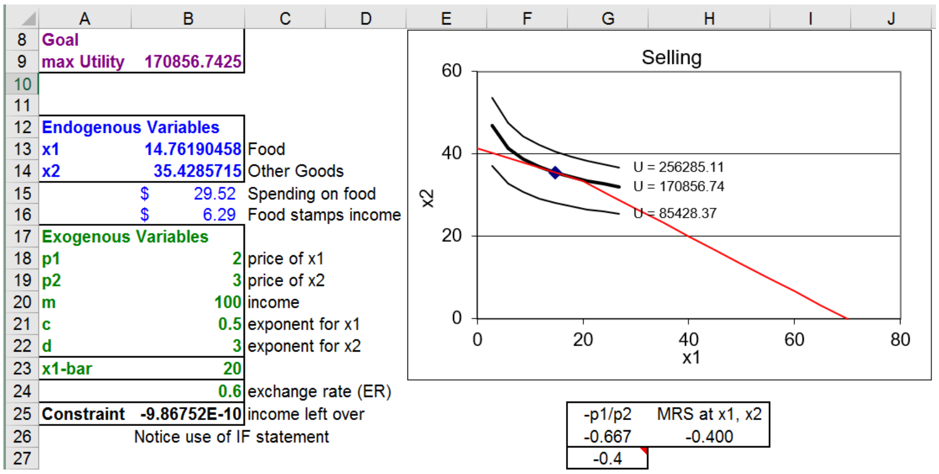


Figure 3.15: Maximizing utility by selling food stamps.
Source: FoodStamps.xls!Selling

The consumer maximizes utility and reaches a higher level of satisfaction than what is attainable by staying on the kink and not selling the food stamps. The ability to get higher satisfaction explains the unintended consequence of an active illegal trade in food stamps.

This analysis does not incorporate the costs of selling food stamps, including the risk of getting caught. There is no doubt that EBT cards make it more difficult to sell food stamps, but the inability to stop the illegal trade testifies to the forces at play—the search for higher satisfaction is powerful indeed.

One Last Question

If the Carte Blanche Principle is true, then why does the government use food stamps instead of cash to help the poor?

Whitmore devotes the conclusion of her paper (p. 38) to answering this question:

> A crucial aspect of the success of the Food Stamp Program is its political popularity. The Food Stamp Program is not an entitlement program, so its budget must be approved annually in the Farm Bill. The program's budget has always been fully funded, due largely to two factors: its popularity as a targeted welfare program among voters, and its popularity among farmers because they think it increases demand for food. (footnote omitted)

As a practical matter, it is not true that, in general, the poor will squander cash subsidies or make terrible buying decisions. Giving aid in the form of food stamps generates a deadweight loss for those distorted consumers who would have been better off with cash. As Whitmore points out, however, it is politically impossible to imagine what is today a \$70 billion program being funded annually as a pure cash giveaway. Economics meets politics and the result is a flawed, but functioning anti-poverty program.

Exercises

1. Which parameter in the *Selling* sheet, with the exchange rate set to 0.9, would have to be changed to represent the case of a distorted consumer who decides not to sell food stamps for cash? What would the value of this parameter be?

2. Explain under what condition the MRS equals the price ratio rule (as a condition that the optimal solution has been found) can be violated.

3. A seller of food stamps would obviously prefer a higher price, but what would be the advantage of a higher price in terms of the Theory of Consumer Behavior?

References

The epigraph comes from the first paragraph of Stigler and Becker's "De Gustibus Non Est Disputandum," *The American Economic Review*, Vol. 67, No. 2 (March, 1977), pp. 76 - 90 (www.jstor.org/stable/1807222). The title is a Latin admonition to not quarrel over tastes—do not continue arguing once you pass the point of rational persuasion (similar to "Let's agree to disagree.").

Stigler and Becker offer, however, a second interpretation, which they prefer: "tastes neither change capriciously nor differ importantly between people." Their key point is this:

> The difference between these two viewpoints of tastes is fundamental. On the traditional view, an explanation of economic phenomena that reaches a difference in tastes between people or times is the terminus of the argument: the problem is abandoned at this point to whoever studies and explains tastes (psychologists? anthropologists? phrenologists? sociobiologists?). On

> our preferred interpretation, one never reaches this impasse: the economist continues to search for differences in prices or incomes to explain any differences or changes in behavior. (p. 76)

The idea that tastes are stable and differences in behavior are to be found in price or income shocks is a hallmark of Chicago School economics.

Diane Whitmore's working paper, "What Are Food Stamps Worth?," is available at arks.princeton.edu/ark:/88435/dsp01z603qx42c. Whitmore goes beyond simply counting the number of distorted consumers and offers estimates of deadweight loss.

For more recent work, see Hillary Hoynes and Diane Whitmore Schazenbach, "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program," *American Economic Journal: Applied Economics*, Vol. 1, No. 4 (October 2009), pp. 109 - 139, available at https://www.jstor.org/stable/25760184.

Taxes upon the necessaries of life have nearly the same effect upon the circumstances of the people as a poor soil and a bad climate.

Adam Smith

3.4 Cigarette Taxes

The Carte Blanche Principle says that cash is always as good as or better than in-kind. There is a corollary from the public finance literature: Lump sum taxes are better than quantity taxes.

Public finance is a field of economics that studies the role of government in the economy. Budgeting, collecting taxes, and government spending are some of the areas studied by public finance economists.

There are, of course, many different kinds of taxes. A *lump sum tax* is a fixed amount that must be paid, regardless of how much is purchased. A head tax, where a fee is charged to each person, is an example of a lump sum tax.

A *quantity tax* is an amount for each unit sold so it is added to the price of the product. Federal, state, and local governments levy quantity taxes on gasoline, alcoholic beverages, and tobacco. Unlike a lump sum tax, if more is bought, more quantity tax is paid.

Most people are familiar with sales tax, but this is yet another tax variant. Like a quantity tax, more is paid as more is purchased, but a sales tax is a percentage of the total purchase value. This is an *ad valorem* tax, which is Latin for "according to value."

The goals of taxation can be complicated. The primary motivation for taxes is to pay for government spending, but taxes can also be used to discourage particular activities. Both of these motivations are at play in the case of cigarettes.

Cigarette Smoking and Taxes

The average number of cigarettes sold per day in the United States and Japan since 1900 is shown in Figure 3.16. Visit ourworldindata.org/smoking to see an interactive version of this chart and add other countries. The pattern is the same around the world—rising smoking rates reach a peak, then a rapid decline.

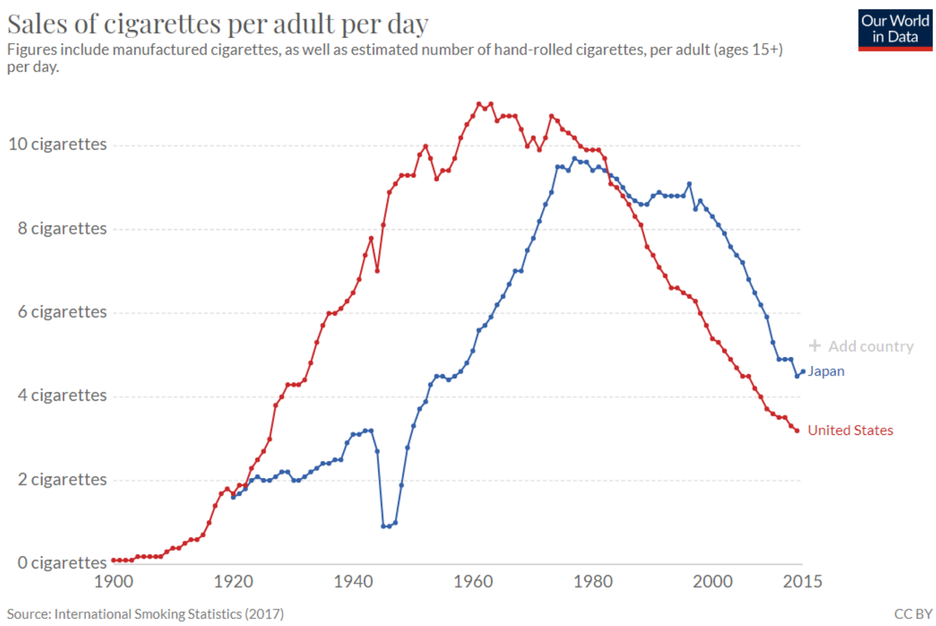


Figure 3.16: Smoking rates in Japan and the United States.
Source: ourworldindata.org/smoking

American soldiers were given cigarettes during the two world wars and this drove the sharp increase in cigarette smoking. The collapse in its smoking rate in the 1940s shows that Japan did not do this. In both countries, awareness of the damaging health effects of smoking triggered the decline.

As consumption underwent this long rise and fall, cigarette tax policies also changed dramatically. Tobacco products have always been taxed, but cigarette taxes have risen dramatically in the last few decades. Figure 3.17 shows tax rates in US states in 2019.

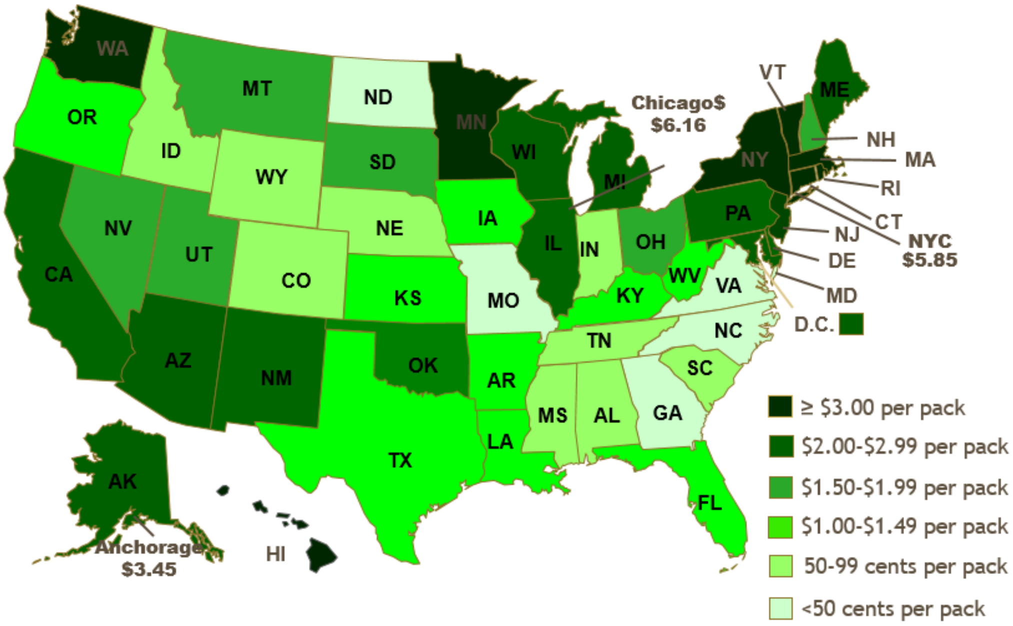


Figure 3.17: State cigarette quantity taxes in 2019.
Source: tobacconomics.org

There is wide variation in state cigarette tax rates. In 2019, New York and Connecticut had the highest state tax of $4.35 per 20-pack of cigarettes. Missouri had the lowest, $0.17 per pack.

Other governmental levels also tax cigarettes. New York City, for example, adds a $1.50 per pack tax, bringing state and local taxes to $5.85 per pack. To this we add the federal tax rate of $1.0066 per pack. Finally, smokers pay a sales tax on the total price paid (including the quantity taxes). In New York City, a pack of cigarettes cost over $10 in 2019.

We will analyze the quantity tax by using the Theory of Consumer Behavior. We will also compare it to a lump sum tax—an option that is not currently being used by the government. To make a good comparison, we have to make sure that the taxes are *revenue neutral.* This means that the tax revenues generated by the tax proposals are the same. It would not be fair to compare a quantity tax that generated $50 in revenues to a $100 lump sum tax.

Quantity Tax

STEP Open the Excel workbook *CigaretteTaxes.xls*, read the *Intro* sheet, and proceed to the *QuantityTax* sheet.

Cell B21 enables us to levy a quantity tax. The sheet opens with cell B21 = 0, which means there is no tax.

The sheet also opens with the consumer considering the bundle 20,60. The MRS is greater than the price ratio (in absolute value) and the consumer can move down the budget constraint so we know utility is not being maximized.

STEP Utility is maximized at 1250 by consuming 25 units of cigarettes (x_1) and 50 units of other goods (x_2). Run Solver to confirm this result.

Suppose we impose a \$1/unit quantity tax on cigarettes. What effect does this have on the consumer?

STEP You can find the consumer's optimal solution after levying the tax by changing cell B21 to 1 and running Solver.

Notice how the chart updated when B21 was set to one. The red budget constraint shows how the line rotated and swung in when the tax was imposed. This is the same as increasing the price of good 1. After running Solver, you can see that the consumer responds by buying fewer cigarettes.

We can also find the optimal solution using analytical methods by solving the following constrained optimization problem:

$$\max_{x_1,x_2,\lambda} U(x_1,x_2) = x_1x_2$$
$$\text{s.t. } 100 = 2(x_1 + Q\_Tax) + x_2$$

The consumer wishes to maximize utility (which is Cobb-Douglas with both exponents equal to 1), subject to the budget constraint, with parameter values for income and prices plugged in.

We leave Q_Tax as an exogenous variable so we can find the optimal solution as a function of Q_Tax. We have worked on this problem before, except $p_2 = 1$ (instead of 3) and we have added the quantity tax.

The Lagrangean procedure remains the same and we walk through the four steps to find the answer.

1. Rewrite the constraint so that it is equal to zero.

$$0 = 100 - 2(x_1 + Q\_Tax) - x_2$$

2. Form the Lagrangean function.

$$\max_{x_1,x_2,\lambda} L = x_1x_2 + \lambda(100 - (2 + Q\_Tax)x_1 - x_2)$$

Notice that we are working with a mixed concrete and general problem. We have numerical values for prices, income, and the utility function exponents, but we have the amount of the quantity tax as a variable. We use this strategy whenever we want to find the optimal solution as a function of a particular exogenous variable.

3. Take partial derivatives with respect to x_1, x_2, and λ.

$$\frac{\partial L}{\partial x_1} = x_2 - (2 + Q\_Tax)\lambda$$
$$\frac{\partial L}{\partial x_2} = x_1 - \lambda$$
$$\frac{\partial L}{\partial \lambda} = 100 - (2 + Q\_Tax)x_1 - x_2$$

4. Set the derivatives equal to zero and solve for $x_1{}^*$, $x_2{}^*$, and λ^*.

$$\frac{\partial L}{\partial x_1} = x_2 - (2 + Q\_Tax)\lambda = 0$$
$$\frac{\partial L}{\partial x_2} = x_1 - \lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = 100 - (2 + Q\_Tax)x_1 - x_2 = 0$$

We use the usual solution method, moving the lambda terms to the right-hand side and then dividing the first equation by the second, which allows us to cancel the lambda terms.

$$x_2 = (2 + Q\_Tax)\lambda$$
$$x_1 = \lambda$$
$$\frac{x_2}{x_1} = \frac{(2 + Q\_Tax)\lambda}{\lambda}$$
$$x_2 = (2 + Q\_Tax)x_1$$

Finding an expression for x_2 seems like an answer, but it is not because it is a function of x_1. To be a solution (which is called a reduced form), we must solve for x_1 asa function of exogenous variables alone. We must keep working. Canceling the lambda terms has moved us closer to an answer—we have reduced the three equation, three unknown system to two equations in two unknowns.

$$x_2 = (2 + Q\_Tax)x_1$$
$$100 - (2 + Q\_Tax)x_1 - x_2 = 0$$

We substitute the first equation into the second and solve for the optimal amount of good 1.

$$100 - (2 + Q\_Tax)x_1 - [(2 + Q\_Tax)x_1] = 0$$
$$100 = 2(2 + Q\_Tax)x_1$$
$$x_1^* = \frac{50}{(2 + Q\_Tax)}$$

Then, we substitute this into our expression for x_2 to get the optimal amount of good 2.

$$x_2^* = (2 + Q\_Tax)\left[\frac{50}{(2 + Q\_Tax)}\right] = 50$$

We can check this solution with Solver's result by substituting $Q\_Tax = 1$ into the reduced form solution for the two goods. Optimal cigarette consumption is $\frac{50}{3}$ or $16\frac{2}{3}$. Because Q_Tax does not appear in the optimal solution for good 2, its value is simply 50 for any value of Q_Tax.

Lump Sum Tax

Let's see how the consumer would optimize with a lump sum tax that raised the same tax revenue for the government.

STEP Making sure that you have run Solver in the Quantity sheet with B21 = 1 so that B11 is approximately $16\frac{2}{3}$, proceed to the *LumpSumTax* sheet.

The quantity tax imposed in the *QuantityTax* sheet has been replaced with a revenue-neutral lump sum tax. With a $1/unit quantity tax, the consumer purchases $16\frac{2}{3}$ units of x_1, which means the state generates $16.67 of revenue from the quantity tax. It could have generated the same revenue by taxing the consumer $16.67, regardless of how much x_1 or x_2 the consumer bought. This is called a lump sum tax because you pay a fixed amount (that's the "lump sum" part) no matter what you decide to buy.

The difference in the way the lump sum tax operates is reflected in the budget constraint equation. Instead of being part of the price of good 1 like a quantity tax, the lump sum tax is subtracted from income.

$$100 = 2(x_1 + Q\_Tax) + x_2$$
$$100 - Lump\_Tax = 2x_1 + x_2$$

The two charts show how the lump sum tax works differently than the quantity tax. Instead of rotating, the new budget line (in red) in the *LumpSum* sheet has shifted inwards. How would the consumer respond to this tax?

STEP Run Solver to find the optimal solution with the lump sum tax.

Before we compare the quantity and lump sum tax solutions, we confirm Solver's answer in the *LumpSum* sheet by solving the problem analytically.

STEP Try your hand at this problem. Check your work (or peek if you get stuck) by clicking the Show Math button.

Remember, Solver gave you an answer so can be quite sure you are correct if your analytical work gives the same result.

Comparing Quantity and Lump Sum Taxes

We now have the data needed to compare the two tax schemes, as shown in Figure 3.18.

Tax	Revenue	x_1*	x_2*	Utility*
No tax	$0	25	50	1250
Q tax = $1/unit	$16.67	16 2/3	50	833 1/3
Lump tax = $16.67	$16.67	20 5/6	41 2/3	868

Figure 3.18: Comparing the tax schemes.

The first row shows that the consumer will buy the bundle 25,50 when there is no tax, generating an optimal utility of 1250. Obviously, there is no revenue because there is no tax.

The second row shows that utility falls to $833\frac{1}{3}$ with an optimal solution of $16\frac{2}{3}$,50 with a $1/unit of x_1 quantity tax. The tax produces $16.67 of revenue for the government.

The last row shows that a revenue-neutral lump sum tax of $16.67 would result in purchases of $21\frac{5}{6}$ and $41\frac{2}{3}$, which would give a level of utility of 868.

The primary lesson is that, for this consumer, if the government needed to raise $16.67 of tax revenue, the lump sum tax is better than the quantity tax because the consumer's maximum utility is higher under the lump sum tax.

Notice that we are not violating the rule against interpreting utility values as being meaningful. We are not comparing two consumers. We are not treating utility as if it were on a cardinal scale by saying, for example, that there is a gain of 868 minus $833\frac{1}{3}$ equals $34\frac{2}{3}$ utils of increased satisfaction. We are merely saying that satisfaction is higher under the lump sum tax scheme than the revenue-neutral quantity tax.

A graph can be used to explain this rather curious result that lump sum taxes enable higher utility than equivalent revenue quantity taxes. It is a complicated graph, so we will build up to it in stages.

The first layer is simply the initial solution, before any tax is applied. It is shown in Figure 3.19.

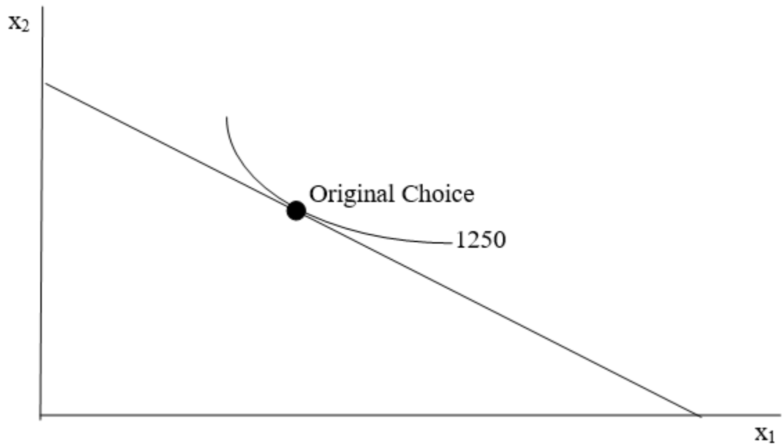


Figure 3.19: The initial optimal solution.

Figure 3.20 shows what happens with a quantity tax. The budget constraint rotates in because the price paid by the consumer (composed of the price of the product plus the tax) has increased. The consumer is forced to re-optimize and find a new optimal solution, labeled *Quantity Tax*. Utility has clearly fallen since we are on a lower indifference curve.

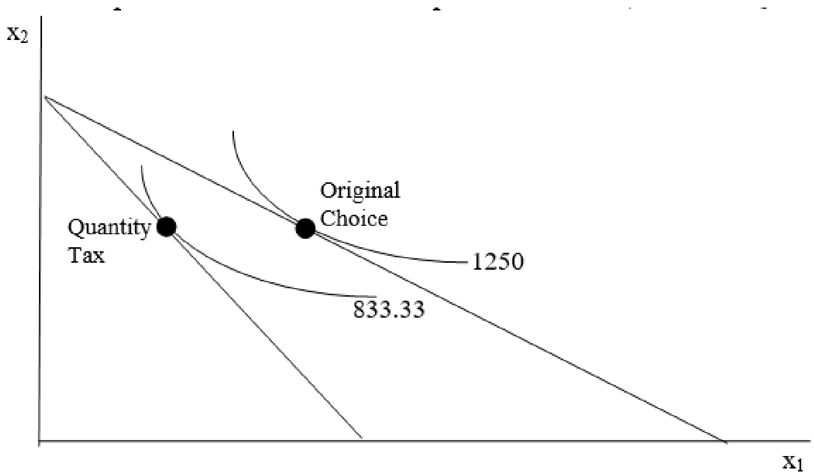


Figure 3.20: Applying a quantity tax.

Then we add a final layer to show the lump sum tax, as shown in Figure 3.21. This enables comparison of the two tax schemes.

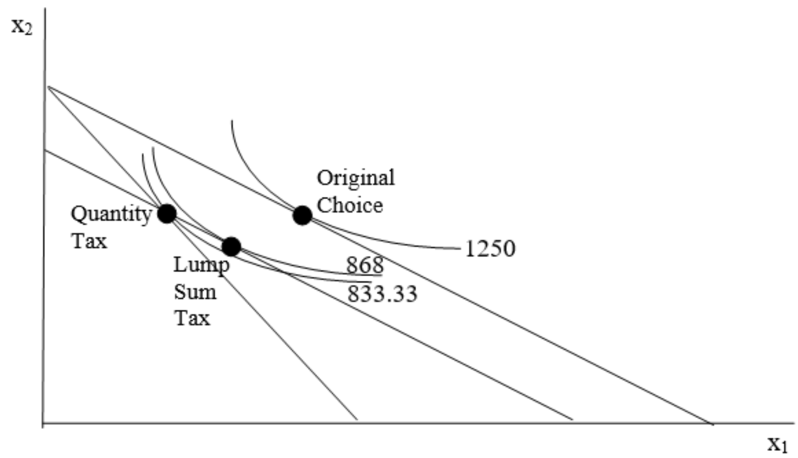


Figure 3.21: Adding a lump sum tax.

The lump sum tax budget constraint has to go through the optimal choice bundle with the quantity tax so that the lump sum tax raises the same revenue as the quantity tax. It also has to be parallel to the original budget constraint. Because it cuts the indifference curve at the quantity tax's optimal solution, we know we can move down the budget line and reach a higher indifference curve than the quantity tax solution.

Figure 3.21 shows that, starting from the *Original Choice* point, we can compare a quantity tax and a revenue-neutral lump sum tax. Figure 3.21 makes clear that the lump sum tax enables attainment of a higher level of utility than the quantity tax because the indifference curve attainable under the lump sum tax is higher than the indifference curve that maximizes utility with the quantity tax.

The reason why the lump sum tax is better is due to the fact that it is *non-distorting*. It leaves the relative prices of the two goods unchanged.

The Lesson and a Follow-up Question

The lesson is that the Theory of Consumer Behavior has been used to show that lump sum taxes are better than quantity taxes. Generating the same amount of revenue, lump sum taxes enable the consumer to reach a higher level of satisfaction than quantity taxes.

This begs a question: Why do we see quantity taxes instead of lump sum taxes? Why are cigarettes (and alcohol and gasoline) so heavily quantity taxed?

The answer lies in the diversity of consumers. The lesson holds only for each individual consumer. It is a fact that there is a revenue-neutral lump sum tax that leaves each individual consumer better off. The amount, however, of the preferable lump sum tax is different, in general, for each consumer. It depends on how many cigarettes (or alcohol or gasoline) each consumer buys. In other words, the lesson does not hold for all consumers taken as a whole. Thus, a single lump tax for all consumers will not necessarily yield higher utility than a quantity tax for each consumer.

This point is obvious if you consider a consumer who does not buy the taxed product at all. This consumer would prefer any size quantity tax to a lump sum tax. After all, if you do not smoke, you do not have to pay any quantity tax on tobacco. The collapse in smoking (see Figure 3.16) goes a long way to explaining why cigarette taxes have soared.

Lump Sum Corollary to the Carte Blanche Principle

We used the Theory of Consumer Behavior to demonstrate a corollary to the *Carte Blanche Principle*: for consumers of a particular product, a lump sum tax is better than a revenue-neutral quantity tax.

If given the option between a quantity and a revenue-neutral lump sum tax, a consumer who buys the taxed good would prefer the lump sum tax because it will leave the consumer with a higher level of utility. Unlike the quantity tax, the lump sum tax will not distort the relative prices faced by the consumer.

Although the *Lump Sum Corollary* is true, we see quantity taxes for various products because the *Lump Sum Corollary* does not apply to all consumers taken as a group. It is not true that there is a single lump sum tax that is preferred to a quantity tax by all consumers.

Exercises

1. Return to the *CigaretteTaxes.xls* workbook and apply a $2/unit quantity tax. Run Solver. Find the solution by evaluating the reduced form. Show your work. Do the two methods agree?

2. Repeat this for the lump sum tax. Find the revenue-neutral solution via Solver, evaluate the reduced form expression at the new *Lump_Tax*, and compare the two methods. Do the two methods agree?

3. Would the percentage change in the consumer's consumption of x_1 be more affected by a quantity tax if her indifference curves were flatter, assuming a Cobb-Douglas utility function? Describe your procedure in answering this question.

References

The epigraph is from the online version of *The Wealth of Nations* by Adam Smith, who is well known as the father of economics. You can access *The Wealth of Nations* (and many other texts) online at www.econlib.org/. If you want a physical copy of the book, used copies abound or you can get a new, inexpensive copy at www.libertyfund.org/.

Cigarette sales data were obtained from Hannah Ritchie and Max Roser (2020) "Smoking," published online at ourworldindata.org/smoking. Our World in Data is a website with striking, thought-provoking data visualizations on a range of topics.

The state tax map is from Frank J. Chaloupka's PowerPoint presentation, "The Truth about Tobacco Economics," available at tobacconomics.org/. Tobacco data, research, and news from around the world makes this a good source for information on smoking and tobacco policy.

The Centers for Disease Control and Prevention publishes data from 1970 on state-level cigarette prices and taxes from the Tax Burden on Tobacco.

In addition to these data sources, the economics literature on cigarette smoking is vast. Frank A. Sloan, V. Kerry Smith, and Donald H. Taylor, "Information, Addiction, and Bad 'Choices': Lessons from a Century of Cigarettes," *Economics Letters*, Vol. 77 (2002), pp. 147-155, is an accessible, informative starting point.

For a broader, historical review, see Allan M. Brandt, *The Cigarette Century: The Rise, Fall, and Deadly Persistence of the Product That Defined America* (2007).

Finally, e-cigarettes are also taxed. Cotti, Chad D., Charles J. Courtemanche, Johanna Catherine Maclean, Erik T. Nesson, Michael F. Pesko, and Nathan Tefft (January 2020), "The Effects of E-Cigarette Taxes on E-Cigarette Prices and Tobacco Product Sales: Evidence from Retail Panel Data," *NBER*, www.nber.org/papers/w26724. From the abstract, "We then calculate an e-cigarette own-price elasticity of -2.6 and a positive cross-price elasticity of demand between e-cigarettes and traditional cigarettes of 1.1, suggesting that e-cigarettes and traditional cigarettes are economic substitutes."

Chapter 4

Comparative Statics

4.1 Engel Curves

The Theory of Consumer Behavior is built on an optimization problem: maximize utility subject to a budget constraint. It is written in equation form like this:

$$\max_{x_1, x_2} U(x_1, x_2)$$
$$\text{s.t. } p_1 x_1 + p_2 x_2 = m$$

This problem can be solved analytically or with numerical methods and the solution can be displayed by a canonical graph, as in Figure 4.1. But it turns out that this is just a first step in how economists think.

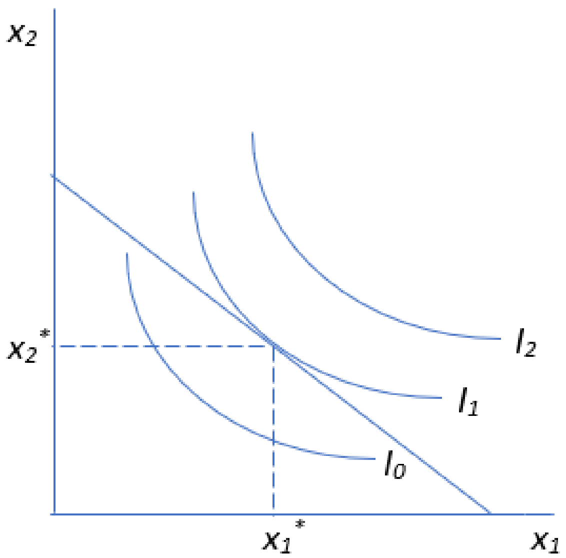


Figure 4.1: Displaying the optimal solution.

The material in this chapter gets to the heart of the economic approach: we explore how the optimal solution responds to a shock, a change in an exogenous variable, holding everything else constant. This is called comparative statics.

The most important comparative statics exercise is based on changing a price, enabling us to derive a demand curve. We start, however, by shocking income and tracking the response. This produces an Engel curve. Starting here gives you a chance to absorb and master the logic of comparative statics before diving into the demand curve.

Initial, Shock, New, Compare

To do comparative statics analysis, we follow a four-step procedure.

1. We find the ***initial*** solution.

2. We change a single exogenous variable, called the ***shock***, holding all other exogenous variables constant. Economists use a Latin phrase, ceteris paribus, as shorthand. This literally means *with other things held equal* and economists use the phrase to mean *everything else held constant*.

3. We find the ***new*** optimal solution.

4. Finally, we ***compare*** the new to the initial solution to see how the optimal solution responded to the shock.

Comparative statics is the fundamental methodology of economics. It gives a framework for interpreting observed behavior. This framework has been given many names, including: the method of economics, the economic approach, the economic way of thinking, and economic reasoning.

While *comparative* clearly points to the comparison between the new and initial solution, the meaning of *statics* (not be confused with statistics) is less obvious. It means that we are going to focus on positions of rest and not worry about the path of the solution as it moves from the initial to the new point.

There are a few complications and additional issues to be aware of when doing comparative statics analysis. Analytical and numerical methods can be used, but they do not always exactly agree. In addition, we have several ways of comparing the new and initial solutions. A qualitative comparison focuses only on direction (up or down), while quantitative comparisons compute magnitudes of the change in response (either as a difference or a percentage change). Finally, we can display the comparative statics analysis in the canonical graph itself or a separate chart. These three issues will be demonstrated via example.

Elasticity Basics

Elasticity is a pure number (it has no units) that measures the sensitivity or responsiveness of one variable when another changes. Elasticity, responsiveness, and sensitivity are synonyms. An elasticity number expresses the impact one variable has on another. The closer the elasticity is to zero, the more insensitive or inelastic the relationship.

Elasticity is often expressed as "the something elasticity of something," like the price elasticity of demand. The first something, the price, is always the exogenous variable; the second something, in this case demand (the amount purchased), is the response or optimal value being tracked.

A less common, but perhaps easier, way is to say, "the elasticity of something with respect to something." The elasticity of demand with respect to price clearly shows that demand depends on and responds to the price.

Unlike the difference between the new and initial values, elasticity is computed as the ratio of percentage changes in the values. The endogenous or response variable always goes in the numerator and the exogenous or shock variable is always in the denominator.

The percentage change, $\frac{new - initial}{initial}$, is the change (or difference), $new - initial$, divided by the initial value. This affects the units in the computation. The units in the numerator and denominator of the percentage change cancel and we are left with a percent as the units. If we compute the percentage change in apples from 2 to 3 apples, we get 50%. The change, however, is +1 apple.

If we divide one percentage change by another, the percents cancel and we get a unitless number. Thus, elasticity is a pure number with no units. So if the price elasticity of demand for apples is −1.2, there are no apples, dollars, percents, or any other units. It's just −1.2.

The lack of units in an elasticity measure means we can compare wildly different things. No matter the underlying units of the variables, we can put the dimensionless elasticity number on a common yardstick and interpret it. Figure 4.2 shows the possible values that an elasticity can take, along with the names we give particular values.

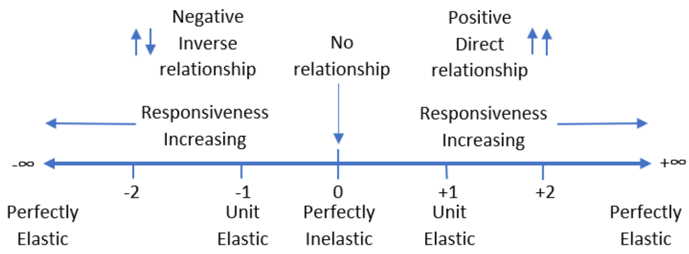


Figure 4.2: Elasticity on the number line.

Empirically, elasticities are usually low numbers around one (in absolute value). An elasticity of +2 is extremely responsive or elastic. It means that a 1% increase in the exogenous variable generates a 2% increase in the endogenous variable.

The sign of the elasticity indicates direction (a qualitative statement about the relationship between the two variables). Zero means that there is no relationship—i.e., that the exogenous variable does not influence the response variable at all. Thus, −2 is extremely responsive like +2, but the variables are inversely related so a 1% *increase* in the exogenous variable leads to a 2% *decrease* in the endogenous variable.

One (both positive and negative) is an important marker on the elasticity number line because it tells you if the given percentage change in an exogenous variable results in a smaller percentage change (when the elasticity is less than one), an equal percentage change (elasticity equal to one), or greater percentage change (elasticity greater than one) in the endogenous variable.

Elasticities are a confusing part of economics. Below are six common misconceptions and issues surrounding elasticity. Reading these typical mistakes will help you better understand this fundamental, but easily misinterpreted concept.

1. Elasticity is about the *relationship* between two variables, not just the change in one variable. Thus, do not confuse a negative elasticity as meaning that the response variable must decrease. The negative means that the two variables move in opposite directions. So, if the age elasticity of time playing sports is negative, that means both that time playing sports falls as age increases and time playing sports rises as age decreases.

2. Elasticity is a *local phenomenon*. The elasticity will usually change if we analyze a different initial value of the exogenous variable. Thus, any one measure of elasticity is a local or point value that applies only to the change in the exogenous variable under consideration from that starting point. You should not think of a price elasticity of demand of -0.6 as applying to an entire demand curve. Instead, it is a statement about the movement in price from one value to another value close by, say \$3.00/unit to \$3.01/unit. The price elasticity of demand from \$4.00/unit to \$4.01/unit may be different. There are constant elasticity functions, where the elasticity is the same all along the function, but they are a special case.

3. Elasticity can be calculated for *different size changes*. To compute the x elasticity of y, we can go from one point to another, $\frac{\%\Delta y}{\%\Delta x}$, or use the derivative's infinitesimally small change at a point, $\frac{dy}{dx}\frac{x}{y}$. These formulas will be explained below, but the point now is that economists are sloppy in their language and do not bother to distinguish elasticity calculated at a point via calculus (for an infinitesimal change) and elasticity calculated for a finite distance from one point to another. If the function is nonlinear, these two methods give different results. If an economist mentions a point elasticity, it is probably calculated via calculus as an infinitesimally small change.

4. Elasticity always puts the *response variable in the numerator*. Do not confuse the numerator and denominator in the computation. In the x elasticity of y, x is the exogenous or shock variable and y is the endogenous or response variable. Students will often compute the reciprocal of the correct elasticity. Avoid this common mistake by always checking to make sure that the variable in the numerator responds or is driven by the variable in the denominator.

5. You already know this, but remember that elasticity is *unitless*. The x elasticity of y of 0.2 is not 20%. It is 0.2. It means that a 1% increase in x leads to a 0.2% increase in y.

6. Perhaps the single most important thing to remember about elasticity is: *Do not confuse elasticity with slope.* This may be the most common confusion of all and deserves careful consideration.

Economists, unlike chemists or physicists, often gloss over the units of variables and results. If we carefully consider the units involved, we can ensure that the difference between the slope and elasticity is crystal clear.

The slope is a quantitative measure in the units of the two variables being compared. If $Q^* = \frac{P}{2}$, then the slope, $\frac{dQ^*}{dP} = \frac{1}{2}$. This says that an increase in P of \$1/unit will lead to an increase in Q^* of $\frac{1}{2}$ a unit. Thus, the slope would be measured in units squared per dollar (so that when multiplied by the price, we end up with just units of Q).

Elasticity, on the other hand, is a quantitative measure based on percentage changes and is, therefore, unitless. The P elasticity of $Q^* = 1$ says that a 1% increase in P leads to a 1% increase in Q^*. It does not say anything about the actual, numerical \$/unit increase in P, but speaks of the percentage increase in P. Similarly, elasticity focuses on the percentage change in Q^*, not the change in terms of number of units.

Thus, elasticity and slope are two different ways to measure the responsiveness of a variable as another variable changes. Elasticity uses percentage changes, $\frac{\%\Delta y}{\%\Delta x}$, while the slope does not, $\frac{\Delta y}{\Delta x}$. They are two different ways to measure the effect of a shock and mixing them up is a common mistake.

Comparative Statics Analysis of Changing Income

STEP Open the Excel workbook *EngelCurves.xls*, read the *Intro* sheet, and proceed to the *OptimalChoice* sheet.

We have run Solver and the initial solution, $x_1{}^* \approx 25$ and $x_2{}^* \approx 16\frac{2}{3}$, is displayed.

Our first attempt at comparative statics analysis is straightforward: change income, ceteris paribus, and compute the response in $x_1{}^*$ and $x_2{}^*$.

STEP Change cell B18 to 150 (this is the shock) and then run Solver to find the new optimal solution.

The budget line shifts out and the consumer takes advantage by re-optimizing and moving to a new, highest attainable indifference curve.

STEP Compare the initial and new values of x_1* and x_2* given the \$50 increase in income.

In qualitative terms, we would say that the increase in income has led to an increase in optimal consumption of the two goods.

In quantitative terms, we can compute the response as the change in the own units of the two variables.

The *own units* statement of comparative statics for x_1* is $\frac{\Delta x_1{}^*}{\Delta m}$.

Income rose by \$50 and optimal consumption of each good went up by 12.5 units. We compute $\frac{37.5-25}{150-100}$ so we say that we get an increase of $\frac{1}{4}$ unit for every \$1 increase in income.

Elasticity is another a way to present a quantitative comparative statics result. We use a formula that multiplies the slope by the initial values.

Income elasticity of $x_1{}^* = \frac{\Delta x_1{}^*}{\Delta m}\frac{m}{x_1{}^*} = [\frac{37.5-25}{150-100}][\frac{100}{25}] = 1$. This elasticity is unit elastic. This means that a 1% change in income leads to a 1% change in the optimal purchase of good 1. We had a 50% increase to income and that produced a 50% increase in x_1*.

The elasticity formula seems mysterious, but it is easily derived from the definition of the ratio of percentage changes.

$$\frac{\%\Delta x_1{}^*}{\%\Delta m} = \frac{\frac{\Delta x_1{}^*}{x_1{}^*}}{\frac{\Delta m}{m}} = \frac{\Delta x_1{}^*}{x_1{}^*}\frac{m}{\Delta m} = \frac{\Delta x_1{}^*}{\Delta m}\frac{m}{x_1{}^*}$$

The algebra above shows how slope and elasticity are connected. Multiplying the slope by an initial position is the same as computing a percentage change.

While it is certainly possible to do comparative statics analysis by running Solver to find the initial solution, changing a parameter on the sheet, running Solver again to find the new solution, and then comparing the initial and new solutions, the tediousness of this manual approach is obvious.

Fortunately, there is a better way. It involves using the Comparative Statics Wizard Excel add-in.

STEP Click the Reset button to make sure you start from the initial parameter values.

STEP Install the Comparative Statics Wizard add-in, *Cswiz.xla*, from the *MicroExcel* archive.

Instructions and documentation are available in the *CompStatics.doc* file in the *SolverCompStaticsWizard* folder. You can see which add-ins are installed by accessing the Add-ins Manager dialog (In Excel 2019, File: Excel Options: Add-ins: Go).

STEP Once the Comparative Statics Wizard add-in is installed, from the *OptimalChoice* sheet, click the Add-ins tab on the Ribbon, then click Wizard and Comp Statics (in earlier versions, execute Tools: Wizard: Comp Statics) to bring up the main dialog box of the CSWiz add-in, shown in Figure 4.3.

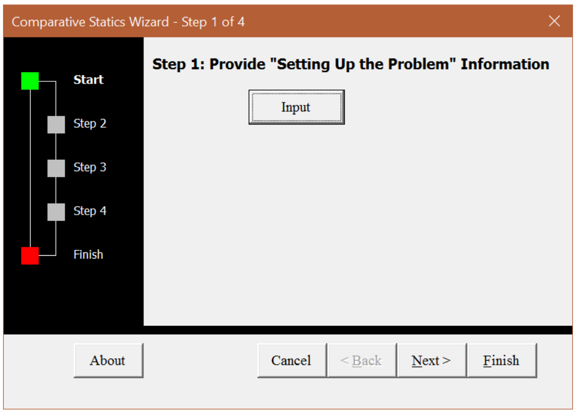


Figure 4.3: First step in the Comparative Statics Wizard.

STEP Click on the Input button and answer the three questions posed.

You are providing Excel with the information it needs to organize the results. Clearly, the goal is cell B7 so you will click on cell B7 when prompted by the first question. Excel enters the absolute reference to that cell ($B$7) in the dialog box and you click OK. Follow the same procedure for the next two questions. The endogenous variables are in cells B11:B12 and the exogenous variables are in cells B16:B20 so can click and drag to select those cells.

Notice how the Comparative Statics Wizard add-in presumes that you have properly organized and set up the problem on the spreadsheet.

STEP Once you have provided the goal, endogenous and exogenous variable cells, click the Next button.

Step 2 uses Excel's Solver to find the initial solution. It temporarily hides the Comparative Statics Wizard and brings up Solver so you can use it to find the optimal solution.

STEP At the Step 2 screen, click the Run Solver button to bring up the Solver dialog box. Click Solve to have Solver find the initial solution.

Read the message in the box after you have run Solver. It explains what you have done so far.

Having found the initial solution, we are ready to input the shock.

STEP At the Step 3 screen, click the Input button.

As in the first screen, you are asked three questions. The first question asks for the shock variable itself. In this case, click on cell B18 (the income variable value, not the label). The second question is the amount of change. Enter 50. The third question is the number of shocks. The default value is 5. Accept this value by clicking the OK button.

You have asked Excel to change income, holding the other variables constant, from 100 to 150 to 200 to 250 to 300 to 350—five jumps of 50 each from the 100 initial value.

STEP After verifying that you have entered the shock information correctly, click the Next button to continue.

The Step 4 screen is the heart of the add-in. You have provided the goal, endogenous and exogenous variable information, Solver found the initial solution, and you have told Excel which variable to shock and how. Excel is ready to run the problem over and over again for each of the shock variable values you provided. It is essentially the manual approach, but Excel does all of the tedious work.

STEP Click the Run Comparative Statics button. The bar displays Excel's progress through the repeated optimization problems. It runs Solver at each value of income, but it is very fast.

STEP Click the Next button, read the information in the box, and click the Finish button.

Excel takes you to a sheet it has inserted into the workbook with all of the comparative statics results. This sheet is similar to the *CS1* sheet. Notice how the results are arranged. It begins with the initial parameter values (widen column A if needed), then displays a table with income in column A, followed by maximum utility and the optimal values of the two goods.

The results produced by the Comparative Statics Wizard can be further processed as shown in the *CS1* sheet.

STEP Proceed to the *CS1* sheet. Columns F and G contains slope and elasticity calculations. Click on the cells to see the formulas.

Notice that you have to be careful with parentheses when doing percentage change calculations in Excel. Simply entering "= C14 – C13/C13" will not do what you want because Excel's order of operations rule will divide C13 by C13 (which is 1) and subtract that from C14.

Income Consumption and Engel Curves

There are two graphs on the *CS1* sheet. They appear to be the same, but they are not. One graph is an income consumption curve and the other is an Engel curve. They are related and understanding their connection is important.

Ernst Engel (not to be confused with Karl Marx's benefactor and friend, Friedrich Engels) was a 19^{th} century German statistician who analyzed consumer expenditure data. He found that food purchases increased as income rose, but at a decreasing rate. This became known as *Engel's Law.* A graph of quantity demanded for a good as a function of income, ceteris paribus, is called an *Engel curve.*

The *income consumption curve* (ICC) shows the effect of the increase in income in the canonical indifference-curves-and-budget-constraint graph. In other words, the ICC shows the comparative statics analysis on the underlying, canonical graph. Panel A in Figure 4.4 shows the income consumption curve.

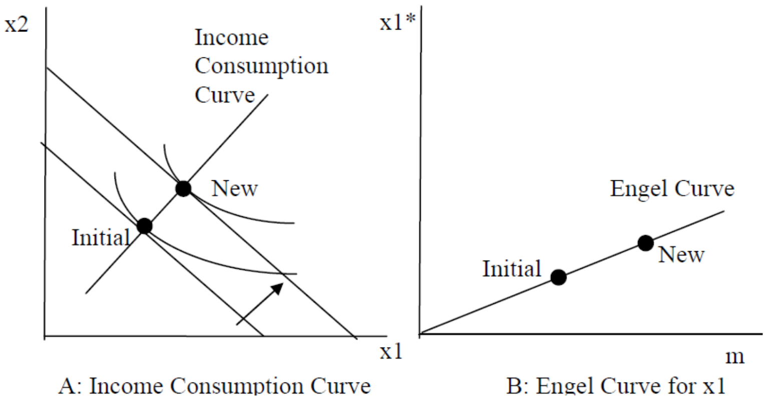


Figure 4.4: Displaying the results of a shock in income.

Panel B shows that the Engel curve for x_1 plots the relationship between income and optimal x_1. This presentation graph shows only the optimal value of the endogenous variable (x_1) as a function of the shock variable (m) and hides everything else. There is an Engel curve graph for x_2, but it is not displayed.

STEP Use your comparative statics results to make Engel and income consumption curves. This will help you understand the relationship between the two curves.

For the Engel curve, select data in m (in column A) and x_1 (in column C). For the ICC, you need to select x_1 and x_2 (in columns C and D). After selecting the data, click the Insert tab in the Ribbon and choose the Scatter chart type in the Charts group.

The slope of the Engel curve reveals if the good is normal or inferior. A *normal good*, as in Figure 4.4, has a positively sloped Engel curve: when income rises, so does optimal consumption. An inferior good has a negatively sloped Engel curve, increases in income lead to decreases in optimal consumption of the good. Figure 4.5 shows this case.

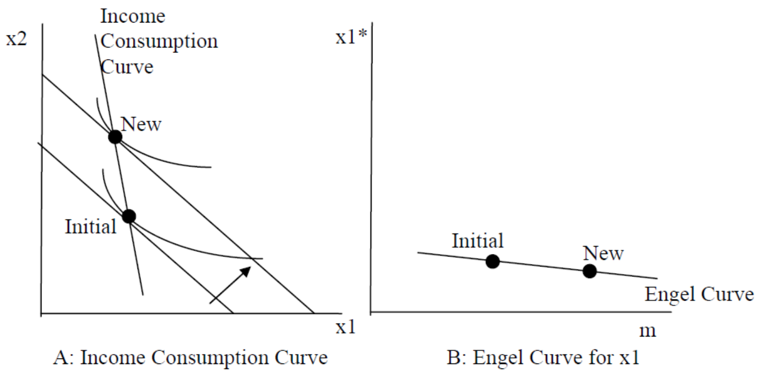


Figure 4.5: x_1 as an inferior good.

Hamburger is the classic inferior good example. As income rises, the idea is that you eat less hamburger meat and more of better cuts of beef. The example also serves to point out that goods are not either normal or inferior due to some innate characteristic, but that the relationship is a local phenomenon. Figure 4.6 shows how a consumer might react across the full range of income. Do you understand the story this graph is telling?

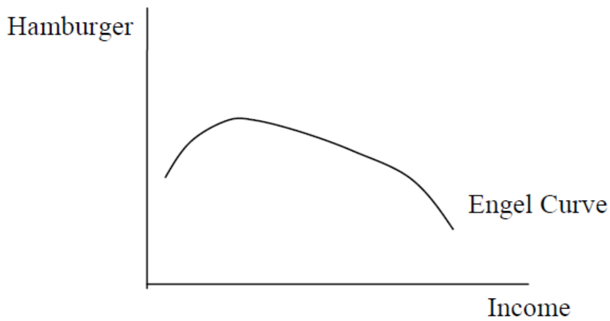


Figure 4.6: A hypothetical Engel curve for hamburger.

Figure 4.6 shows that hamburger is normal at low levels of income (with increasing consumption as income rises), but inferior at higher levels of income. Our Cobb-Douglas utility function cannot generate this complicated Engel curve.

Analytical Comparative Statics Analysis of Changing Income

We can derive the Engel curve for the problem in the *EngelCurves.xls* workbook via analytical methods.

As usual, we rewrite the constraint and form the Lagrangean, then take derivatives, and solve the system of equations. The novelty this time is that we leave m as a letter so that our final answer is a function of income. This enables us to derive an Engel curve.

1. Rewrite the constraint so that it is equal to zero.

$$0 = m - 2x_1 - 3x_2$$

2. Form the Lagrangean function.

$$\max_{x_1, x_2, \lambda} L = x_1 x_2 + \lambda(m - 2x_1 - 3x_2)$$

We take derivatives and set them equal to zero.

$$\frac{\partial L}{\partial x_1} = x_2^* - 2\lambda^* = 0$$
$$\frac{\partial L}{\partial x_2} = x_1^* - 3\lambda^* = 0$$
$$\frac{\partial L}{\partial \lambda} = m - 2x_1^* - 3x_2^* = 0$$

To solve for the optimal values of x_1 and x_2, move the lambda terms in the top two equations to the right-hand side and divide the first equation by the second to eliminate lambda (and give the familiar MRS = $\frac{p_1}{p_2}$ condition. Then solve for optimal x_2 in terms of x_1.

$$\frac{x_2^*}{x_1^*} = \frac{2}{3}$$
$$x_2^* = \frac{2}{3} x_1^*$$

Substitute this expression for x_2 into the third first-order condition and solve for optimal x_1.

$$m - 2x_1^* - 3\left[\frac{2}{3}x_1^*\right] = 0$$
$$4x_1^* = m$$
$$x_1^* = \frac{1}{4}m$$

We can evaluate this expression at any value for m. If we substitute in $m = 100$, we get $x_1{}^* = 25$ which is what we got when we solved this problem with an income of $100.

Our reduced form expression for $x_1{}^*$ agrees with the values in columns A and C of the *CS1* sheet that we produced via the numerical approach using the Comparative Statics Wizard. The numerical method picks individual points off the Engel curve function that we derived here.

There is also an Engel curve for $x_2{}^*$. It is $x_2{}^* = \frac{1}{6}m$.

Of course, these Engel curves are for this particular consumer, with this particular utility function and set of exogenous variables. Different preferences will give different Engel curves.

If we make the problem more general, in the sense of substituting letters for numbers in the Lagrangean, then these exogenous variables will appear in the reduced form expression. In other words, the one-quarter and one-sixth constants in the Engel curves will be changed into an expression with the exogenous variables. Evaluating that expression at the current values of the exogenous variables will give one-quarter and one-sixth.

If you change an exogenous variable other than income, you will no longer move along the Engel curve. Instead, you will shift the entire Engel curve.

To compute an own units response in $x_1{}^*$ given a change in income, we can simply take the derivative with respect to m, which is simply $\frac{1}{4}$. This means the slope of the reduced form is constant at any value of m.

The elasticity at a given value of m can be computed via the following formula:

$$\frac{dx_1{}^*}{dm}\frac{m}{x_1{}^*}$$

Because it is calculated *at* a particular point, this is called *point elasticity*, as opposed to an elasticity measured from one point to another. Economists usually compute and report point elasticities, but they often omit the adjective and simply call the result an elasticity.

Notice how the point elasticity formula is similar to the elasticity formula from one point to another, $\frac{\Delta x_1{}^*}{\Delta m}\frac{m}{x_1{}^*}$. We have simply replaced the delta with a d—this shows that the two formulas are the same except for the size of the change in m. Instead of a discrete-size change, the point elasticity formula is based on an infinitesimally small change in m.

At $m = 100$, the point income elasticity of $x_1{}^* = (\frac{1}{4})(\frac{100}{25}) = 1$. Good x_2 also has a constant unit income elasticity. Rays from the origin always have constant unit elasticities.

The utility function plays a crucial role in comparative statics outcomes. Cobb-Douglas utility functions always yield linear Engel curves with constant unit income elasticities. We do not believe that, in the real world, Engel curves are always linear and unit income elastic. While there are other utility functions with less restrictive results, they are more difficult to work with mathematically. Ease of algebraic manipulation helps explain the popularity of the Cobb-Douglas functional form.

An Engel Curve is Comparative Statics Analysis

This chapter introduced comparative statics analysis. It focused on tracking the optimal solution as income changes. This is called an Engel curve.

Comparative statics analysis, including elasticities, can be done via numerical and analytical methods. The Comparative Statics Wizard handles much of the tedious work in the numerical approach.

We can compute an elasticity in two ways: *at* a point and *from* one point to another. The former uses the derivative and latter is based on a discrete-size change in the exogenous variable. A point elasticity is one based on the derivative. Both elasticities are based on percentage changes, but the derivative uses infinitesimally small changes in the exogenous variable.

We will often compare the two methods. In this case, the two methods agreed perfectly. This will not always be true.

Exercises

1. Change the price of good 1 from 2 to 3 in the *OptimalChoice* sheet of the *EngelCurves.xls* workbook. From $m = 100$, use the Comparative Statics Wizard to create a graph of the Engel curve for good 1. Title the graph and label the axes. Take a picture of your graph and paste it in your Word document.

2. Why is the slope of your graph different than the one in the *CS1* sheet?

3. Compute the income elasticity of demand for good 1 from $m = 100$ to 200. Show your work.

4. Compute the income elasticity of demand for good 1 at $m = 100$. Show your work.

5. Why are your answers in question 3 and 4 the same?

References

The epigraph is from H. S. Houthakker, "Engel's Law," in J.Eatwell, M.Milgate and P.Newman (eds.) *The New Palgrave Dictionary of Economics*, (London: McMillan, 1987), pp. 143-144.

The *Palgrave* is much more than a simple dictionary. It is a reference resource with articles on specific terms or phrases. The 2008 version of the Palgrave Dictionary is edited by Stephen N. Durlauf and Lawrence E. Blume. It is available online at www.dictionaryofeconomics.com.

I shall also argue that the most secure propositions and the most reliable predictions, even though they are conditional predictions, arise out of comparative statics, and that when we are asked the awkward question "what good is economics to anyone," apart from its usefulness in providing a gainful occupation for economists, the defense rests mainly on the achievements of rather old-fashioned comparative statics.

Kenneth E. Boulding

4.2 More Practice with Engel Curves

This section derives Engel curves via numerical and analytical methods for different utility functions. It applies the same logic as the previous chapter. This is mastery by repetition. Recognizing how the same steps are used is essential to thinking like an economist.

Quasilinear Preferences

This example uses a quasilinear utility function, $U = x_1^{\frac{1}{2}} + x_2$. The budget constraint is $140 = 2x_1 + 10x_2$.

We begin with the analytical approach. We rewrite the constraint and form the Lagrangean, leaving m as a letter (since we want to derive an Engel curve).

$$\max_{x_1, x_2, \lambda} L = x_1^{1/2} + x_2 + \lambda(m - 2x_1 - 10x_2)$$

We take derivatives and set them equal to zero.

$$\frac{\partial L}{\partial x_1} = \frac{1}{2}x_1^{-1/2} - 2\lambda = 0$$
$$\frac{\partial L}{\partial x_2} = 1 - 10\lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = m - 2x_1 - 10x_2 = 0$$

To solve for the optimal values of x_1 and x_2, we follow our usual approach, moving the λ terms over to the right-hand side and dividing the two equations to cancel the λs.

$$\frac{1}{2}x_1^{-1/2} = 2\lambda$$
$$1 = 10\lambda$$
$$\frac{\frac{1}{2}x_1^{-1/2}}{1} = \frac{2\lambda}{10\lambda}$$
$$\frac{\frac{1}{2}x_1^{-1/2}}{1} = \frac{2}{10}$$

Notice that the MRS is a function of x_1 alone. This is a property of the quasilinear utility function. We can solve for $x_1{}^*$ from the MRS equal to the price ratio equation.

$$\frac{\frac{1}{2}x_1^{-1/2}}{1} = \frac{2}{10}$$
$$\left[x_1^{-1/2}\right]^{-2} = \left[\frac{4}{10}\right]^{-2}$$
$$x_1^* = 6.25$$

Next, we plug this value into the third first-order condition and solve for $x_2{}^*$.

$$m - 2[6.25] - 10x_2 = 0$$
$$10x_2 = m - 12.5$$
$$x_2^* = \frac{1}{10}m - 1.25$$

To compute an own units response in $x_1{}^*$ given a change in m, we can simply take the derivative with respect to m, which is zero (because m does not appear in the $x_1{}^*$ reduced form). Thus, increases in income leave $x_1{}^*$ unchanged. In other words, the Engel curve for good 1 is horizontal at 6.25.

The own units response for $x_2{}^*$ is $\frac{dx_2{}^*}{dm}\frac{m}{x_2{}^*} = \frac{1}{10}$. This means that an additional dollar in income leads to a $\frac{1}{10}$ increase in good 2.

We can use the income elasticity formula, $\frac{dx_1{}^*}{dm}\frac{m}{x_1{}^*}$, to compute the income elasticity. At $m = 140$, the income elasticity of $x_1{}^* = (0)(140/6.25) = 0$, which is perfectly inelastic. This means that changes in m have no effect at all on $x_1{}^*$.

These results seem a little strange. Perhaps the numerical approach and Excel can shed some light on what's going on here.

STEP Open the Excel workbook *EngelCurvesPractice.xls*, read the *Intro* sheet, then go to the *QuasilinearChoice* sheet. It shows the optimal solution, 6.25, 12.75, for $m = 140$. Change income to 160.

As expected the budget line shifts out.

STEP Run Solver to find the new initial solution. The resulting chart looks like Figure 4.7.

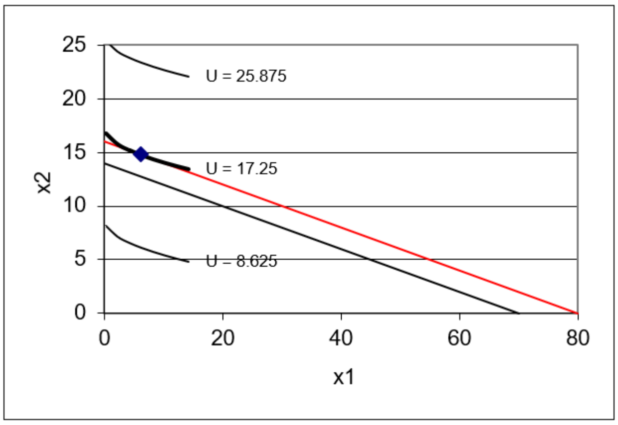


Figure 4.7: Income shock with quasilinear preferences.

Figure 4.7 and your screen show that the value of $x_1{}^*$ remained unchanged as income rose from \$140 to \$160. This consumer maximizes utility by using all of the extra \$20 in income on good 2.

Figure 4.7 also displays a key property of the quasilinear functional form: the indifference curves are vertically shifted and actually parallel to each other. Thus, when we increase income, the new point of tangency is found directly, vertically up from the original solution.

STEP Return income to its initial value of $140. Run the Comparative Statics Wizard, applying 5 shocks to income in $10 dollar increments.

Your results should look like the *CS1* sheet.

STEP Create Engel and income consumption curves. For the Engel curves, this requires making a chart of x_1* as a function of m and another chart of x_2* as a function of m. For the income consumption curve, the chart is x_2* as a function of x_1*. Each point on this chart is a point of tangency between the budget line and maximum attainable indifference curve.

Your first attempt at making a chart of x_1* as a function of m will not yield a horizontal line at 6.25. Look closely, however, at the y axis scale. The problem is that Solver is reporting numbers very close to, but not exactly, 6.25 as income changes.

But these slight differences in optimal x_1 are not meaningful. They are Solver noise. In fact, for all of these values of m, optimal x_1 really is exactly 25. We need to clean up Solver's results.

Simply changing the display to fewer decimals will not work. This will change the display of the y axis, but Excel will still have the same number in its memory. Instead, we have to use Excel's ROUND function to change the numbers produced by Solver.

The ROUND function has two arguments, the cell you want to round and the number of decimal places. So, ROUND(123.456,1) evaluates to 123.5.

STEP Enter this formula in a blank cell, "=ROUND(123.456,-2)" to see what a negative argument does.

We can use the ROUND function to round Solver's results to the hundredths place. Cell F12 shows how this strategy is implemented.

STEP Apply Excel's Round function to your comparative statics results and then make a chart of the Engel curve for good 1 using the rounded data. Your final chart should look like the one in the *CS1* sheet.

Finally, we can use the CSWiz results to examine the responsiveness of the endogenous variables to the changes in income we applied.

STEP Compute the response to the income changes in own units and income elasticities for x_1* and x_1*. Check your work with the results in the *CS1* sheet.

Notice that the responsiveness results from the numerical method are the same as that via the analytical approach.

Perfect Complements

STEP Proceed to the *PerfCompChoice* sheet to practice on another utility function. This function reflects preferences in which the two goods are perfect complements. This gives L-shaped indifference curves, but our analysis proceeds as usual.

The problem is to maximize the perfect complements utility function subject to the budget constraint. The *PerfCompChoice* sheet shows that $p_1 = 2, p_2 = 10, a = b = 1$.

We do the problem first via the analytical method, leaving m as a letter so we can find $x_1* = f(m)$ and $x_2* = f(m)$—these are Engel curves for goods 1 and 2.

In section 3.2, we showed how to solve this problem by finding the intersection of two lines on which the solution must lie. Since $a = b = 1$, the optimal solution must be where $x_1 = x_2$ (a ray from the origin with slope +1). Of course, the solution must also lie on the budget line, so we can solve this system of two equations and two unknowns by substituting in x_1 for x_2 in the budget constraint equation.

$$\left.\begin{array}{l} x_2 = x_1 \\ m = 2x_1 + 10x_2 \end{array}\right\} \Rightarrow m = 2x_1 + 10[x_1] \Rightarrow m = 12x_1 \Rightarrow x_1^* = \frac{m}{12}$$

Since x_2 must equal x_1 at the optimal solution, we know $x_2* = \frac{m}{12}$.

To compute an own units response in x_1* given a change in income, we can simply take the derivative with respect to m, which is simply $\frac{1}{12}$. This slope is constant and the Engel curve is linear.

The income elasticity at a given value of m can be computed via the point elasticity formula, $\frac{dx_1*}{dm}\frac{m}{x_1*}$. At $m = 50$, the income elasticity of $x_1* = \frac{1}{12}\frac{50}{4.167} = 1$. This means that a 1% change in m will result in a 1% change in x_1*.

STEP Run the Comparative Statics Wizard on the *PerfCompChoice* sheet (you can make the change in income $10) and create Engel and income consumption curves.

STEP Compute the response to the income changes in own units and income elasticities for x_1* and x_2*.

Check your work with the results in the *CS2* sheet. Notice that the results in Excel are the same as the analytical approach.

The Utility Function Determines the Shape of the Engel Curve

This section ran a comparative statics analysis of a change in income on quasilinear and perfect complement utility functions. This enabled practice in deriving Engel curves and income consumption curves, along with computing responsiveness in own units and elasticities.

The quasilinear function has the peculiar result that the income elasticity of x_1* is zero. This happens because the indifference map of a quasilinear utility function is a series of vertically parallel curves. Thus, when the budget line shifts out, the new optimal solution is found directly above the initial solution and x_1* remains unchanged.

With the perfect complements utility function, we were able to find an analytical solution even though we could not use the Lagrangean method. The Engel curve for x_1* has a constant slope and a unit income elasticity. These are the same properties for the Engel curve we found in the previous chapter using the Cobb-Douglas functional form.

The shape of the Engel curve, its slope and income elasticity are all influenced by the consumer's utility function. The relationship is complicated, so there is no rule or simple statement about how the functional form of utility determines the Engel curve.

Ernst Engel wanted to know how spending on food changed as income rose. He believed food purchases would increase at a decreasing rate as income increased, as shown in Figure 4.8. This makes common sense. As you get richer and richer, you can buy a much nicer house and cars, but it is difficult to spend a lot more on food. This is known as *Engel's Law.*

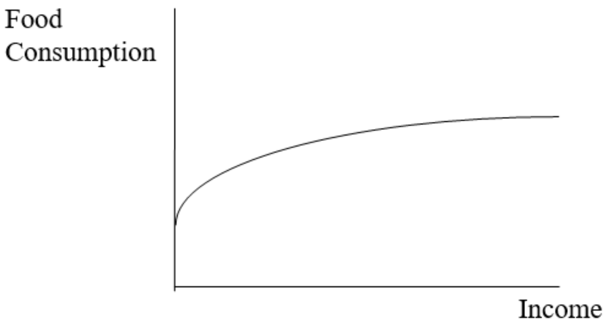


Figure 4.8: Engel's Law.

None of three utility functions we have encountered thus far (Cobb-Douglas, quasilinear, and perfect complements) are capable of generating an Engel curve that conforms to Engel's Law for food purchases. If we were interested in food, we would have to find and use a utility function with an Engel curve that conformed to Engel's Law. Such functions exist, but as you can imagine, they are more complicated than the computationally simple functions we have used thus far.

Exercises

1. In the *QuasilinearChoice* sheet, copy cell B11 and paste it in cell C11. Set income to $200 and run Solver to find the new optimal solution. In cell D11, enter a formula to find the difference between cell C11 and B11. Is this tiny difference meaningful? Explain.

2. Having changed income and run Solver in question 1, if you connected the initial and new solutions on the chart, you would get a vertical line. Why is this happening? Will this happen with every consumer?

3. Having changed income and run Solver in question 1, is good 1 a normal or an inferior good? Explain.

4. Use Word's Equation Editor to solve the general version of the perfect complements problem. In other words, find x_1* and x_2* for

$$\max_{x_1,x_2} U = \min\{ax_1, bx_2\}$$
$$\text{s.t. } m = p_1x_1 + p_2x_2$$

References

The epigraph is from pages 487 and 488 of Kenneth E. Boulding, "In Defense of Statics," *The Quarterly Journal of Economics*, Vol. 69, No. 4 (November, 1955), pp. 485–502 (www.jstor.org/stable/1881991). As you can tell from the quotation, Boulding had a well-deserved reputation for witty, biting comments. His defense of comparative statics in the article just cited notwithstanding, he once quipped, "Mathematics brought rigor to Economics. Unfortunately, it also brought mortis."

The first "empirical" demand schedule was published in 1699 by Charles Davenant.

George Stigler

4.3 Deriving a Demand Curve

We know how to find the initial optimal solution in the Theory of Consumer Behavior and we have explored the comparative statics properties of a change in income.

We are well prepared to embark on the most important comparative statics analysis in the Theory of Consumer Behavior: deriving a demand curve.

Numerical Comparative Statics Analysis of Changing Price

STEP Open the Excel workbook *DemandCurves.xls* and read the *Intro* sheet,then go to the *OptimalChoice* sheet.

The problem is set up, but the consumer is not optimizing because the MRS does not equal the price ratio and the consumer can move to higher indifference curves by traveling up the constraint.

STEP Run Solver to find the initial solution: $x_1* = 25$ and $x_2* = 16\frac{2}{3}$.

Next, we explore how this initial optimal solution changes as the price of good 1 changes, ceteris paribus. This comparative statics analysis will produce a demand curve.

Before we actually do it, can you anticipate what will happen when we increase the price of good 1? Believe it or not, if you try to figure it out first—before actually seeing it—you will learn more. Take a moment to think: what will happen to the graph on your screen when we increase the price of x_1?

Let's see how you did.

STEP Shock: Change cell B16 to 3.

Figure 4.9 shows how your screen should look. With a higher p_1,the budget constraint rotates in, pivoting on the x_2 intercept. The consumer now has fewer consumption possibilities and needs to re-optimize to find the new optimal solution.

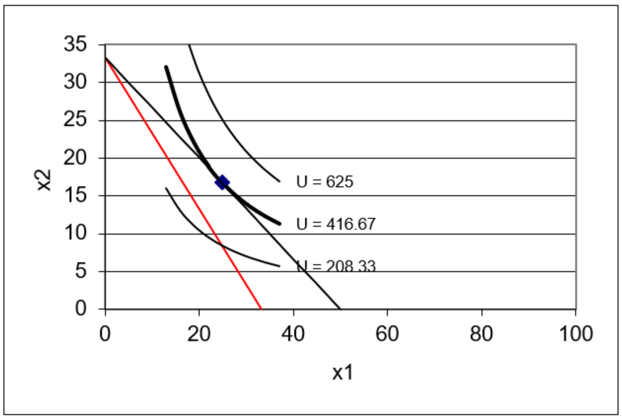


Figure 4.9: New budget line when p_1 rises.

STEP New: Run Solver to find the new optimal solution.

We have completed initial, shock, and new—the last step is to compare. Figure 4.10 shows a table that displays the comparative statics results.

p_1	x_1^*	x_2^*	$\Delta x_1^*/\Delta p_1$	$\%\Delta x_1^*/\%\Delta p_1$	$\Delta x_2^*/\Delta p_1$	$\%\Delta x_2^*/\%\Delta p_1$
2	25	$16\frac{2}{3}$				
3	$16\frac{2}{3}$	$16\frac{2}{3}$	$-8\frac{1}{3}$	-0.67	0	0

Figure 4.10: Comparative statics results of an increase in p_1.

In qualitative terms, we can see that x_1* falls as p_1 rises, but x_2* remains unchanged.

Quantitatively, we can compute the own units response in good 1 as new minus initial x_1*, which is $16\frac{2}{3} - 25 = -8\frac{1}{3}$ divided by 1 (from $3-2$). This is the value displayed in the table. The own units response in x_2 is zero since it did not change.

Responsiveness in percentage terms is the price elasticity of demand. We need to compute the percentage change in $x_1{}^*$ divided by the percentage change in p_1. The numerator is -33% because $\frac{16\frac{2}{3}-25}{25} = -\frac{1}{3}$. The denominator is $\frac{3-2}{2} = 0.5$ or 50%. So, a 50% increase in price, from $p_1 = 2$ to 3, caused a 33% decrease in quantity demanded. Thus the price elasticity of demand is $\frac{-0.33}{0.5} = -\frac{2}{3}$ or roughly -0.67. This number is displayed in the table in Figure 4.10.

The same calculation can be performed on x_2. Since we are considering the effect on *good 2* from a shock to the price of *good 1*, we call this a *cross price* analysis. The term *cross* is used in economics when we examine the effect of i on j; an *own* effect, for example, would be p_1 on x_1.

We quickly realize that the cross price elasticity, the p_1 elasticity of x_2, is zero because the numerator is zero. This is perfectly inelastic or completely unresponsive.

Comparative statics via numerical methods is easier with the Comparative Statics Wizard add-in. If it is not installed, return to the beginning of this chapter to load the CSWiz add-in.

STEP Analyze the effect of a change in p_1 by running the CSWiz add-in and changing the price of good 1 by \$1 increments (for five shocks).

You can see a slightly different comparative statics analysis in the *CS1* sheet. Instead of changing price by one dollar increments, the CS1 sheet was performed with a shock size of 0.1.

STEP Use your comparative statics results to make a *demand curve*, a graph of $x_1{}^* = f(p_1)$. To do this, select the p_1 data in column A, then hold down the *ctrl* key (and keep holding it), while selecting the x_1 data in column C. With cells in columns A and C selected, select the Scatter chart type. Title the graph and label the axes.

Another way to display the comparative statics results is via the *price consumption (or offer) curve*, as shown in Panel A of Figure 4.11 for a utility function that is not Cobb-Douglas and not meant to display the increasing price analysis that you just completed. Instead, a price decrease is shown.

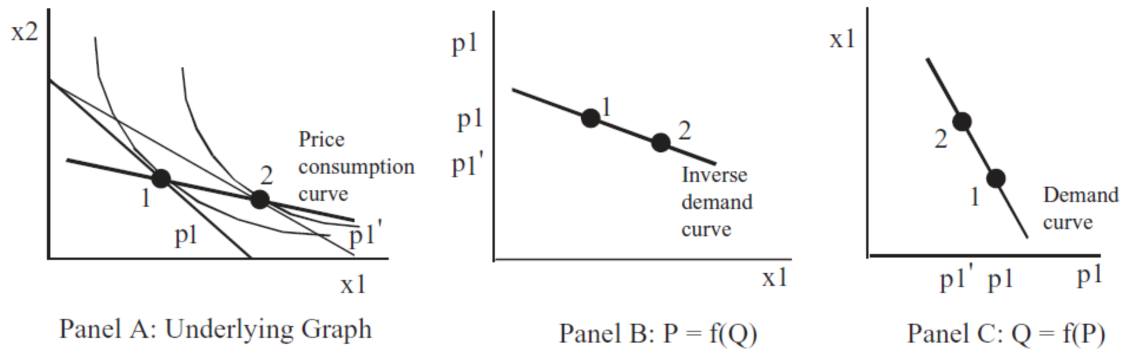


Figure 4.11: Three ways to show effects of p_1 shock.

There is a lot going on in Figure 4.11. The graph on the left (Panel A) shows a price decrease swinging the budget constraint out. It uses numbers to indicate the initial and new optimal solutions.

Panels B and C show demand, but look closely, the axes have been flipped. Instead of graphing x_1 as a function of p_1, the exogenous variable (p_1) is on the y axis in Panel B. This is a backwards, but common presentation in economics. The roots of this strange way of presenting the results can be traced back in the history of economics to Alfred Marshall in 1890.

Modern economists call the graph in Panel B of Figure 4.11 an *inverse demand curve* because it is plotted as $P = f(Q)$. The demand curve, the mathematically correct version, is $Q = f(P)$ because we plot $y = f(x)$ with y as the dependent variable that is determined by x.

In introductory economics, the inverse demand curve is used. The professor just draws a downward sloping line or curve and pronounces that it is obvious that as price goes up, quantity demanded falls (we will soon see that this is not guaranteed). As the level of sophistication rises, especially if we are doing econometrics and trying to estimate a demand curve, economists use the mathematically correct demand curve. Economists are used to both ways of presenting demand. It is confusing at first, but you can get the hang of it pretty quickly.

STEP Read the information in the *CS1* sheet. It explains how the ROUND function was used to create the price consumption curve from the comparative statics results.

Notice that the price consumption curve for changes in p_1 in the Excel workbook is horizontal. This is a property of the Cobb-Douglas utility function and is not especially realistic. The indifference map in Figure 4.11 is not based on a Cobb-Douglas utility function because the price consumption curve is not horizontal.

Another useful Excel skill to master that is especially relevant right now involves controlling the x and y axes. Excel's default is that the leftmost column of selected data goes on the x axis. If we want to make a demand curve with the data in the *CS1* sheet, this is convenient. We select the data in column A (p_1), hold down the *ctrl* key and select the data in column C (x_1). When you make a Scatter chart, Excel puts price on the x axis and quantity on the y axis.

But what if we want to make an inverse demand curve, with p_1 on the y axis? One easy way to do it is by directly editing the SERIES formula in the chart.

STEP Visit vimeo.com/econexcel/using-series-formula to watch a quick, 5-minute video of how the SERIES formula works.

After you watch the video, try it on your demand curve chart. Can you flip the axes by directly editing the SERIES formula? Click on your demand curve, then switch columns A and C in the x and y arguments in the SERIES formula. To see an example of this, click on the series in the chart in the *CS1* sheet.

Analytical Comparative Statics Analysis of Changing Price

We take the opportunity here to extend our previous analytical work. We could just leave p_1 as a letter since we want to derive a demand curve, but we will be more aggressive and leave all exogenous variables as letters. This will give us the most general answer we can get.

We rewrite the constraint and form the Lagrangean.

$$\max_{x_1, x_2, \lambda} L = x_1^c x_2^d + \lambda(m - p_1 x_1 - p_2 x_2)$$

Although it seems more formidable than when numbers are used in place of letters, we can apply the usual strategies for taking derivatives and solving the first-order conditions to find the optimal solution.

We take derivatives and set them equal to zero.

$$\frac{\partial L}{\partial x_1} = cx_1^{c-1}x_2^d - p_1\lambda = 0$$
$$\frac{\partial L}{\partial x_2} = dx_1^c x_2^{d-1} - p_2\lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = m - p_1x_1 - p_2x_2 = 0$$

To solve for the optimal values of x_1 and x_2, we move the lambda terms to the right-hand side and divide the first equation by the second. This gets rid of lambda and gives the familiar MRS $= \frac{p_1}{p_2}$ condition, which can then be solved for optimal x_2 as a function of optimal x_1.

$$\frac{cx_2^*}{dx_1^*} = \frac{p_1}{p_2}$$
$$x_2^* = \frac{d}{c}\frac{p_1}{p_2}x_1^*$$

We substitute this expression into the third first-order condition (the budget constraint) and solve for optimal x_1.

$$m - p_1x_1^* - p_2\left[\frac{d}{c}\frac{p_1}{p_2}x_1^*\right] = 0$$
$$\left(1 + \frac{d}{c}\right)p_1x_1^* = m$$
$$x_1^* = \left(\frac{c}{c+d}\right)\frac{m}{p_1}$$

This expression contains the demand curve for x_1 because it shows the quantity demanded at a given p_1. It also contains an Engel curve because it shows how x_1 varies with income. It also shows how x_1 moves when c or d, the consumer's tastes and preferences, change—although, such a graph is unnamed.

Furthermore, this expression can be evaluated for any combination of exogenous variable values. For example, suppose $c = d = 1, p_1 = 2$, and $m = 100$. Then it can be seen easily that optimal $x_1 = 25$. In fact, you can readily see that the reduced form expression for optimal x_1 agrees with the numerical approach using the Comparative Statics Wizard to recalculate the optimal solution at given values of p_1.

We can use our reduced form expression to calculate an own units response to a shock in p_1 by taking the derivative with respect to p_1.

$$x_1^* = \left(\frac{c}{c+d}\right)\frac{m}{p_1}$$

$$x_1^* = \left(\frac{c}{c+d}\right)m\left(p_1\right)^{-1}$$

$$\frac{dx_1^*}{dp_1} = -1\left(\frac{c}{c+d}\right)m\left(p_1\right)^{-2}$$

This formidable-looking expression is the instantaneous rate of change of the demand curve at a particular point. Because $x_1{}^*$ is a nonlinear function of p_1, its derivative with respect to p_1 contains p_1. The fact that the demand curve is not a line explains why we get different results when we compute responsiveness with Δ versus d.

STEP Read the *CS1* sheet carefully. Your primary goal is to understand the relationship between Δ in cells F14 and G14 versus the derivative in cells I13 and J13.

The key idea is this: as Δ gets smaller, it approaches d. Thus, earlier, we computed the price elasticity of demand from $p_1 = 2$ to 3 and got -0.67. But the *CS1* sheet shows an elasticity of -0.95 (in G14) as we go from $p_1 = 2$ to 2.1 and when we use the derivative formula, which is based on an infinitesimally small change in p_1, we get an elasticity of -1.

Notice that, unlike the demand curve, $x_1{}^* = f(p_1)$, the Engel curve, $x_1{}^* = f(m)$ is a line for the Cobb-Douglas utility function. We say, "x one star is nonlinear in p one" and "x one star is linear in m." Because the Engel curve is a line, Δm and the derivative with respect to m give identical results. The size of the change in m does not matter if the relationship is linear.

The unit price elasticity is a property of a Cobb-Douglas utility function. We can use the reduced form expression for $x_1{}^*$ to show that we always get a -1 price elasticity.

$$\frac{dx_1^*}{dp_1} \cdot \frac{p_1}{x_1^*} = -1\left(\frac{c}{c+d}\right) m\,(p_1)^{-2}\,\frac{p_1}{\left(\frac{c}{c+d}\right)\frac{m}{p_1}}$$

$$\frac{dx_1^*}{dp_1} \cdot \frac{p_1}{x_1^*} = -1$$

So Cobb-Douglas produces three constant elasticities:

1. Unit income elasticity
2. Unit own price elasticity
3. Zero cross price elasticity

None of these are especially realistic. Cobb-Douglas is common because it is easy to work with, not because it produces sensible elasticities.

A Point Off the Demand Curve?

Unlike an introductory economics course where demand curves appear out of the blue as downward sloping lines or curves, understanding where demand curves come from and what they actually represent are major goals for us.

So far, we have a mechanical understanding of the derivation of demand. Yes, it is true that changing p_1, ceteris paribus, and tracking how $x_1{}^*$ changes is how a demand curve is derived. And, yes, it is true that at every price, quantity demanded is the solution to an optimization problem for that price. But let's try a thought experiment not included in introductory economics.

If we consider what it means to be at a point off the demand curve, such as point Z in Figure 4.12, it helps us understand that the demand curve is really like a ridgeline across the top of a mountain range.

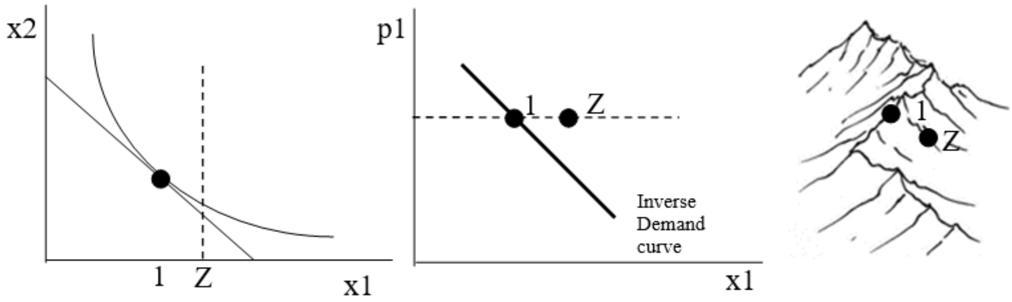


Figure 4.12: Interpreting a point off the (inverse) demand curve..

With a point Z to the right of the inverse demand curve, we know that the consumer is buying too much x_1, as shown by the vertical dashed line in the graph on the left of Figure 4.12. We cannot precisely plot the point Z on the indifference curve graph because we do not know how much good 2 the person is buying at point Z. We do know, however, that she is not optimizing. In other words, at point Z, this consumer is failing to maximize satisfaction and is not on the tangency of the budget line and highest attainable indifference curve.

Considering the meaning of a point off the demand curve reveals that a demand curve is a geometrical object with a special characteristic—every point on the demand curve is a point of maximum utility given prices and income. If we added an axis for utility, the demand curve would show itself as a 3D object that displayed the maximum utility at each given price. In other words, the demand curve is a ridgeline that connects mountain peaks, as shown in the sketch on the right in Figure 4.12.

A Demand Curve Is a Comparative Statics Exercise

Deriving a demand curve is the most important comparative statics exercise in the Theory of Consumer Behavior. Demand and supply (the most important comparative statics exercise in the Theory of the Firm) are at the heart of the market mechanism.

Given a particular functional form for utility, demand curves can be derived via numerical methods, picking off individual points on the demand curve for explicit values of price, ceteris paribus. Slopes and elasticities can be computed.

Demand curves can also be derived via analytical methods by finding the reduced form expression as a function of price (and any other exogenous variables). Slopes and elasticities can be computed by using the derivative.

For Cobb-Douglas utility, we found that $x_1{}^* = (\frac{c}{c+d})\frac{m}{p_1}$. For this reduced form, the numerical and analytical methods yield different values for slopes and elasticities based on changing p_1 because the demand curve is a curve, instead of a line (like the Engel curve). The smaller the discrete change in p_1 used in the numerical method, the closer it gets to the analytical result.

We can also "derive" a demand curve with graphs, as shown in Figure 4.11. We can display the effect of a price change by rotating the budget line and showing the initial and new points of tangency. If we display the p_1 and corresponding optimal amount of x_1 in a separate graph, we have graphically derived a demand curve (or inverse demand curve, if we flip the axes).

Finally, if we work out the implications of a point off the demand curve, we can see the demand curve in a new light—it is actually a 3D object represented in 2D space. All of the points on the demand curve are actually points of maximum utility subject to the budget constraint.

Exercises

1. In the *OptimalChoice* sheet, click the Reset button and reproduce Figure 4.10 with a decrease (instead of an increase) in p_1 from \$2/unit to \$1/unit. Use Word's Table feature to create the table and fill in the cells.

2. Use Word's Drawing Tools to create a graph of the price consumption curve and demand curve for x_1 (as in Figure 4.11) that accurately reflects the shock and results from question 1.

3. What is the difference between a demand curve and an inverse demand curve?

References

The epigraph is from page 103 of George J. Stigler, "The Early History of Empirical Studies of Consumer Behavior," *The Journal of Political Economy*, Vol. 62, No. 2 (April, 1954), pp. 95–113 (www.jstor.org/stable/1825569)

Most economists do not care who first came up with the concept of a demand schedule. Most of those who do care believe that it was Gregory King, a century after Charles Davenant. Stigler was a winner of the Nobel Prize in Economics and a professor at the University of Chicago. He had a lifelong passion for the intellectual history of economics. In this article, he showed that Davenant actually preceded King.

It took a long time to translate demand (and supply) schedules as tables (with columns for price and quantity) into graphs. Fleeming (pronounced flem-ming) Jenkin in 1870 is often given credit for drawing the first demand curve, but there were precursors. Alfred Marshall's *Principles of Economics* (1890) popularized supply and demand graphs. His graphs appeared, however, only in footnotes.

Marshall's *Principles* was the most popular economics book of its era. It is freely available online at www.econlib.org/library/Marshall/marP.html.

Modern economists sometimes mock Marshall for switching the axes, claiming he made a mistake, but this assertion is incorrect. Marshall put price on the vertical axis because he wanted to show market demand and supply curves on a graph as the horizontal sum of individual demand and supply curves, as in footnote 70 from Book III, Chapter IV. Future generations of introductory economics students became locked in to the Marshallian inverse demand and supply curves.

Although you may conclude that Marshall's violation of accepted mathematical convention (i.e., independent variables belong on the x axis) is confusing, the decision was not due to a lack of math knowledge. In fact, Marshall was a brilliant mathematician, earning Second Wrangler (to the future Lord Rayleigh) as an undergraduate at Cambridge in the Tripos competition.

To understand how the role of mathematics has changed in economics, consider the recipe Marshall gave a friend for using math in economics: "1) Use mathematics as a shorthand language, rather than as an engine of inquiry. 2) Keep to them till you have done. 3) Translate into English. 4) Then illustrate by examples that are important in real life. 5) Burn the mathematics. 6) If you can't succeed in 4 burn 3. This last I did often." (A. C. Pigou, Memorials of Alfred Marshall, 1925, p. 427.)

Quasilinear utility functions are not particularly realistic, but they are very easy to work with.

Hal Varian

4.4 More Practice with Deriving Demand

This section derives the demand curve from two different utility functions, quasilinear preferences and perfect complements, to provide practice deriving demand curves. Nothing new here, just practice applying the tools, techniques, and concepts of the economic way of thinking.

Quasilinear Preferences

We begin with the analytical approach. Rewrite the constraint and form the Lagrangean, leaving p_1 as a letter so we can derive a demand curve.

$$\max_{x_1, x_2, \lambda} L = x_1^{1/2} + x_2 + \lambda(140 - p_1 x_1 - 10x_2)$$

STEP Follow the usual Lagrangean procedure to solve this problem. For help, refer back to section 4.2 where we solved this same problem except with m instead of p_1.

You should find reduced form expressions like this:

$$x_1^* = \frac{25}{p_1^2}$$

$$x_2^* = 14 - \frac{2.5}{p_1}$$

The first expression, $x_1{}^* = \frac{25}{p_1^2}$, is a demand curve for $x_1{}^*$ because it gives the quantity demanded of x_1 as a function of p_1. If we rewrite the equation in terms of p_1 like this, $p_1^2 = \frac{25}{x_1{}^*} \rightarrow p_1 = \frac{5}{\sqrt{x_1{}^*}}$ then we have an inverse demand curve, with price on the y axis as a function of quantity on the x axis.

The derivative of x_1* with respect to p_1 tells us the slope of the demand curve at any given price.

$$x_1^* = 25p_1^{-2}$$
$$\frac{dx_1^*}{dp_1} = -2 \cdot 25p_1^{-3} = -\frac{50}{p_1^3}$$

The own price elasticity of demand is:

$$\frac{dx_1^*}{dp_1} \cdot \frac{p_1}{x_1^*} = -\frac{50}{p_1^3}\frac{p_1}{\frac{25}{p_1^2}} = -2$$

The constant elasticity of demand for good 1 is a property of the quasilinear utility function. Notice that 2 is the reciprocal of the exponent on x_1 in the utility function. In fact, with $U = x_1^c + x_2$, the price elasticity of demand for x_1 is $-\frac{1}{1-c}$ for values of c that yield interior solutions.

The expression for optimal x_2 is a cross price relationship. It tells us how the quantity demanded for good 2 varies as the price of good 1 changes. The equation can be used to compute a cross price elasticity, like this:

$$\frac{dx_2^*}{dp_1} \cdot \frac{p_1}{x_2^*} = \frac{2.5}{p_1^2}\frac{p_1}{14 - \frac{2.5}{p_1}} = \frac{2.5}{p_1\left(14 - \frac{2.5}{p_1}\right)} = \frac{2.5}{p_1\left(\frac{14p_1 - 2.5}{p_1}\right)} = \frac{2.5}{14p_1 - 2.5}$$

Unlike the own price elasticity, the cross price elasticity is not constant—it depends on the value of p_1. It is also positive (whereas the own price elasticity was negative). When p_1 rises, optimal x_2 also rises. This means that goods 1 and 2 are *substitutes*.

Complements, on the other hand, are goods whose cross price elasticity is negative. This means that an increase in the price of good 1 leads to a decrease in consumption of good 2.

Demand can also be derived via numerical methods.

STEP Open the Excel workbook *DemandCurvesPractice.xls*, read the *Intro* sheet, then go to the *QuasilinearChoice* sheet.

The consumer is maximizing satisfaction at the initial parameter values because the marginal condition, MRS $= \frac{p_1}{p_2}$, is met at the point 6.25,12.75 (ignoring Solver's false precision) and income is exhausted.

We can explore how this initial optimal solution varies as the price of good 1 changes via numerical methods. We simply change p_1 repeatedly, running Solver at each price, while keeping track of the optimal solution at each price. The Comparative Statics Wizard add-in handles the tedious, cumbersome calculations and outputs the results in a new sheet for us.

STEP Run the Comparative Statics Wizard on the *QuasilinearChoice* sheet. Increase the price of good 1 by 0.1 (10 cent) increments.

You can check your comparative statics analysis by comparing your results to the *CS1* sheet, which is based on 1 (instead of 0.1) dollar size shocks. Of course, the numbers will not be exactly the same since the Δp_1 shock size is different.

The columns of price and optimal x_1 are points on the demand schedule. The numerical approach via the CSWiz essentially picks individual points on the demand curve for the given prices. If you plot these points, you have a graph of the demand curve.

The analytical approach, on the other hand, gives the demand function as an equation. You can evaluate the expression at particular prices and generate a plot of the demand curve.

The two approaches, if done correctly, will always yield the same graphical depiction of the demand curve. They may not, however, yield the same slopes or elasticities.

STEP Using your results, create demand and price consumption curves. Compute the own unit changes and elasticities for x_1* and x_2*.

The *CS1* sheet shows how to do this if you get stuck. You can click on cells to see their formulas. Think about how the formulas work and how they compute the answer.

It is critical that you notice that your own unit changes and elasticities are closer to the instantaneous rates of change in columns I and J of the *CS1* sheet because you have smaller changes in p_1 and, for this utility function, $x_1{}^*$ is nonlinear in p_1.

Take a moment to reflect on what is going in the calculations presented in the *CS1* sheet. The color-shaded cells invite you to compare those cells.

Now, let's walk through this slowly.

STEP Click on cell F13 to see its formula.

It is computed as the change in optimal x_1 for a \$1 increase in p_1. There is a decrease of about 3.47 units when price increases by 1 unit.

STEP Click on cell I12 to see its formula.

It is computed by substituting the initial price, \$2/unit, into the expression for the derivative (displayed as an equation above the cell). The result of the formula, -6.25, is the instantaneous rate of change. In other words, there will be a 6.25-fold decrease in optimal x_1 given an infinitesimally small increase in p_1.

STEP Go to your CSWiz results and, if you have not done so already, compute the change in optimal x_1 for a \$0.1 increase in p_1.

You should find that your slope is about -5.8. The change in optimal x_1 is about 0.58, but you have to divide by the change in price, 0.1, to get the slope. Notice that your answer is much closer to the derivative-based rate of change (-6.25). This is because you took a much smaller change in price, 0.1, than the one dollar change in price in the *CS1* sheet and you are working with a curve.

STEP Return to the *CS1* sheet and compare cells G13 and J12.

The same principle is at work here. Because the demand curve is nonlinear, the two cells do not agree. Cell G13 is computing the elasticity from one point to another, whereas cell J12 is using the instantaneous rate of change (slope of the tangent line) at a point.

If you compute the price elasticity from 2 to 2.1 (using your CS results), you will find that it is much closer to -2.

Finally, you might notice that unlike the Cobb-Douglas utility function, which produced a horizontal price consumption curve (PCC), the quasilinear utility function in this case is generating a downward sloping price consumption curve. In fact, the slope of the price consumption curve tells you the price elasticity of demand: Upward sloping PCC means that demand is inelastic, horizontal PCC yields a unit elastic demand (as in the Cobb-Douglas case), and downward sloping PCC gives elastic demand (as in this case).

Perfect Complements

We begin with the analytical approach.

$$U(x_1, x_2) = min\{ax_1, bx_2\}$$

For $a = b = 1$, we know that we can find the intersection of the optimal choice and budget lines to get the reduced form expressions for the endogenous variables, $x_1{}^* = \frac{m}{p_1+p_2}$ (which is the same for $x_2{}^*$ since $x_1{}^* = x_2{}^*$).

This solution says that when a and b are the same in a perfect complements utility function, the optimal amounts of each good are equal and found by simply dividing income by the sum of the prices.

The reduced form expression contains Engel and demand curves. Holding prices constant, we can see how m affects consumption. Likewise, holding m and p_2 constant, we can explore how optimal x_1 varies as p_1 changes. This, of course, is a demand curve for x_1.

As usual, we find the instantaneous rate of change by taking the derivative with respect to p_1. The p_1 elasticity of x_1 is the derivative multiplied by $\frac{p_1}{x_1{}^*}$.

$$\frac{dx_1^{x^*}}{dp_1} = -\frac{m}{(p_1+p_2)^2}$$

$$\frac{dx_1^{x^*}}{dp_1} \cdot \frac{p_1}{x_1^{x^*}} = -\frac{m}{(p_1+p_2)^2}\frac{p_1}{\frac{m}{p_1+p_2}} = -\frac{p_1}{p_1+p_2}$$

We can also derive demand for a perfect complements utility function via numerical methods.

STEP Proceed to the *PerfCompChoice* sheet and run the Comparative Statics Wizard with an increase in the price of good 1 of 0.1 (10 cents).

Can you guess what we will do next? The procedure is the same every time: we solve the model then explore how the optimal solution responds to shocks.

STEP Create demand and price consumption curves based on your comparative statics results. Compute the own units changes and elasticities for x_1* and x_2*. The *CS2* sheet shows how to do this if you get stuck.

As before, you will want to concentrate on how your own units changes and elasticities are closer to the instantaneous rates of change than the Δp_1 in columns F and G of the *CS2* sheet because you have smaller changes in p_1 and we are dealing with a nonlinear relationship.

The lesson is clear: whenever the demand curve is not a line, that is, x_1* is nonlinear in p_1, then Δp_1 will not exactly equal dp_1. As the size of the discrete change in price gets smaller, the numerical method result will approach the result based on the derivative.

Although the two methods might not exactly agree, they are usually pretty close. How close depends on the curvature of the relationship and the size of the discrete shock. This means you can always check your analytical work by doing a manual Δ shock and computing the change from one point to another.

Notice also that the price consumption curve is upward sloping and the price elasticity is less than one (in absolute value).

Deriving Demand from the Consumer's Utility Maximization Problem

The primary purpose of this section was to provide additional practice in deriving demand with different utility functions. Clearly, the demand curve is strongly influenced by the utility function that is being maximized given a budget constraint.

Two examples were used to demonstrate how the analytical and numerical methods are related. Calculus is based on the idea of infinitesimally small changes. You can see calculus in action by using the CSWiz to take smaller changes in price—which drives the numerical method ever closer to the derivative-based result.

Exercises

1. Return to the *QuasilinearChoice* sheet and click the Reset button. Now change the exponent on good 1 from 0.5 to 0.75. Use the Comparative Statics Wizard to derive a demand curve for this utility function.

2. Working with the same utility function as in the first question, derive the demand for $x_1{}^*$ via analytical methods. Use Word's Equation Editor as needed. Show your work.

3. Using your results from questions 1 and 2, compute the own price elasticity via numerical and analytical methods. Do they agree? Why or why not? Show your work and take screen shots as needed.

References

The epigraph is from page 63 of Hal Varian's best-selling, undergraduate textbook, *Intermediate Microeconomics* (7th edition, 2006). In the preface, Varian tackles head on the issue of calculus. "Many undergraduate majors in economics are students who should know calculus, but don't—at least not very well. For this reason, I have kept calculus out of the main body of the text."

The book you are reading at this moment takes a different approach. Calculus is used extensively, but it is made accessible by consistent repetition along with the substantial support of numerical methods. If you are a student who struggles with analytical methods, you will never have a better opportunity to master calculus and algebra. Do the practice problems with care and match the analytical and numerical approaches in each application.

> To my knowledge, no one has described heroin as a Giffen good. But the description may be appropriate for those users who are addicted.
>
> Neal Kumar Katyal

4.5 Giffen Goods

Demand curves are derived by doing comparative statics on the consumer's optimization problem: Change price, ceteris paribus, and track optimal consumption of a good.

In introductory economics courses around the world, demand is always drawn downward sloping so that as price rises, ceteris paribus, quantity demanded falls. Economists have long been intrigued, however, by a perplexing possibility: quantity demanded rising as price rises. An upward sloping demand curve! Can this happen? Yes, but it is quite rare and it took decades to figure it out.

We begin with a definition: *Giffen goods* are goods that have upward sloping demand curves. Giffen's connection to this counter intuitive demand relationship—price rises and you want to buy more?—is controversial.

Giffen and the Irish Potato Famine

The Great Irish Famine took place during 1845-1848.

> To put the disaster in proper perspective, the famine killed at least 12 percent of the population over a three-year period. Another 6-8 percent migrated to other countries. In terms of the percentage of population affected, the 1845-48 famine is one of the largest ever recorded. Other famines have killed more people in total because the affected populations were larger, not the percentage of exposure. For instance, the 30 million or more people who perished in the Chinese famine of 1958-62 were 5 percent or 6 percent of the population. (Rosen, 1999, p. S303)

Why did so many people die? This is a difficult question to answer comprehensively. The economics of famine are complicated. The proximate answer is that the Irish ate a *lot* of potatoes and a potato blight destroyed the food source. Rosen (1999, p. S303) says this:

> As difficult as it is to imagine today, on the eve of the famine, per capita consumption of potatoes is reliably estimated to have averaged 9 pounds (40-50 potatoes) per person per day (Bourke 1993). Diets were astonishingly concentrated on potatoes, especially in rural areas. Grain was grown in rural Ireland but was either sent to towns or exported abroad.

When blight wiped out the potato crop, why didn't the Irish eat something else or just import food? This is hard to understand. Books have been written on the subject. The *Biblio* sheet in *GiffenGoods.xls* has references. In fact, Amartya Sen won a Nobel Prize in Economics for his work on famine. It turns out that it is not simply a matter of too little food—amazingly, food can be just a few miles away and yet many people can be starving!

But our focus is on Giffen goods and the story picks up decades after the famine. Although there is no evidence that he ever said anything close to "price increase led to higher quantity demanded," Sir Robert Giffen (1837–1910) is credited with using the behavior of potato prices and quantities to state the claim that quantity demanded rose as prices rose.

Figure 4.13 shows Irish potato prices before, during and after the famine. Although consumption fell when price spiked in 1847 to more than double the 1846 price, somehow the legend grew that quantity demanded increased as prices rose in this time period. Thus, the Irish potato became the canonical example of a Giffen good—even though there is no evidence that price and quantity moved in the same direction.

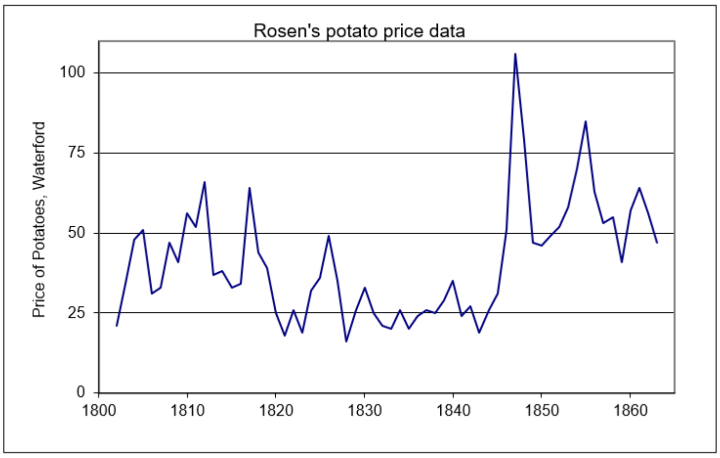


Figure 4.13: Potato price in Waterford, Ireland.
Source: OptimalChoice.xls!OptimalChoice

Economists began arguing over whether or not quantity demanded rose as the price spiked and, even if it did not, whether it was theoretically possible. It would take decades of contentious debate before the matter was settled.

Two Common Mistakes in the Giffen Debate

Before explaining how we could, in theory, get a Giffen good, we need to clear up two mistakes in thinking about Giffen goods. Both mistakes involve violating the strict ceteris paribus requirement that underlies a demand curve. The first mistake has a long history in econometrics and the second is easily corrected once we remember that we must hold everything else constant.

Estimating demand from observed prices and quantities is quite difficult. It turns out that plotting price and quantity data over time and fitting a line is no way to estimate a demand curve.

Suppose that the observed quantity of potatoes sold and consumed really had increased as the price spiked in 1847. Would that have been a good way to support the Giffen good claim? Absolutely not.

The problem is that the price and quantity data in different time periods do not fulfill the ceteris paribus requirement. It is true that price and quantity changed over time, but presumably so did other factors that affect demand and supply.

STEP Open the Excel workbook *GiffenGoods.xls* read the *Intro* sheet, then go to the *ID* sheet and read it carefully. Make sure to click the buttons and think about the charts that are displayed.

This sheet walks you through a simple example and shows why fitting a line to observed market price and quantity data is a really bad move. The heart of the confusion lies in the inability to extract the individual supply and demand curves that produce the observed data. This is called the *identification problem.*

So, even if it is true that we see prices and quantities moving together, that is not a demonstration of Giffen behavior.

The second mistake is less easy to forgive. No complicated issues of estimation are involved. We simply forget that demand requires that the ceteris paribus condition hold. Suppose you notice that a particular brand of jeans has become increasingly popular and suddenly more people want it as its price rises. Have we discovered a Giffen good?

Absolutely not. We are violating the crucial ceteris paribus part of the definition of a demand curve by failing to hold constant everything except a change in price. In this case, the increased popularity of a particular brand is a shock to the demand curve, shifting it right. This is not a Giffen good because we are not working with a single, fixed demand curve. Instead, as in the second chart in the *ID* sheet, changes in demand are driving new equilibrium price–quantity combinations.

Having seen two common mistakes in trying to understand and show Giffen behavior, both involving violation of the strict ceteris paribus condition, the natural question then is: Can true Giffen goods, ones that meet the specific requirements of a demand function, exist? The answer is yes.

Giffen Goods in Theory

The left graph in Figure 4.14 shows the canonical graph of the Theory of Consumer Behavior displaying a Giffen good, while the right shows its associated upward sloping demand curve. Notice that the indifference curves require a little tweaking and somewhat odd placement to make x_1 be a Giffen good. Remember that indifference curves cannot cross, but they do not have to be similarly shaped and equally separated. For x_1 to be Giffen, point 2 in Figure 4.14 has to lie to the left of point 1 so that the decrease in p_1 leads to a decrease in optimal x_1.

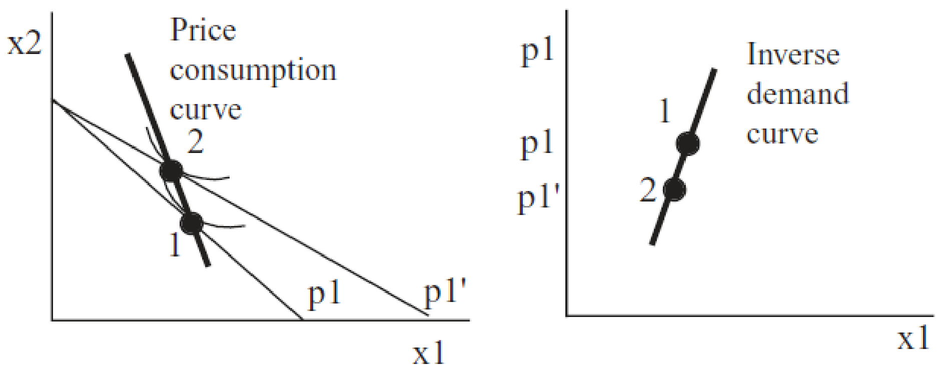


Figure 4.14: A Giffen good.

Do not be confused by the decrease in x_1. Quantity demanded fell, but so did price. Thus, we have a *positive relationship* between price and quantity demanded (they are moving together) and an upward sloping demand curve. This is a Giffen good.

To be crystal clear, it is not the fact that optimal x_1 decreased that tells us we have a Giffen good, but that it decreased as price fell. If we started at point 2 and raised the price, the budget constraint would swing in, and we would move to point 1, with an increase in optimal x_1. We would have *Giffenness* because x_1 rose as p_1 increased, We would be traveling up the upward sloping demand curve.

A version of Figure 4.14 is depicted in every microeconomics book that discusses Giffen goods and, make no mistake, this is a canonical graph in micro theory. But dead graphs on a printed page (or computer screen) force the reader to reconstruct individual elements and can be difficult to disentangle. With Excel at our disposal, we can walk through a numerical example to gain complete mastery of the concept of Giffenness.

STEP Proceed to the *Optimal1* sheet and look at the utility function.

The sheet models a Giffen good. The utility function is admittedly quite complicated, but a simple functional form like Cobb-Douglas or quasilinear is never going to produce Giffenness.

$$u(x_1, x_2) = \left\{ \begin{array}{ll} ax_1 - \frac{b}{2}x_1^2 + cx_2 + \frac{d}{2}x_2^2 & \text{for } 0 \leq x_1 \leq a/b \\ \frac{a^2}{2b} + cx_2 + \frac{d}{2}x_2^2 & \text{for } x_1 > a/b \end{array} \right\}$$

The *U1* sheet shows that this functional form meets the requirements of well-behaved preferences. The coefficients have been set to values that do not violate the axioms of revealed preference in the range we are working in. The indifference curves, for example, will never intersect.

Another example of a utility function that exhibits Giffen behavior is $U = ax_1 + \ln x_1 + \frac{x_2^2}{2}$. This is implemented in the *Optimal2* sheet. We will use the *Optimal1* sheet here and save the *Optimal2* sheet for Q&A work. These are just two of the many functional forms that meet the requirements of well-behaved utility that could exhibit Giffen behavior.

The *Optimal1* sheet opens with $x_1 = 44$ and $x_2 = 11$. A single indifference curve is displayed and it does not have the curvature we have been used to seeing. Recall that perfect substitutes are straight lines, so we can infer that this utility function is expressing preferences with a high degree of substitutability between the two goods.

Without running Solver, we know this is the optimal solution because the MRS equals the price ratio.

STEP It is hard to see that the budget line is just touching the indifference curve, but if you click the Zoom In button, you will see that the tangency condition is clearly met.

Since we are working on Giffen behavior, we want to explore the effects of a change in price on the quantity demanded. We will increase the price of x_1 and see how the consumer responds. Before we do, think through what will happen. How will the constraint change and where must the new tangency point lie if x_1 is a Giffen good?

STEP Change p_1 to 1.1. What happens?

The budget line pivots around the y intercept. It may look like a parallel shift, but it really is not.

STEP Click the Zoom Out button to see that the price increase has, as expected, rotated the budget line in.

The 44,11 initial optimal bundle is no longer affordable. The consumer must re-optimize.

STEP Run Solver. What happens?

Figure 4.15 shows the result. Optimal consumption of good 2 has collapsed from 11 to around 1.5 and the consumer now wants to buy 48.6 units of good 1, which is more than the initial amount of 44.

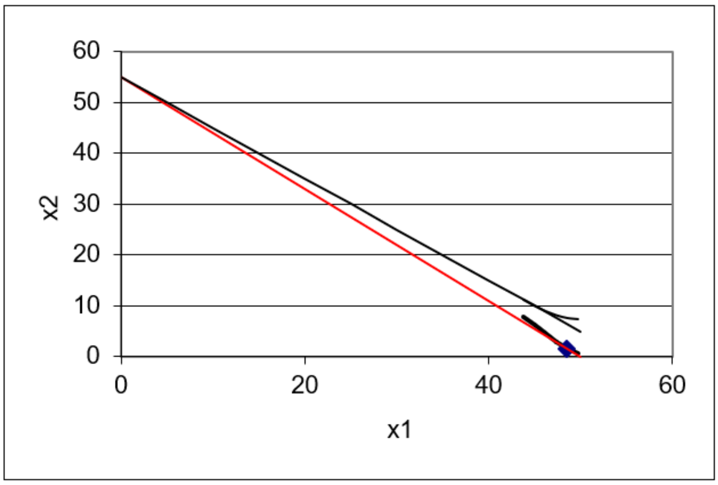


Figure 4.15: A numerical example of Giffen behavior.
Source: GiffenGoods.xls!Optimal1

This is amazing! The price of good 1 went *up* by 10 cents (from 1 to 1.1) and the optimal amount of good 1 *increased* by 4.6 units (from 44 to 48.6). Price rose, ceteris paribus, and so did quantity demanded!

This is a concrete, numerical example of a Giffen good. We can use the Comparative Statics Wizard to explore more carefully the demand curve resulting from this bizarre utility function.

STEP Use the Comparative Statics Wizard to trace the demand curve from 0.1 to 3. Set cell B16 to 0.1, then apply 300 (yes, 300) shocks by increments of 0.01 with the CSWiz add-in. Finally, create a graph of the inverse demand curve, p_1 as a function of x_1*.

Your results should look like Figure 4.16, which is also in the *CS1* sheet. That is certainly a strange looking demand curve. It is Giffen in a range. In other words, a Giffen good is not intrinsically and everywhere a Giffen good. Giffenness is a local phenomenon. The demand curve pictured in Figure 4.16 has three different behaviors. As price rises from zero, quantity demanded falls. This continues until a price of about 70 cents. From there, penny increases lead to increased consumption of good 1. In this range, x_1 is a Giffen good. There is a third region, at prices such as \$2 and \$3, where the good is not Giffen.

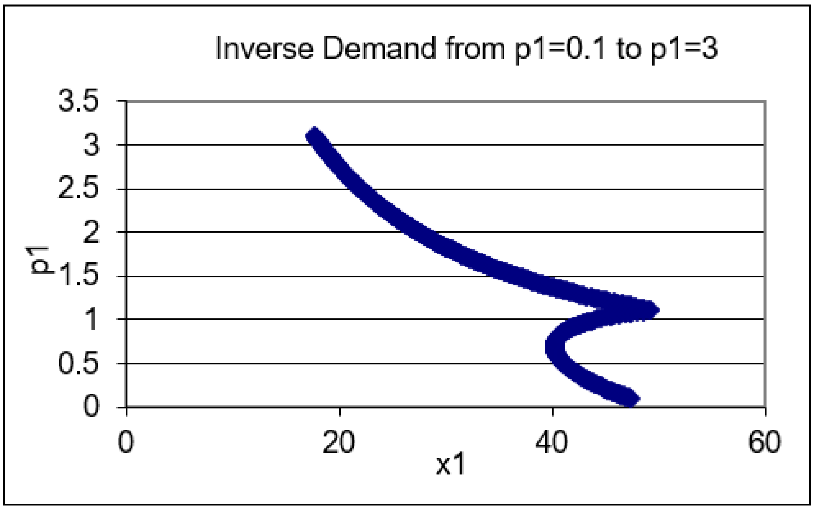


Figure 4.16: The inverse demand curve for x_1.
Source: GiffenGoods.xls!CS1

So, this example has shown that Giffen goods are not only possible, they can be modeled by the Theory of Consumer Behavior. We now know that there are utility functions that reflect well-behaved preferences that generate Giffen behavior.

Giffen Goods in Theory and Practice

A Giffen good is a strange creature in economics. The phenomenon of quantity demanded rising as price increases was first purportedly sighted during the Irish potato famine and named after Sir Robert Giffen, even though there is no evidence that Giffen actually claimed seeing quantity demanded rise as prices rose, ceteris paribus.

Certainly there are utility functions that give rise to Giffen goods. Certainly individual consumers may have well-behaved preferences that yield Giffen behavior. But has a Giffen good ever been spotted? Do Giffen goods exist in the real world in the sense that a market demand curve is upward sloping? This is the subject of much debate. Ceteris paribus is a difficult requirement to meet.

The actual sighting of a Giffen good in the real world remains contentious. We know for sure that the original example, potatoes during the Great Irish

Famine, was flawed and there is little evidence that it was a Giffen good. The *Biblio* sheet has a few references that can start you learning more about the history of Giffen goods in economics.

The next section gives an even deeper explanation for Giffen goods. It establishes the specific conditions needed for Giffenness to occur.

Exercises

1. Use the results in the *CS1* sheet to find the price range for which we see Giffen behavior. Report your answer and describe your procedure.

2. Use the *Optimal1* sheet utility function and parameter values to find the optimal solution via analytical methods. Show your work. Note that $x_1 < \frac{a}{b}$,so the utility function is

$$U = ax_1 - \frac{b}{2}x_1^2 + cx_2 + \frac{d}{2}x_2^2$$

3. Use Word's Drawing Tools to reproduce Figure 4.14, depicting x_1 as a Giffen good, but use a p_1 increase (instead of a decrease).

References

The epigraph comes from page 2436 of Neal Kumar Katyal, "Deterrence's Difficulty," *Michigan Law Review*, Vol. 95, No. 8. (August, 1997), pp. 2385–2476, repository.law.umich.edu/mlr/vol95/iss8/3/.

The *Biblio* sheet in *GiffenGoods.xls* has a list of references on Giffen goods. Scroll down to see suggested readings on the Irish potato famine, the history of Giffen goods in economics, and modern-day efforts at finding Giffen goods. Click on a link if anything catches your eye and seems worth exploring.

Eugene (or Eugen or Yevgeni) Slutsky [1880 – 1948] intended to become a mathematician, but he was expelled from the University of Kiev for participating in student revolts.

Gonçalo L. Fonseca

4.6 Income and Substitution Effects

Without a doubt, the demand curve is the most important idea in the Theory of Consumer Behavior. We have derived the demand curve analytically and numerically. The demand curve tells us the optimal amount to buy at a given price. It also tells us how quantity demanded will change as price changes, ceteris paribus.

This section remains focused on the demand curve, extending the analysis of the consumer's optimal response to a change in price. The core concept is that the total effect on quantity demanded (given by the demand curve) for a given change in price can be broken down into two separate effects, called *income and substitution effects*.

Our attention is still on the change in quantity demanded as price changes, ceteris paribus, but by breaking apart the observed response when price changes, we get a deeper explanation of demand. We also explain how we might get a Giffen good.

Intuition

Before diving into complicated graphs and math, let's review the story behind income and substitution effects. Seeing the big picture improves your chances of really understanding what income and substitution effects are all about.

Suppose that, ceteris paribus, price rises. We know the consumer has to re-optimize. We know the consumer will choose a new optimal combination of goods. We can see the consumer buy a different amount after the price changes. If we simply compute the change in the amount purchased of x_1 before and after the price change, we are comparing two points on the demand curve. This is called the *total effect* of a price change.

The breakthrough idea is that the increase in price has two channels by which it affects the consumer. One channel focuses on the fact that a price increase is like a decrease in purchasing power. After all, given an income level, if prices double, then I can buy half of what I bought before. My income has not changed, but my purchasing power has fallen just the same as if my income had been cut in half. The *income effect* reflects the fact that price changes affect optimal quantity demanded by altering purchasing power.

The other channel is called the *substitution effect*. The idea is that a price change in one good alters the relative prices faced by the consumer and induces substitution of the relatively cheaper good for the relatively more expensive one. When p_1 rises, x_1 is relatively more expensive than x_2 and so I am naturally going to avoid x_1 and be attracted to x_2.

Figure 4.17 shows the two channels below the total effect—they are submerged and not directly observed. Added together, they make up the total effect.

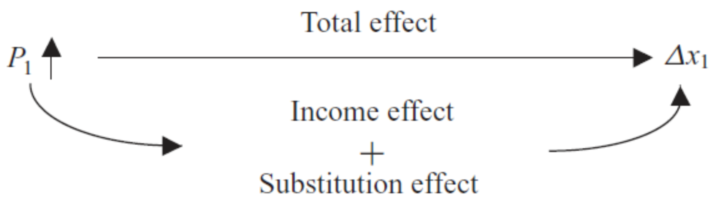


Figure 4.17: The basic idea behind income and substitution effects.

We will see that the income effect can be either positive or negative, but the substitution effect is *always negative* (assuming well-behaved preferences). When price goes up, the substitution effect says "buy less." Of course, if price falls, the reverse occurs and, according to the substitution effect alone, consumption increases.

The reason the income effect is ambiguous in sign is the fact that there are normal and inferior goods. If the good is normal, then optimal x_1 rises as income increases, but if the good is inferior, then consumption and income are inversely related.

Finally, it helps to know the underlying motivation behind the discovery of income and substitution effects. Economists were arguing about the existence of Giffen goods. The Law of Demand said price and quantity were inversely related. Income and substitution effects explained under which conditions Giffen behavior (an upward sloping demand curve) is possible. We will see that if the income and substitution effects work together, then the demand curve is guaranteed to be downward sloping. Understanding income and substitution effects will allow us to give a more refined, precise definition of the Law of Demand.

Numerical Example of Income and Substitution Effects

STEP Open the Excel workbook *IncSubEffects.xls*, read the *Intro* sheet, and proceed to the *OptimalChoice* sheet.

We have the usual Cobb-Douglas utility function with a conventional budget line. We have done this problem before and the initial optimal solution is $25{,}16\frac{2}{3}$.

STEP Decrease p_1 by 1 to \$1/unit (in cell B17).

Figure 4.18 displays what is on your screen. The red line is the familiar new budget line (after the price decrease). There is, however, a dashed line that has not been used before. This dashed line represents the outcome of a thought experiment.

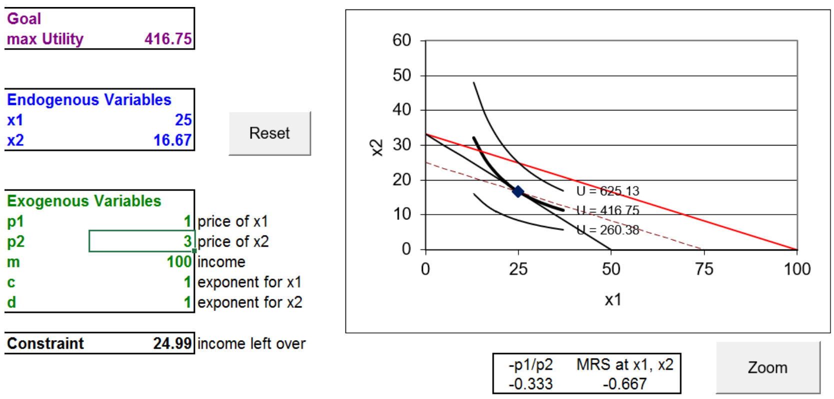


Figure 4.18: Decreasing p_1.
Source: IncSubEffects.xls!OptimalChoice

STEP Click the Zoom button to see a second graph of the situation. It has the axes scale adjusted so you can see better what is going on.

The dashed line is critical to understanding the splitting of the total effect into income and substitution effects. It has the same slope as the new budget line, yet it goes through the initial optimal solution. What we have done is pretend to take away enough income from the consumer to enable him to buy the initial bundle with the new, lower p_1.

We took away income (shifting down the budget constraint relative to the new budget line) because the fall in price implies an increase in purchasing power. Had there been a price rise, we would have had to increase income to compensate for the price increase.

We will find a tangency solution on the dashed line and this will allow us to split the total effect into the income and substitution effects.

Of course, nothing like this actually happens in the real world. When the price falls, the consumer re-optimizes, buying a new optimal bundle, and that is the end of the story. But for the purposes of understanding the demand curve, we figure out what the consumer would buy at the imaginary dashed line and we use that to split the total effect into the substitution and income effects.

But this is all way too abstract. Let's actually do it so you can see how it works. To figure out how much income to take away to cancel out the changed purchasing power from the price change, we use the *Income Adjuster Equation.*

$$\Delta m = x_1{}^* \Delta p_1$$

Applied to this problem, we know that $x_1{}^*$ is 25 (from the initial optimal solution) and the change in p_1 is -1 (because the price fell from 2 to 1, so $new - initial$ is $1 - 2$); thus, we have:

$$\Delta m = x_1{}^* \Delta p_1$$

$$\Delta m = [25][-1] = -25$$

The minus tells us that we have to take away income. The dashed line is based on an income of \$75, $p_1 = 1$, and $p_2 = 3$.

In summary, we have three budget lines when we work with income and substitution effects: (1) the usual initial line, (2) the usual new line from the change in price, and (3) the imaginary (dashed) line that has been adjusted to pass through the initial optimal solution.

We find the usual new optimal solution so we can compute the total effect first, then we use the dashed line to find the income and substitution effects.

STEP With $p_1 = 1$, run Solver.

Figure 4.19 shows that the consumer chooses the 50,$16\frac{2}{3}$ combination. Thus, we have two points to consider so far:

- Point A: Initial: At $m = 100, p_1 = 2, x_1{}^* = 25, x_1{}^* = 16\frac{2}{3}$.
- Point C: New: At $m = 100, p_1 = 1, x_1{}^* = 50, x_1{}^* = 16\frac{2}{3}$.

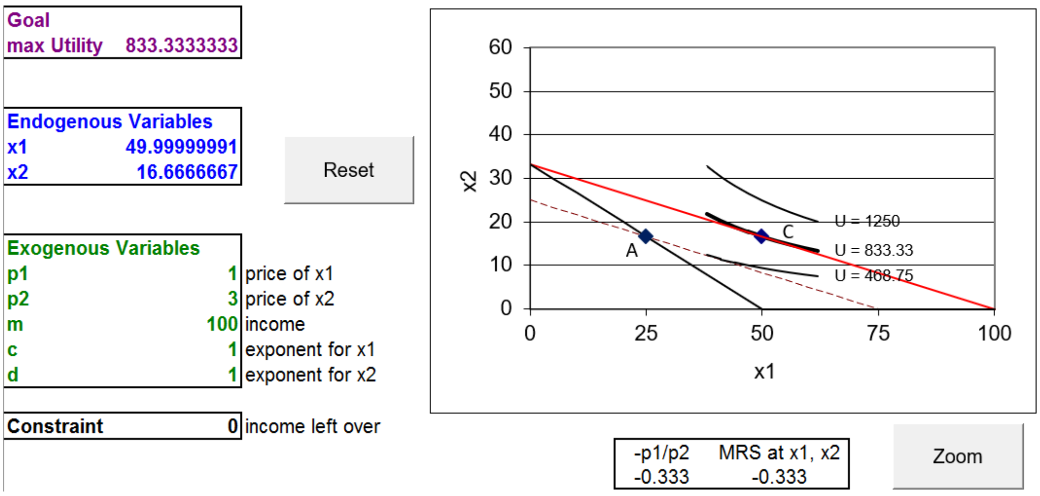


Figure 4.19: New optimal solution at $p_1 = 1$.
Source: IncSubEffects.xls!OptimalChoice

Notice that Excel displays three difference curves around the current optimal solution, but there are actually an infinite number of curves going through every point in the quadrant. With $c = d = 1$ being held constant, the indifference map is not changing in any way. We are simply displaying different indifference curves whenever x_1 and x_2 in cells B12 and B13 change.

Points A and C are two points on the price consumption curve and two points on the demand curve. The total effect of a $1/unit decrease in the price of good 1 can be found by measuring the movement from A to C: for x_1, the total effect is +25 units and for x_2, the total effect is zero ($x_2{}^* = 16\frac{2}{3}$ before and after the price shock).

The total effect can be directly observed. With the initial price, we can see the consumer purchase 25 units of good 1 and $16\frac{2}{3}$ of good 2. We see the price of good 1 fall by \$1/unit and watch the consumer respond by buying 25 units more of x_1 and leaving the amount of x_2 unchanged.

We are now ready for the key move. We will hypothetically take away exactly \$25 of income so we can find the optimal solution on the imaginary, dashed line. The consumer does not actually have income taken away. It is a thought experiment. Working out what the consumer would do in this hypothetical situation allows us to split the total effect into its constituent parts.

STEP Change income to \$75 (notice that the budget line now lies on top of the dashed budget line) and run Solver.

You can safely ignore the steeper line in the chart—all we want is point B, the optimal solution with the dashed budget line. Solver tells us that point B is 37.5,12.5. This gives us three points to consider:

- Point A: Initial: At $m = 100, p_1 = 2, x_1{}^* = 25, x_2{}^* = 16\frac{2}{3}$.
- Point B: Unobserved: At $m = 75, p_1 = 1, x_1{}^* = 37\frac{1}{2}, x_2{}^* = 12\frac{1}{2}$.
- Point C: New: At $m = 100, p_1 = 1, x_1{}^* = 50, x_2{}^* = 16\frac{2}{3}$.

Look carefully at the three points and concentrate on how points B and C differ: C uses new p_1 with original m, while B is based on new p_1 with adjusted m (adjusted in a special way so that the dashed line goes through point A).

With these three points, we can compute total, income, and substitution effects for x_1 and x_2. The three effects are shown by arrows on the axes of Figure 4.20. This is a complicated graph. Take your time and read it with care. Try to separate the different elements and lines to different parts of the problem: initial (A), new (C), and intermediate positions (B).

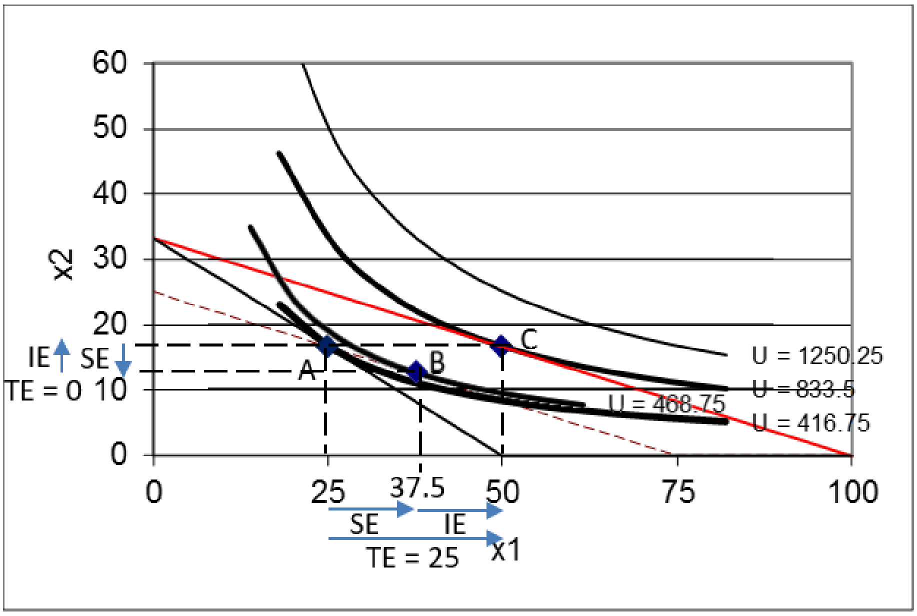


Figure 4.20: Total (TE), income (IE), and substitution (SE) effects.

There are effects measured from one point to another for both x_1 and x_2. These Δs are calculated the usual way as $new - initial$. For x_1, we find:

- SE: A to B: $37\frac{1}{2} - 25 = 12\frac{1}{2}$
- IE: B to C: $50 - 37\frac{1}{2} = 12\frac{1}{2}$
- TE: A to C: $50 - 25 = 25$

Notice that the total effect (TE) can be found by computing the difference from A to C ($50 - 25 = 25$) or taking advantage of the fact that SE + IE = TE, so $12.5 + 12.5 = 25$. The effects for x_1 are all computed along the x axis in terms of units of x_1.

Analyzing the effect on x_2 of a change in p_1 gives us cross income and substitution effects for x_2, which are shown by arrows on the y axis, in Figure 4.20.

- SE: A to B: $12\frac{1}{2} - 16\frac{2}{3} = -4\frac{1}{6}$
- IE: B to C: $16\frac{2}{3} - 12\frac{1}{2} = 4\frac{1}{6}$
- TE: A to C: $16\frac{2}{3} - 16\frac{2}{3} = 0$

On x_2, the income and substitution effects work against each other. The substitution effect, from A to B, lowers the amount of x_2 since p_1 fell, making x_2 more expensive relative to x_1. But when we move from B to C, the income effect exactly cancels out the SE. The fall in p_1 has increased our purchasing power and, since x_2 is a normal good, we want to buy more of it.

It is a property of the Cobb-Douglas utility function that the cross IE and SE effects cancel each other out, leaving a zero total effect. This is not a usual or common result and it demonstrates how the functional form imposes structure on the demand curve.

Let's return now to x_1 and focus on its substitution effect, which we know is always negative. This leads immediately to a question: If the SE is always negative, then why is it +12.5 in Figure 4.20?

The answer to this apparent contradiction is that the negative refers to the relationship, not the actual value of the SE. Given that price fell, an increase in quantity purchased is consistent with a negative effect because it is the relationship between the two variables that is being described as negative.

Likewise, the sign of the income effect can be tricky. The key is to pay attention to which shock variable is being considered. The income effect measured as the response to a change in income is positive, in this case, because as I move from B to C, my income is increased and I respond by increasing my optimal consumption of good 1.

Now you might ask, "If the two effects work together, then how is the substitution effect negative and the income effect positive?" This is because we defined the income effect as the response to a change in income, like the movement from point B to C in Figure 4.20. But, if you remember, this example began with a decrease in the price of good 1. The decrease in the price of good 1 can be interpreted as an increase in income, in the sense of greater purchasing power. If we tie the 12.5 increase in good 1 from the income effect to the *decrease* in price of good 1, we see that this negative relationship reinforces the negative substitution effect and gives a negative total effect.

Now that we know how the income and substitution effects combine to form the total effect of a price change, we can show how easy it is to compute them from a reduced form solution.

We first have to solve the model analytically and get a reduced form expression as a function of m and p_1. We have done this before for a Cobb-Douglas utility function and found

$$x_1{*} = (\frac{c}{c+d})\frac{m}{p_1}$$

If we substitute in $c = d = 1$, we have

$$x_1{*} = \frac{m}{2p_1}$$

At $m = 100$ and $p_1 = 2$, $x_1{*} = 25$. This is the initial solution (point A).

If p_1 falls to \$1/unit, then we plug in $m = 100$ and $p_1 = 1$, which gives the new solution (point C), $x_1{*} = 50$. The total effect is $50 - 25 = 25$.

To find the SE, we need point B. We use the reduced form expression to compute quantity demanded with adjusted m (\$75) and new p_1 (\$1/unit).

$$x_1{*} = \frac{m}{2p_1} = \frac{[75]}{2[1]} = 37.5$$

Once we have point B, we have split the total effect from A to C and we can compute the SE and IE by going from A to B and B to C, respectively. The SE is $37.5 - 25 = 12.5$ and the IE is $50 - 37.5 = 12.5$. These results agree with our earlier work.

Income and Substitution Effects via Graphs

Income and substitution effects are complicated. Figure 4.20 is not easy to understand. There are three budget lines and a lot going on. So what is so important about income and substitution effects that makes it worthwhile to master them?

Income and substitution effects hold the key to explaining how we can get a Giffen good. They mark real progress in economics, settling a long debate about whether or not upward sloping demand curves are possible. We will deconstruct the income and substitution effect graph (Figure 4.20), examining each layer one at a time, to show the source of Giffen behavior.

We begin with Figure 4.21. On the left we have the initial optimal solution and the right displays a single point on the demand curve (not shown).

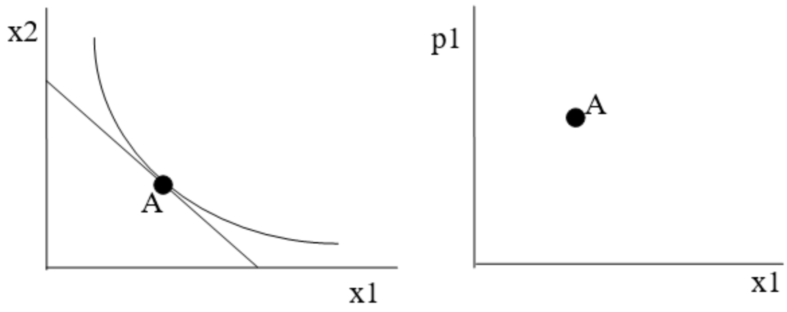


Figure 4.21: The initial solution.

Next, we decrease the price of good 1, as shown in Figure 4.22, which creates a new budget line. We know the consumer will re-optimize and choose a new optimal solution along the new, flatter line, but Figure 4.22 does not show this new solution quite yet. Instead, it shows the point B solution on a dashed line with the income that would have to be taken away to cancel out the increased purchasing power from the price decrease.

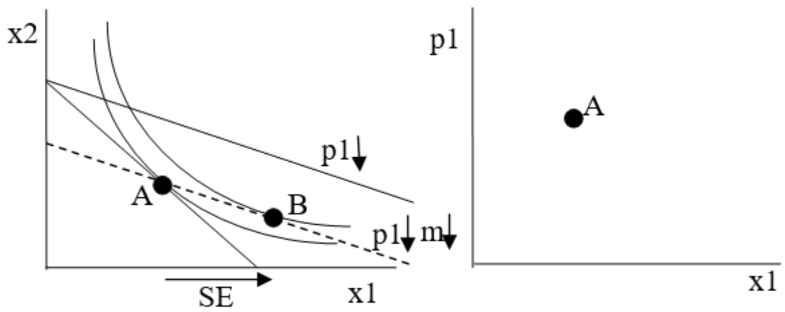


Figure 4.22: A p_1 decrease and imaginary budget constraint.

Figure 4.22 shows the optimal solution, point B, for the hypothetical situation with lower p_1 and adjusted m. The rightward pointing arrow is the SE for x_1 is the substitution effect, from point A to B on the x axis. The dashed line has a flatter slope (new p_1 is less than initial p_1) through point A. This guarantees that B is to the *right* of A. This is why the SE is always negative.

It is impossible to draw a point B to the left of A without making the indifference curves cross. With MRS $= \frac{p_1}{p_2}$ at A, lowering p_1 and adjusting m so dashed line goes through A, means the consumer must move southeast to find the highest indifference curve tangent to the dashed line.

Now, we are ready to show point C. We have a known negative substitution effect and all that remains to be done is to find the indifference curve tangent to the new budget line (with lower p_1). The key insight is that there are several possible positions for point C. Figure 4.23 shows three possibilities.

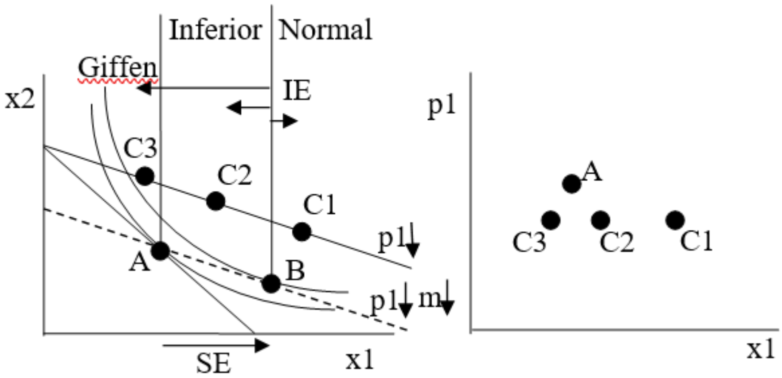


Figure 4.23: Understanding Giffen behavior.

Figure 4.23 shows that the final position of point C depends on whether the good is normal or inferior, with a subcategory of inferior goods that are Giffen.

- C1: Good 1 is a normal good so the income effect from B to C works together with the movement from A to B and we end up at point C1. In this case, and for any point C to the right of B, we get a downward sloping demand curve.

- Good 1 is an inferior good so the income and substitution effects work against each other. The movement from B to C will be to the left and leave us with a point C to the left of B. There are two possibilities:

  1. C2: The income effect pushes the consumer to buy less x_1, but it is less than the substitution effect (which leads to buying more x_1 as p_1 falls). We end up at point C2 between A and B and the demand curve is still downward sloping.
  2. C3: The income effect not only works against the substitution effect, it is stronger, swamping it. Point B to C moves in the opposite direction than A to B and and is bigger than A to B. This leaves the consumer to the left of B at point C3. The demand curve is upward sloping. This is a Giffen good.

It can be difficult to draw a Giffen good correctly because the indifference curves cannot cross. So, in Figure 4.23, the space available for point C3 is tight—C3 can only fit to the left of A and to the right of the indifference curve that is shown tangent to B.

Figure 4.23 also makes clear that it is the indifference curves, which come from the utility function, that determine how quantity demanded responds to a change in price. How a good generates utility (i.e., whether utility is Cobb-Douglas, quasilinear, perfect complements, or another functional form) determines whether it is normal, inferior, or Giffen.

The decomposition of the total effect into income and substitution effects provides the condition which must hold for Giffen behavior: the income effect must work against the substitution effect and be bigger. We can reinforce this key insight with a mathematical expression that gives more detail on exactly how we get Giffenness.

The Slutsky Equation

In 1915, decades after the supposed spotting of a Giffen good during the Irish potato famine, Eugen Slutsky published a paper in an Italian journal that showed how to decompose the total effect of a price change into income and substitution effects. He had a mathematical expression that showed how it was possible to get an upward sloping demand curve!

Unfortunately, his work went unnoticed. Twenty years later, John R. Hicks (a Nobel laureate in 1972) and R. G. D. Allen rediscovered the ideas in Slutsky's paper. Sometimes, the idea of income and substitution effects are referred to as Slutsky-Hicks or Slutsky-Hicks-Allen. We will keep it simple and call it the Slutsky Equation.

The Slutsky Equation, which we will not derive, says in mathematical terms something that we already know: The total effect of a price change can be expressed as the sum of a substitution and an income effect. It turns out that there are several ways to express the decomposition with a Slutsky Equation. Here are two versions:

$$\frac{\Delta x_1}{\Delta p_1} = \frac{\Delta x_1^{SE}}{\Delta p_1} + \frac{\Delta x_1^{IE}}{\Delta p_1}$$

$$\frac{\Delta x_1}{\Delta p_1} = \frac{\Delta x_1^{SE}}{\Delta p_1} - x_1 * \frac{\Delta x_1}{\Delta m}$$

Both equations say the same thing: the total effect, $\frac{\Delta x_1}{\Delta p_1}$, is equal to the substitution effect, $\frac{\Delta x_1^{SE}}{\Delta p_1}$, plus the income effect. Where they differ is how they express the income effect.

Look carefully at the denominators. The income effect in the first equation has a Δp_1 denominator, like the other two terms. What Slutsky figured out was that the income effect of price change, $\frac{\Delta x_1^{IE}}{\Delta p_1}$, could be written as $-x_1{*}\frac{\Delta x_1}{\Delta m}$. In other words, the income effect channel of the price change can be expressed as the amount of good 1 initially purchased times the change in x_1 as income changes (the slope of the Engel curve). Notice the minus sign, which picks up the fact that when price falls, that is like an increase in income.

Now we can really see how to get a Giffen good, which has an upward sloping demand curve so $\frac{\Delta x_1}{\Delta p_1} > 0$. Since the first term, the substitution effect is always negative, we definitely need an inferior good so that $\frac{\Delta x_1}{\Delta m} < 0$ so that the second term is positive. Obviously, if the good is extremely inferior, so that $\frac{\Delta x_1}{\Delta m}$ is much less than zero, we might get a Giffen good.

But the Slutsky Equation reveals another way to get Giffen behavior. A large opposing income effect can be obtained by the good being inferior and the consumer buying a lot of it so that $-x_1{*}\frac{\Delta x_1}{\Delta m}$ is a big positive number to outweigh the negative substitution effect. If the good is merely inferior, but the consumer buys little of it, then it less likely to be Giffen.

This is why we look for Giffen behavior in *staples*, basic commodities that comprise a large share of the budget. Potatoes for the Irish, rice for Asians, and tortillas for Mexicans are three examples that economists have examined for Giffen behavior. For a poor person, these items could be consumed in large quantities, yet, as income rises, quantity demanded falls so they are inferior goods. The combination of a large $x_1{*}$ and $\frac{\Delta x_1}{\Delta m} < 0$ could produce a large, positive $-x_1{*}\frac{\Delta x_1}{\Delta m}$ term that is bigger than the negative substitution effect.

Remember how we generated Giffen behavior with *GiffenGoods.xls* in the previous section? We increased the price from \$1/unit to \$1.1/unit and optimal x_1 rose from 44 to 48.6, while optimal x_2 fell dramatically from 11 to around 1.5. Notice how x_1 is a staple, dominating the amounts purchased of the two goods.

We know its Giffen, but is x_1 also inferior? Let's find out.

STEP Open *GiffenGoods.xls* and proceed to the *Optimal1* sheet. Click the Reset button and run Solver to make sure you are at the optimal initial solution of 44,11. Increase m to 60 and run Solver. What happens?

Yes, as we know must be true (since we know x_1 is a Giffen good), x_1 is an inferior good: optimal x_1 fell (to 39) as income increased to $60. Giffenness requires that x_1 be inferior and this example also reflects the fact that concentration of the consumer's budget on an inferior good contributes to the production of a Giffen response.

The *Biblio* sheet in *GiffenGoods.xls*, from the previous section, had several references to papers trying to find Giffen goods, yet the jury is still out. What is unquestioned, however, is the theoretical requirement: it must be an inferior good so that the IE is in the opposite direction and larger than the SE.

The Slutsky Equation also enables us to fine tune a statement that is, strictly speaking, false. Introductory economics students around the world learn the *Law of Demand*: when price increases, ceteris paribus, quantity demanded must fall. In other words, holding everything else constant, quantity demanded and price are inversely related and demand is always downward sloping.

This is fine, at the introductory level, where we do not want to confuse beginning students, but we know that an upward sloping demand curve is possible—it is called a Giffen good. They are a violation of the "Law" of Demand and we know they could exist. When their price rises, so does quantity demanded.

Can we rehabilitate the Law of Demand so there is no exception? Yes, we can. Our knowledge of income and substitution effects points the way. We can more precisely define the Law of Demand. By inserting a qualifying clause, we can get the Law of Demand to be exactly right: *If the good is normal, then quantity demanded falls as price rises, ceteris paribus.* That is guaranteed to be true because a normal good has an income effect that works together with the substitution effect. Thus, there is no way to get Giffenness.

The Cobb-Douglas utility function cannot give Giffen behavior. The reduced form solution, ${x_1}^* = (\frac{c}{c+d})\frac{m}{p_1}$, means that $\frac{d{x_1}^*}{dm} = (\frac{c}{c+d})\frac{1}{p_1} > 0$ so the income effect, $-x_1{}^*\frac{dx_1{}^*}{dm}$, is negative. This means the IE and SE are both negative and work together so there is no way the Cobb-Douglas utility function can generate Giffenness.

TE = SE + IE

Income and substitution effects are used by economists to better understand the demand curve and to explain Giffen behavior. By disassembling the total effect of a price change, the Slutsky Equation shows how a Giffen good can arise if the income effect opposes and swamps the substitution effect (which generates an upward sloping relationship between price and quantity demanded).

Given a utility function and budget constraint, we find the initial optimal solution (point A). A price change will lead to a new optimal solution (point C) which we can use to compute the total effect. We can then use the Income Adjuster Equation to find a hypothetical point B that splits the total effect into substitution and income effects.

Given a reduced form expression of $x^* = f(p, m)$, we can find points A, B, and C by evaluating the expression at the appropriate p and m values to compute points A, B, and C.

The Slutsky Equation is a mathematical presentation of income and substitution effects. The math gives us the insight that the income effect, $-x_1{}^*\frac{\Delta x_1}{\Delta m}$, is composed of initial optimal x_1 times the response of x_1 to an income change. This reveals that Giffenness is more likely to be found in inferior goods that also attract a high concentration of the consumer's budget.

There are even more ways to express the Slutsky Equation than the two used in this section. Instead of altering income to allow the consumer to buy the initial bundle of goods, you can change income to allow the consumer to be on the initial indifference curve. This is sometimes referred to as the Hicks substitution effect.

Exercises

1. Reproduce, using Word's Drawing Tools, Figures 4.21, 4.22, and 4.23, explaining each graph in your own words.

2. Repeat question 1, with one key change: apply a price *increase* in good 1 (instead of a price decrease).

3. In stating the Law of Demand, some economists choose to include a condition that the good is normal, like this: If the good is a normal good, then price and quantity demanded are inversely related, ceteris paribus. Why is the normal good clause needed?

4. Given the demand function, $x_1{}^* = 20 + \frac{m}{20p_1}$, compute the total, income, and substitution effects when price falls from \$5 to \$4/unit, with income of \$1000. Show your work.

5. Use the *Optimal1* sheet in *GiffenGoods.xls* to find points A, B, and C for a shock in p_1 from \$1 to \$1.1/unit. Compute the TE, SE, and IE for $x1$. Show your work and explain what you did.

References

The epigraph is from the biography of Slutsky available at the New School's History of Economic Thought website, www.hetwebsite.net/het/. The site was created and is maintained by Gonçalo L. Fonseca. There are sketches of hundreds of economists, links to other resources, and descriptions of various schools of thought in economics. The intellectual history of economics is fascinating and this website is a wonderful place to browse.

I never saw Slutsky's work until my own was very far advanced . . . Slutsky's work is highly mathematical, and he does not give much discussion about the significance of his theory.

J. R. Hicks

4.7 More Practice with IE and SE

This chapter uses a quasilinear utility function to provide practice working with income and substitution effects. There is a surprising twist when using the quasilinear functional form. See how fast you can figure it out.

STEP Open the Excel workbook *IncSubEffectsPractice.xls*, read the *Intro* sheet, then go to the *OptimalChoice* sheet.

Notice that the absolute value of the MRS is less than the price ratio. Because the slope of indifference curve at 16.25,10.75 is less than the slope of the budget constraint, we know the consumer should travel northwest along the budget constraint, buying more x_2 and less x_1, until the MRS $= \frac{p_1}{p_2}$.

STEP Run Solver to find the initial optimal solution. Figure 4.24 shows this result.

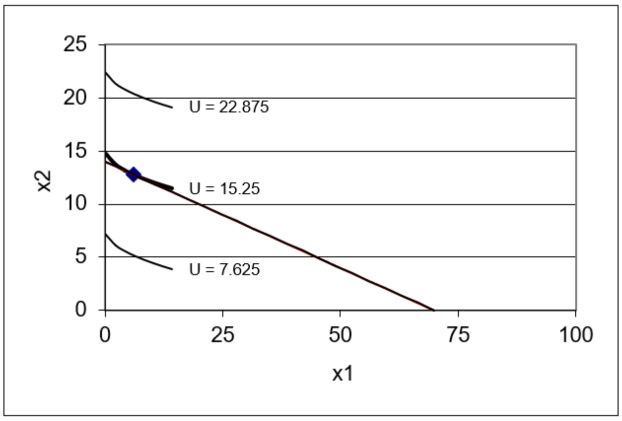


Figure 4.24: Initial optimal solution.
Source: IncSubEffectsPractice.xls!OptimalChoice

STEP Proceed to the *CS1* sheet. It shows a comparative statics analysis of an increase in the price of good 1 from \$2/unit to \$7/unit in \$1 increments. It also charts the results as an inverse demand curve for x_1.

The demand curve tracks the total effect of a price change. When the price of good 1 rises from \$2 to \$3, the quantity demanded falls from $6\frac{1}{4}$ to $2\frac{7}{9}$. By subtracting the new from the initial value, we see that the total effect is a decrease of $3\frac{17}{36}$ units of x_1, displayed in cell F13 as -3.47222.

Income and substitution effects explain how this total effect came to be by dismantling the total effect into two parts that add up to the total.

The substitution effect tells us how much less the consumer would have purchased when price rises strictly from the fact that the relative prices of the two goods have changed. We compute how much income we have to give the consumer to cancel out the reduced purchasing power caused by the price increase to focus exclusively on the relative price change. The substitution effect is always negative.

Figure 4.25 shows a typical decomposition of the total effect (TE) into the substitution effect (SE) and income effect (IE) with indifference curves suppressed to highlight the budget lines under consideration.

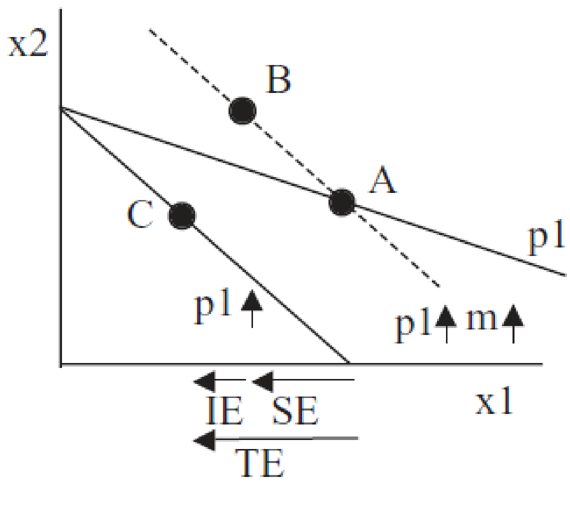


Figure 4.25: Typical TE, SE, and IE with p_1 increase.

From point A, price rose and the consumer will now be at point C on the new budget line (labeled $p_1 \uparrow$). The dashed line is the result of a hypothetical scenario in which the consumer has been given enough income to purchase the initial bundle A. Notice how the original budget line and the dashed line go through point A. The dashed line has a higher price, but also a higher income. Thus, the movement from point A to point B reflects solely the different relative prices in the goods, without any change in purchasing power. This is the substitution effect.

While the substitution effect is focused on relative prices, the income effect is that part of the response in quantity demanded when price changes that is due to changed purchasing power. From point B, a decrease in income from the dashed to the new budget line leads to a decrease in x_1 (at point C). Thus, x_1 is a normal good from point B to C in Figure 4.25 and the two effects are working in tandem. The demand curve is guaranteed to be downward sloping for this price change.

In the *CS1* sheet, we have seen that the demand curve is downward sloping because quantity demanded falls when price rises. But an open question still remains: Do the income and substitution effects work as in Figure 4.25?

We know point A, the initial optimal solution, is $x_1{}^* = 6.25$ when $p_1 =$ \$2/unit and point C is about 2.78 units of x_1 when price rises to \$3/unit. We need point B to do the income and substitution effects analysis.

The first step in finding point B is to use the Income Adjuster Equation to compute how much income to give the consumer in order to cancel out the effect of the reduced purchasing power.

$$\Delta m = x_1{}^* \Delta p_1$$

$$\Delta m = [6.25][+1]$$

STEP On the *OptimalChoice* sheet, set cell B16 to 3.

The chart updates, showing the new budget constraint in red (swinging in since price rose) and the dashed line. To find point B, we need the optimal solution for the dashed line constraint so we need to change in income on the sheet.

STEP Set cell B18 to 146.25. This applies the dashed line budget constraint to this problem. Run Solver to find point B.

Your result might surprise you. Solver says the optimal solution is about 2.78 for x_1, but that is the same answer we had for point C. What is going on here?

We turn to analytical work to shed light on this mysterious result. Following the procedure in section 3.2, we found this reduced form solution for the quasilinear utility function, $U = x_1^c + x_2$:

$$x_1* = (\frac{p_1}{cp_2})^{\frac{1}{c-1}}$$

We use the initial values of c and p_2 in the *OptimalChoice* sheet to simplify things a bit:

$$x_1* = (\frac{p_1}{[0.5][10]})^{\frac{1}{[0.5]-1}} = (\frac{p_1}{5})^{\frac{1}{-0.5}} = (\frac{p_1}{5})^{-2} = (\frac{5}{p_1})^2 = \frac{25}{p_1^2}$$

This is the same kind of expression, $x_1* = f(p_1, m)$, that we used in the previous section for a Cobb-Douglas utility function, $x_1* = \frac{m}{2p_1}$, to find points A, B, and C.

You might be puzzled. Exactly where is m for the quasilinear reduced form expression for x_1? It is not there, although a mathematician might say that we could easily include it by writing the reduced form expression like this:

$$x_1* = \frac{25}{p_1^2} + 0m$$

The fact that m does not affect optimal x_1 for a quasilinear utility function is the source of the surprising result for point B. We can apply the usual procedure for finding points A, B, and C with a reduced form expression to show this.

Point A is the initial optimal x_1 solution so we plug in $p_1 = 2$ and find $x_1* = \frac{25}{2^2} = 6.25$.

Point C is the new optimal x_1 solution so we plug in $p_1 = 3$ and find $x_1* = \frac{25}{3^2} = \frac{25}{9} = 2\frac{7}{9}$.

Point B is found using new p_1 and adjusted m, \$146.25. But notice that adjusted m is irrelevant because it does not affect x_1. Point B is $x_1* = 2\frac{7}{9}$, the same as point C.

Figure 4.26 shows what is going on here. Unlike the typical case, there is no income effect at all with quasilinear utility, so TE = SE. As usual, the substitution effect is the move from point A to B and the income effect is the movement from B to C. The IE is zero because C is directly below B. The total effect is A to C.

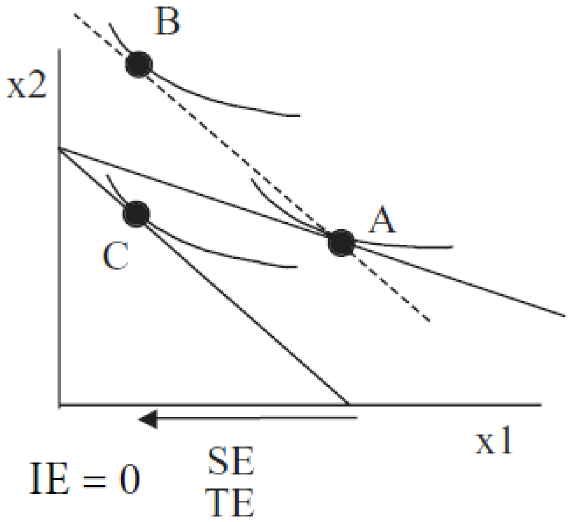


Figure 4.26: TE, SE, and IE with quasilinear utility.

It is the utility function that is driving this result. A utility function with the functional form $U = f(x_1) + x2$ has no income effect because the indifference curves are vertically parallel. If you shift the budget line via an income shock, the new tangency point will be directly above or below the initial point. In other words, the income consumption curve is vertical. Thus, the total effect is composed entirely of the substitution effect. This is the curious twist produced by the quasilinear functional form.

We saw that the income consumption curve is vertical and Engel curve is horizontal in section 4.2 (see Figure 4.7). Economics is certainly cumulative and ideas learned are often worth remembering because they tend to show up again.

Finally, notice that we now know that quasilinear preferences cannot yield Giffen behavior. After all, if the substitution effect is always negative and the income effect is zero, there is no way for the total effect to ever be positive.

Quasilinear Preferences Yield Zero Income Effects

Splitting a total effect into income and substitution effects works for any utility function. After finding the total effect, the Income Adjuster Equation can be used to determine the income needed to cancel out the change in purchasing power from the price change (i.e., setting the imaginary, dashed budget line). Finding the optimal solution with the new price and adjusted income budget constraint determines point B and allows us to split the total effect in two parts.

Of course, the component parts, SE and IE, need not be equal nor share the same sign. We know that Giffen goods arise when the income effect opposes and swamps the always negative substitution effect.

In the case of quasilinear preferences, we have a situation where there is no income effect. The Slutsky decomposition still applies, however, with the total effect being entirely composed of the substitution effect.

Exercises

1. Click the Reset button on the *OptimalChoice* sheet and apply a price decrease for good 1 from \$2/unit to \$1.90/unit. Compute the total, substitution, and income effects. Show your work.

2. Use Word's Drawing Tools to draw a graph similar to Figure 4.26 that shows the total, substitution, and income effects from the 10 cent decrease in price from question 1.

Questions 3 and 4 are difficult. Revisit questions 2 and 3 in *EngelCurvesPracticeA.doc* (in the *Answers* folder in the *MicroExcel* archive) for more detail on the corner solution for this utility function at low levels of income.

3. With quasilinear utility, the income consumption curve is vertical and the Engel curve horizontal only above a threshold income level. At very low levels of income, we get a corner solution. Click the Reset button on the *OptimalChoice* sheet and set income to 10. This will generate a corner solution. Compute the total, substitution and income effects from a 10 cent price increase in good 1 (from 2 to 2.1). Show your work.

4. Use Word's Drawing Tools to draw a graph depicting your results for question 3.

References

The epigraph comes from page 19 of the second edition of *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory* by John R. Hicks. This remarkable book was explicitly cited in the press release announcing that Hicks had won the Nobel Prize in Economic Science in 1972 (with Kenneth Arrow). "In his most well-known work, the monograph, *Value and Capital*, published in 1939, Hicks abandoned this [formal] tradition and gave the [general equilibrium] theory an increased economic relevance." See www.nobelprize.org/prizes/economic-sciences/1972/press-release/.

As mentioned in the previous section, the history of income and substitution effects is complicated. Hicks (and Allen) figured out that the total effect could be decomposed into income and substitution effects in the 1930s, two decades after Slutsky's work. Once Slutsky was rediscovered, Hicks and Allen gave him credit and made the economics profession aware of his contribution. Hicks wrote in *Value and Capital* that "The present volume is the first systematic exploration of the territory which Slutsky opened up" (p. 19).

Usually the first question anyone asks about a proposed new tax is "Who pays?" and about a tax cut is "Who benefits?"

Joel Slemrod and Jon Bakija

4.8 A Tax-Rebate Proposal

This section examines a tax-rebate plan that provides further practice with the logic of income and substitution effects. This application shows that they are more than an intellectual curiosity.

The heart of the idea is for the government to reduce consumption of a particular good, for example, gasoline, without hurting the consumer.

The idea is to tax a good and then turn around and rebate (give back) all of the tax revenue to the consumer. Can we alter the consumer's choices without lowering satisfaction? We keep things simple by ignoring administrative costs of collecting the tax and rebating it so the tax and rebate leaves the consumer's income unchanged. Proponents point out that the government is not making any money (all of the tax revenue raised is refunded back) so the consumer is not going to be hurt.

Opponents contend that this scheme will have no effect because the rebated tax will immediately be spent on the taxed good and we will end up right where we started.

Who is right? We use the Theory of Consumer Behavior to find out. Along the way, income and substitution effects will come into play.

A Concrete Example

STEP Open the Excel workbook *TaxRebate.xls* and read the *Intro* sheet, then go to the *QuantityTax* sheet.

We have a Cobb-Douglas utility function with an option to apply a per unit (quantity) tax on good 1. The workbook opens with no tax and the consumer maximizing satisfaction by buying the bundle 25,50, yielding $U^* = 1250$.

We begin by applying a quantity tax.

STEP Change cell B21 to 1. Notice that a new budget line appears. The consumer cannot afford the original bundle and must re-optimize. Run Solver to find the new optimal solution.

You should find that the consumer will now buy the bundle $16\frac{2}{3}$,50 and maximum utility falls to 833.33. Cell B22 shows that the government collects \$16.67 (\$1/unit tax on the 16.67 units purchased).

The idea behind the tax-rebate proposal called for rebating the tax revenue so that the consumer would not be hurt by the tax. We need to implement the rebate part of the proposal.

STEP Change cell B18 to 116.67. This shifts the budget constraint out. Run Solver to find the optimal solution.

You should find that the consumer optimizes by purchasing 19.445 units of x_1 and 58.335 units of x_2.

This result presents us with a problem. This is not the tax-rebate scheme the government envisioned. After all, the government is collecting more tax revenue (\$19.445) than the consumer is getting as a rebate (\$16.67).

Instead of giving the consumer \$16.67, let's give her \$19.445. What does the consumer do in this case?

STEP Change cell B18 to 119.445. This shifts the budget constraint out a little bit more. Run Solver to find the optimal solution.

Now the consumer buys a little more x_1, just over 19.9 units. But we still do not have a revenue neutral policy. We need to increase m again. This process of repeatedly doing the same thing is called *iteration*.

STEP Set the cell B18 value to \$100 (initial m) plus the amount of tax revenue in cell B22. Run Solver.

You can see that we are converging because the increases to income keep getting smaller and smaller. There is a tax rebate that yields an optimal x_1 that generates a tax revenue that exactly equals the tax rebate. The value of this tax rebate is \$20.

STEP Set cell B18 to \$120. Run Solver.

You should see that the optimal solution is 20,60 and maximum utility is 1200. If Solver is off by a little bit (this is false precision), you can enter 20 and 60 in cells B11 and B12. Since they buy 20 units of x_1, the consumer is paying \$20 in tax. Since they are getting a tax rebate of \$20 (m is set is 120), the tax they pay is exactly canceled out. We are ready to evaluate this program.

Who's Right?

Proponents argued that by taxing the good and then turning around and rebating (giving back) the tax revenues to the consumer, we can alter the consumer's choices without lowering satisfaction. Since the government is not making any money (all of the tax revenue raised is refunded back), the consumer is not going to be hurt.

Clearly the supporters of the tax-rebate proposal are wrong. The consumer had an initial $U^* = 1250$ and now has a new $U^* = 1200$. While we cannot meaningfully say that utility has fallen by 50 (because utility is measured on an ordinal, not cardinal scale), we can say that utility has fallen. Thus, in fact, the consumer is hurt by the tax-rebate proposal.

Critics, on the other hand, believed that this scheme will have no effect since the rebated tax will immediately be spent on the taxed good and we will end up right where we started.

Because the consumer went from an initial bundle of 25,50 to 20,60 after the \$20 tax-rebate, it is obvious that the critics are wrong also. This consumer has altered purchasing plans and is, in fact, buying less x_1.

So, wait, who's right—the critics or the supporters of the scheme? Neither. They are both wrong. Income and substitution effects will help us explain why.

We return to the original problem without a tax or rebate and the initial solution of 25,50. The \$1/unit tax is just like a price increase. We can find point B and compute the substitution and income effects from such a price change.

We first use the Income Adjuster Equation.

$$\Delta m = x_1{}^* \Delta p_1$$

$$\Delta m = [25][+1]$$

This result says that a $25 increase in income to $125 will allow us to buy the initial bundle.

STEP Set income in cell B18 to 125 (and confirm that there is a $1/unit tax in cell B21) and run Solver.

The optimal solution is $20\frac{5}{6}, 62\frac{1}{2}$. We have points A, B, and C so we can compute total, substitution, and income effects of the $1/unit price increase due to the tax without any rebate.

- SE (A to B): $20\frac{5}{6} - 25 = -4\frac{1}{6}$
- IE (B to C): $16\frac{2}{3} - 20\frac{5}{6} = -4\frac{1}{6}$
- TE (A to C): $16\frac{2}{3} - 25 = -8\frac{1}{3}$

Figure 4.27 displays these results with each point signifying a tangency between the budget line and an indifference curve (not drawn in to make it easier to read the graph).

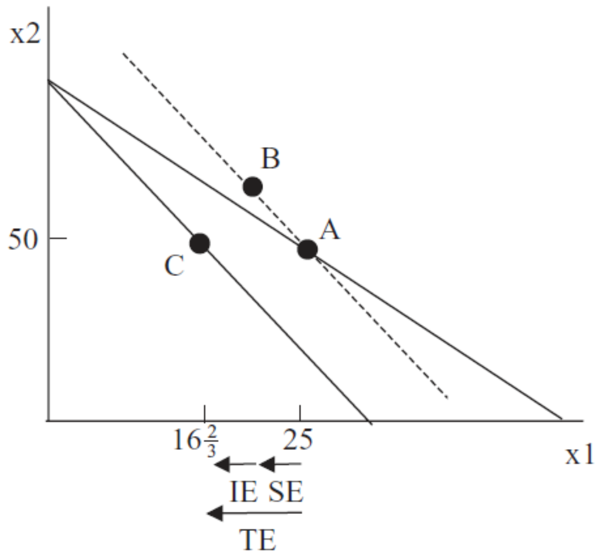


Figure 4.27: TE, SE, and IE for tax without rebate.

The tax-rebate proposal is closely related to Figure 4.27. The tax is like a price increase that moves the consumer from A to C and the rebate is like an income effect that moves the consumer from C to B.

However, if you look carefully, the changes in income are not the same. In the tax-rebate proposal, the revenue-neutral rebate is \$20, whereas in our income and substitution effect work we gave the consumer \$25 to be able to purchase the original bundle. A \$25 rebate is not revenue neutral because the consumer buys only $20\frac{5}{6}$ units of x_1 so the government ends up losing revenue. The rebate has to be \$20 to be consistent with the break-even logic of the proposal.

In addition to the income and substitution effects, Figure 4.28 adds point D, which shows the optimal solution given the tax-rebate proposal. Point D (at coordinate 20,60) has utility of 1200, which is, of course, lower than point B (the combination $20\frac{5}{6}, 62\frac{1}{2}$ yields just over 1300 units of utility). More importantly for the purposes of evaluating the proposal, utility at point D is less than utility at point A (where 25,50 generates $U^* = 1250$).

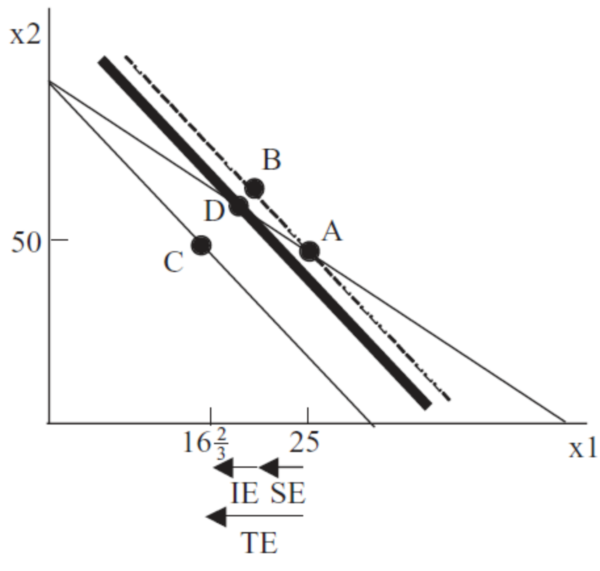


Figure 4.28: Understanding the tax-rebate proposal.

The key to the analysis lies with point D in Figure 4.28. It has to be on the initial budget line to fulfill the revenue-neutral condition of the proposal. But we know point A was the initial optimal solution on that budget line, so we can deduce that the consumer prefers point A to point D (and any other point on the initial budget line) and will suffer a decrease in satisfaction if the tax-rebate proposal is implemented.

Tax-rebate Schemes

Taxes are often used to pay for government services and fund programs deemed worthy by society, but they can also be corrective. Taxes on specific products can discourage particular activities (think cigarettes and smoking).

Simultaneously taxing a good and rebating the tax revenue periodically appears as a policy proposal (often with regard to gasoline). Proponents claim the rebate cancels out the price increase from the tax. The scheme is related to income and substitution effects. The tax is like a price increase and the rebate is like an income effect.

Although similar to income and substitution effects, there is one important difference in tax-rebate proposals: a revenue-neutral rebate does not return enough income to allow the consumer to buy the pre-tax bundle or to reach the pre-tax level of satisfaction. Thus, the consumer cannot reach the initial level of satisfaction.

It is true, however, that a tax-rebate policy will alter consumption patterns. Whether the loss in utility is compensated by the changed consumption pattern is a different question.

Exercises

1. Analytically, we can show that the demand curves for goods 1 and 2 with a Cobb-Douglas utility function (where $c = d$) are $x_1{}^* = \frac{m}{2(p_1+Q_Tax)}$ and $x_2{}^* = \frac{m}{2p_2}$. Use these demand functions to compute the income, substitution, and total effects for x_1 for a \$1/unit tax. Show your work.

2. We know that the tax-rebate scheme gives back too little income to return the consumer to the initial level of utility (1250 units). With a \$1/unit tax, find that level of rebate where the consumer is made whole in the sense that $U^* = 1250$. Describe your procedure in answering this question.

3. At point D in Figure 4.28, is the MRS greater or smaller in absolute value than the price ratio before the tax-rebate scheme is implemented? How do you know this?

References

The epigraph is from page 87 of the fifth edition of *Taxing Ourselves: A Citizen's Guide to the Debate over Taxes* published in 2017 by Joel Slemrod and Jon Bakija. The book does not discuss the tax-rebate proposal covered in this chapter, but it is an excellent, user-friendly guide to the ever-present debate over taxes.

Government spending, taxing, and budgeting is part of the subdiscipline of economics called Public Finance. If you are interested in government's role in the economy or tax reform (including flat or consumption tax proposals), the history of the income tax in the United States, or how economists evaluate and judge taxes, this book is a great place to start.

Chapter 5

Endowment Models

> Our consumers could simply sit down and consume their endowments. But one consumer might, for example, be endowed with a lot of some good that she is not particularly fond of. She may wish to exchange some of that good for something she likes more.
>
> David M. Kreps

5.1 Introduction to the Endowment Model

This chapter introduces a wrinkle to the standard consumer theory model that greatly enhances its applicability. Instead of treating income as a given cash amount, we model the consumer as having a given initial endowment of goods that can be traded for other goods. This transforms the consumer into a combined consumer and seller.

Although the power of this approach may not be immediately obvious, we will see that a wide variety of examples such as saving/borrowing, charitable giving, and much more can be handled with this modification.

The Budget Constraint in an Endowment Model

Instead of the usual income (m) variable, an Endowment Model is characterized by a budget constraint that equates expenditures and revenues from sales out of the initial endowment.

$$p_1x_1 + p_2x_2 = p_1\omega_1 + p_2\omega_2$$

The term on the right-hand side says that the consumer has a given amount of each good, ω_1 and ω_2 (this is Greek letter *omega* so we have omega-one and omega-two). Because the initial amounts of each good are given, ω_1 and ω_2 are exogenous variables.

The starting amount of each good, the coordinate pair ω_1, ω_2, is called the *initial endowment.* If we multiply the initial amount of each good by the price of that good, as done in the right-hand side of the budget constraint equation, we get a dollar-valued amount that represents the total income that can be raised by selling the entire endowment.

Thus, the budget constraint says that spending (on the left-hand side) must equal the value of the consumer's assets (on the right-hand side).

The classic example to illustrate someone operating with an endowment model constraint is a farmer who goes to market with his crop. He sells his produce and, with the revenue obtained by selling, buys other goods. The core idea is that the farmer is a buyer *and* a seller.

Perhaps a more modern example is eBay. People sell all kinds of products and turn around and buy different products. It is a massive online garage-sale community. Once again, the core idea is that eBayers sell *and* buy.

In an Endowment Model, what the agent can buy depends on how much revenue is generated by sales. High prices for goods to be sold are a good thing from the agent's point of view because they generate a lot of revenue with which to buy other goods.

Because Endowment Models transform the consumer into a combined buying-selling agent, we can get different results than we saw in the Standard Model. One critical difference is that price increases lead to decreases in quantity demanded (assuming the good is normal), as usual, but as price keeps rising, we can cross the zero barrier and get *negative* quantity demanded! We will see that the agent switches from being a buyer to being a seller. This is a key idea.

Let's put these abstract ideas into concrete examples so we can understand what is going on with the Endowment Model.

STEP Open the Excel workbook *EndowmentIntro.xls*, read the *Intro* sheet, then go to the *MovingAround* sheet. Follow the instructions on the sheet to learn how we can create a budget line from a single point.

Just like the Standard Model, the agent faces a consumption possibilities frontier, also known as the budget line, that shows the feasible combinations. Bundles beyond the line are unattainable.

STEP Proceed to the *Properties* sheet.

Notice how we can use the value of the endowment to measure the agent's "income." Starting with 35,10 and $p_1 = 2, p_2 = 3$, the value of the endowment is \$100 (\$70 from x_1 and \$30 from x_2). The most x_2 the agent can have is $33\frac{1}{3}$, the y intercept and the maximum x_1, the x intercept, is 50.

The highlighted circle in the graph (reproduced as Figure 5.1) represents the initial endowment. From the initial allocation of 35,10, the agent can move northwest, selling x_1 and buying x_2. Or, the agent can decide to acquire even more x_1 by selling x_2 and buying x_1, which means traveling in a southeasterly direction. The slope of the constraint is the usual price ratio.

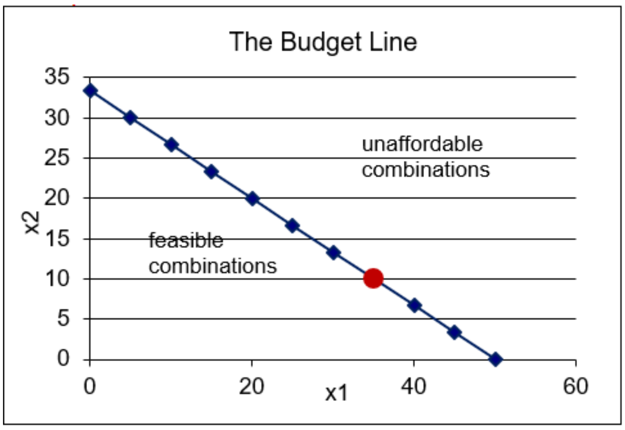


Figure 5.1: Endowment Model budget constraint.
Source: EndowmentIntro.xls!Properties

What will the consumer do in terms of buying and selling? In other words, where will the agent end up on the budget line? We do not know because we do not have any information on this agent's preferences. Before we tackle that problem, however, we need to see how the budget constraint changes when an exogenous variable is shocked.

STEP Proceed to the *Changes* sheet. Change p_1 (in K9) from 2 to 5.

This is different than before. Instead of the budget constraint pivoting about the y intercept (as in the standard, cash-income model), your screen should look like Figure 5.2. The budget constraint has pivoted or rotated as it did before, but the rotation is around the initial endowment. This is a critical difference.

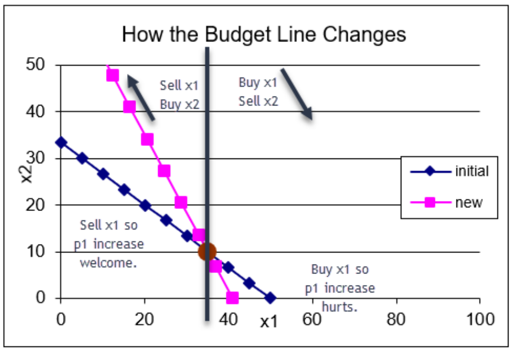


Figure 5.2: Endowment Model p_1 increase.
Source: EndowmentIntro.xls!Changes

The way the budget constraint has changed reveals important information. The price increase has improved the agent's consumption possibilities if she is planning on traveling northwest on the constraint. This makes sense because she would be a seller of good 1 and, with the higher price, she would have more money with which to buy good 2.

On the other hand, if she is a buyer, then we get the usual result that the budget line has rotated in and reduced the consumption possibilities.

STEP Click the Reset button and change p_1 (in K9) from 2 to 1.

Notice how the budget line has swiveled around the endowment again, but this time the agent is worse off if she is a seller and better off if she is a buyer.

STEP Click the Reset button and change p_2 (in K10) from 3 to 6. The result is exactly the same as when you changed p_1 (in K9) from 2 to 1.

This reveals a lesson: All that matters in the Endowment Model are relative prices, $\frac{p_1}{p_2}$. So $p_1 = 1, p_2 = 3$ is the same as $p_1 = 2, p_2 = 6$ and $p_1 = 10, p_2 = 30$ and any p_1 and p_2 whose p_1/p_2 ratio is $\frac{1}{3}$.

Finally, we consider shifts in the budget constraint. We cannot shift m (cash income) like we did in the Standard Model, but we can shock the initial endowment quantities of goods and this acts like a shift in income.

STEP Click the Reset button and change ω_1 (in K13) from 35 to 50.

The chart now looks like the usual increase in income in the Standard Model.

STEP Click the Reset button and change ω_2 (in K14) from 10 to 2.

This generates a downward shift in the budget constraint. So price changes cause rotations (or pivots or swivels) and endowment shocks produce shifts.

The budget constraint in an Endowment Model plays the same role as the budget constraint in the Standard Model. It describes the agent's consumption possibilities. Unlike the Standard Model, however, where price changes caused rotation around the x or y intercept, price shocks in the Endowment Model lead to swiveling around the initial endowment. It makes sense that the initial endowment is going to remain the same as prices change because the agent is neither buying nor selling at the initial endowment so the price does not matter at that point.

To get shifts in the budget constraint, we will have to change either ω_1 or ω_2. This changes the initial endowment point and allows the agent to buy and sell from the new endowment point, creating a new budget line.

Now that you understand the budget constraint, we are ready to solve the agent's constrained utility maximization problem with the Endowment Model.

The Initial Solution

The utility side of the Endowment Model is the same as the Standard Model. The agent's preferences are shown by indifference curves that are represented mathematically by a utility function.

The agent seeks to maximize utility given the budget constraint. As usual, we can solve this problem numerically and analytically.

STEP Proceed to the *OptimalChoice* sheet. Figure 5.3 shows what this sheet looks like when you first open it.

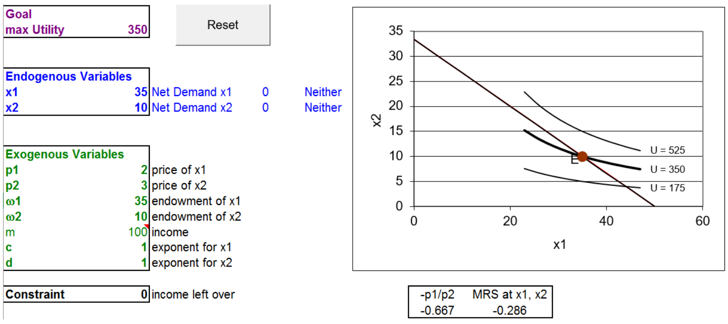


Figure 5.3: The initial view of the optimization problem.
Source: EndowmentIntro.xls!OptimalChoice

Notice how the organization is the same as the Standard Model. There is a goal, endogenous variables (in blue) and exogenous variables (in green). The agent seeks to maximize utility, represented by a Cobb-Douglas functional form, by choosing the amounts of x_1 and x_2 to consume, subject to the budget constraint.

The graph is also similar, with the addition of point E, representing the initial endowment. There are three representative indifference curves (there are an infinity of such curves, one through every point in the quadrant).

Although much is familiar, Figure 5.3 and your computer screen do have some notable innovations. Cells B18 and B19 have been added to the list of exogenous variables. They represent the given initial endowment. Cell B20 has a formula that computes m, which is not bolded to indicate that it is derived from other exogenous variables.

In addition, cells C11:E12 are new. Let's find out what they tell us.

STEP Click on D11 to see its formula, = x1_ - w1_.

The underscore (_) is used to distinguish the names, x1 and w1, from the cell addresses, X1 and W1. Lowercase w is the closest English character to ω.

More substantively, the formula computes *net demand*, how much the consumer wants to buy or sell. It takes *gross demand*, the optimal amount of the good the agent wishes to end up with, that is, the values of x_1 and x_2 and subtracts the initial endowment amounts. There is a gross and net demand for each good.

On opening, the net demand for x_1 is zero because B11 is set at 35, which is equal to the agent's initial endowment of good 1. Suppose the agent decided to buy three units of good 1.

STEP Change B11 to 38.

Net demand for good 1 is now plus three. That makes sense because the consumer started with 35 units of good 1, but wants to have 38, so three more must be purchased.

Of course, the combination 38,10 is unattainable. The consumer must sell some x_2 in order to be able to buy three units of x_1. How much needs to be sold? Two units.

STEP Change B12 to 8.

The agent is back on the budget line and net demand for good 2 is negative. Cell E12 reports that the agent is a seller of good 2. Clicking on cell E12 reveals an IF formula that displays Buyer or Seller depending on whether net demand is positive or negative.

Compare the MRS on your screen to the MRS at the initial position from Figure 5.3. Was buying three units of good 1 with the proceeds from the sale of two units of good 2 a smart move?

No. The MRS at 38,8 is farther away from the price ratio than the MRS at 35,10. The graph reveals that we moved to a lower indifference curve when we moved to 38,8.

We need to head the other way. The agent needs to travel up the budget line, to the northwest, selling good 1 and buying good 2. How much should be sold and bought?

STEP Run Solver to find the initial solution.

Utility is maximized when gross demands are 25 and $16\frac{2}{3}$ of goods 1 and 2, respectively. Net demands are -10 and $6\frac{2}{3}$. This means the agent sells 10 units of good 1 and uses the \$20 in revenue to buy $6\frac{2}{3}$ units of good 2.

This is the same solution as in the Standard Model with $m = \$100$. That makes sense, since the value of the initial endowment is \$100.

We can confirm this result with analytical methods. We follow the recipe for the Lagrangean method of solving constrained optimization problems.

We will work on a general form of this problem, leaving all exogenous variables as letters to get a reduced form expression that we can evaluate for any combination of exogenous values. We rewrite the constraint to that it is equal to zero and form the Lagrangean.

$$\max_{x_1, x_2, \lambda} L = x_1^c x_2^d + \lambda(p_1\omega_1 + p_2\omega_2 - p_1x_1 - p_2x_2)$$

The third step is to take derivatives with respect to each choice variable and in the final step we set the three derivatives equal to zero to get the first-order conditions, which we need to solve for x_1*, x_2*, and $\lambda*$.

$$\frac{\partial L}{\partial x_1} = cx_1^{c-1}x_2^d - p_1\lambda = 0$$
$$\frac{\partial L}{\partial x_2} = dx_1^c x_2^{d-1} - p_2\lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = p_1\omega_1 + p_2\omega_2 - p_1x_1 - p_2x_2 = 0$$

Our solution strategy involves moving the lambda terms to the right-hand side and dividing the first equation by the second to cancel lambda (and giving the familiar MRS $= \frac{p_1}{p_2}$ condition). This equation can then be solved for optimal x_2 as a function of optimal x_1.

$$\frac{cx_2^*}{dx_1^*} = \frac{p_1}{p_2}$$
$$x_2^* = \frac{d}{c}\frac{p_1}{p_2}x_1^*$$

Although it looks like it, this is not the answer for x_2 because it has x_1 in it. The reduced form solution must be a function of exogenous variables alone. Substitute this expression for x_2 into the third first-order condition (the budget constraint) and solve for optimal x_1.

$$p_1\omega_1 + p_2\omega_2 - p_1x_1^* - p_2\left[\frac{d}{c}\frac{p_1}{p_2}x_1^*\right] = 0$$
$$\left(1 + \frac{d}{c}\right)p_1x_1^* = p_1\omega_1 + p_2\omega_2$$
$$x_1^* = \left(\frac{c}{c+d}\right)\frac{p_1\omega_1 + p_2\omega_2}{p_1}$$

This expression can be evaluated for any combination of exogenous variable values. For example, if we use the parameter values in the *OptimalChoice* sheet, we can compute that optimal $x_1 = 25$. This agrees perfectly with the numerical approach.

Furthermore, this expression shows the quantity demanded at a given p_1, ceteris paribus, so it can be used to display a demand curve for x_1. There is, of course, a similar expression for good 2.

In the Standard Model, the reduced form solution was $x_1{}^* = (\frac{c}{c+d})\frac{m}{p_1}$. The Endowment Model's solution is the same, except instead of m in the numerator, we have $p_1\omega_1 + p_2\omega_2$. This makes sense since the value of the initial endowment is $p_1\omega_1 + p_2\omega_2$.

With an Endowment Model, we can subtract the initial amount of good 1 to obtain a net demand curve.

$$nd_1 = x_1{}^* - \omega_1 = \left(\frac{c}{c+d}\right)\frac{p_1\omega_1 + p_2\omega_2}{p_1} - \omega_1$$

Comparative Statics with the Endowment Model

We can do comparative statics analyses analytically or numerically. The reduced form expression can be used to explore the rate of change of optimal x_1 with respect to any exogenous variable. For example, we can take the derivative with respect to p_1.

This is more complicated than usual because p_1 appears in two places. We could use the Product Rule, but it is easier to do some reorganizing and simplify things before we take the derivative.

First, we move p_1 from the denominator. This will enable us to use our usual derivative rule.

$$x_1* = (\frac{c}{c+d})\frac{p_1\omega_1 + p_2\omega_2}{p_1} = (\frac{c}{c+d})(p_1\omega_1 + p_2\omega_2)p_1^{-1}$$

But we can also multiply p_1 through to cancel the p_1 in the $p_1\omega_1$ term.

$$x_1* = (\frac{c}{c+d})(p_1\omega_1 + p_2\omega_2)p_1^{-1} = (\frac{c}{c+d})(\omega_1 + p_1^{-1}p_2\omega_2)$$

Then we can expand to leave p_1 isolated in a single term so that the derivative with respect to p_1 is straightforward.

$$x_1* = (\frac{c}{c+d})(\omega_1 + p_1^{-1}p_2\omega_2) = (\frac{c}{c+d})\omega_1 + (\frac{c}{c+d})p_1^{-1}p_2\omega_2$$

Now, when we take the derivative with respect to p_1, we apply our usual derivative rule and bring the exponent down and subtract one from the second term. The first term has a derivative with respect to p_1 of zero since it does not contain p_1.

$$\frac{dx_1*}{dp_1} = (-1)(\frac{c}{c+d})p_1^{-2}p_2\omega_2$$

We can evaluate this expression at the initial values of the exogenous variables to get an instantaneous rate of change in optimal x_1 as p_1 changes. Plugging in $c = d = 1, p_1 = 2, p_2 = 3$, and $\omega_2 = 10$ gives -3.75. This means that an infinitesimally small increase in p_1 would decrease x_1 by 3.75-fold.

But what does that number tell us? Is it a lot in the sense of a big response to a price shock? The slope provides no answer to this question. We need percentage changes—elasticity—to answer this question.

We can multiply the slope by the initial ratio of $\frac{p_1}{x_1*}$ to compute the p_1 elasticity of x_1*.

$$\frac{dx_1*}{dp_1}\frac{p_1}{x_1*} = ((-1)(\frac{c}{c+d})p_1^{-2}p_2\omega_2)(\frac{p_1}{x_1*})$$

We evaluate this expression at $p_1 = 2$ (and the initial values of the other exogenous variables).

$$\frac{dx_1*}{dp_1}\frac{p_1}{x_1*} = ((-1)(\frac{c}{c+d})p_1^{-2}p_2\omega_2)(\frac{p_1}{x_1*}) = -3.75(\frac{2}{25}) \approx -0.3$$

The elasticity does tell us that the quantity demanded of x_1 is quite price insensitive at the initial solution. An elasticity less than one (in absolute value) is said to be inelastic and the closer to zero, the lower the responsiveness.

Unlike the Standard Model, where a Cobb-Douglas utility function gives a unit price elasticity, we get a non-unitary elasticity here because a change in p_1 appears in the denominator and numerator in the reduced form. In the numerator, the change in price is affecting the value of the agent's endowment whereas in the Standard Model, income is fixed.

We can also use numerical methods to explore the comparative statics properties of an own price change.

STEP Use the Comparative Statics Wizard to *decrease* p_1 by 0.1 (10 cents) for 15 shocks (from 2 to 0.5). Be sure to keep track of net demands and the buyer/seller position in the endogenous variables by using the *ctrl* key to select non-contiguous cells, as depicted in Figure 5.4. You want to track cells B11:B12 and D11:E12.

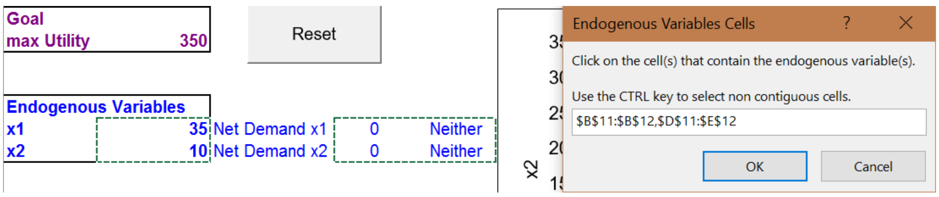


Figure 5.4: Selecting endogenous variables with CSWiz.

The *CSP1* sheet shows what your results should look like. There are several notable outcomes.

When the price fell from 90 cents to 80 cents, the agent switched from selling x_1 and buying x_2 to buying x_1 and selling x_2. The price of x_1 got so low that even though the agent starts with a lot of x_1 (compared to x_2), it is better to buy more x_1. The budget line gets flatter as p_1 falls, making buying x_1 a better choice than selling it.

Notice the behavior of maximum utility (column B) as price falls. The agent was a seller at first so falling prices hurt. Below 90 cents, however, the agent is a buyer of x_1 and falling p_1 increases utility.

The *CSP1* sheet also shows slope and elasticity computations. From p_1 \$2/unit to \$1.90, the slope (yellow background) and elasticity (orange background) measures are close, but different than at $p_1 = 2$ (using the derivative). This is due to the fact that optimal x_1 is non-linear in p_1. In other words, $x_1{}^* = f(p_1)$ is not a line, but a curve (as clearly shown in the chart below the data).

The Endowment Model Extends the Standard Model

The Endowment Model is the Standard Model of the Theory of Consumer Behavior with an initial endowment of goods instead of cash income. This transforms the consumer into the dual-role of seller and buyer of goods. The driving force in the agent's decision making remains utility maximization. Many of the ideas behind the Standard Model (such as equating the MRS and price ratio) carry over to the Endowment Model. Of course, the framework for presenting and understanding the model, comparative statics analysis, remains the same.

It may seem that replacing income with an initial endowment is a minor twist, but we will see that the Endowment Model enables analysis of a wide range of choice problems.

Exercises

1. Perform a comparative statics analysis of c, the exponent on x_1, using the Comparative Statics Wizard. Use increments in c of 0.1. State the effect of changing c on $x_1{}^*$. Describe your procedure and take screen shots of your results as needed.

2. Use your comparative statics results to find the c elasticity of $x_1{}^*$ from 1 to 1.1. Show your work.

3. Use the reduced form expression in this chapter to find the c elasticity of $x_1{}^*$. Show your work.

4. Compare your answers from questions 2 and 3. Explain why they are the same or differ.

References

The epigraph is from page 188 of David M. Kreps *A Course in Microeconomic Theory* (1990). If you are interested in graduate study in economics, this book is worth browsing. In the preface, Kreps says (p. xv), "The primary target for this book is a first-year graduate student who is looking for an introduction to microeconomic theory that goes beyond the traditional models of the consumer, the firm, and the market." Kreps allows that it could be used for undergraduate majors taking an "advanced theory" course or "mathematically sophisticated students," but he warns that, "The book presumes, however, that the reader has survived the standard intermediate microeconomics course."

The Endowment Model is taking us close to the next level of microeconomic theory. Google "graduate micro theory" for more advanced micro books.

To learn more about Masters and PhD programs in economics, search for "graduate economics rankings" and be sure to visit the American Economics Association's website at www.aeaweb.org/resources/students/grad-prep.

The term impatience carries
with it the presumption that
present goods are preferred.
But I shall treat the two terms
(impatience and time
preference) as synonymous.

Irving Fisher

5.2 Intertemporal Consumer Choice

Suppose the government wants to stimulate saving by workers so they won't be poor when they retire. Individual Retirement Accounts (IRAs) and 401(k) (their section in the tax code) plans enable savings to grow tax free, so the interest rate earned is higher than if returns were taxed. A higher interest rate should stimulate more saving. But how much more?

Typically, estimates of the interest rate elasticity of savings are positive, but quite small, say 0.15. If someone had this elasticity, would attempts to stimulate saving by increasing the interest rate be effective?

No, because the low interest rate elasticity of savings means that saving is not responsive to changes in the interest rate. Suppose the interest rate doubles so we have a huge 100% change. Because the elasticity is 0.15, that means we will see only a 15% increase in savings. A more realistic 10% increase in the interest rate would generate a small 1.5% increase in savings. The small elasticity tells us that shocks to the interest rate are not going to move the amount saved by very much.

This is an example of interpreting an elasticity. Computing an elasticity is important (and you will continue to see examples of how to do it), but understanding what an elasticity is telling us is even more critical.

Now that we know the elasticity is low and what that means, this leads to a second question: What would make the interest rate elasticity of savings be so small? The rest of this chapter offers an application of the Endowment Model to answer this question. In addition, income and substitution effects play a major role in the explanation. There is no doubt about it, learning economics is a cumulative undertaking—the same ideas keep popping up again and again.

The Intertemporal Choice Model

Intertemporal choice means the agent faces a decision that spans across time periods. Saving over the years working means less consumption, but that allows for more consumption when retired. We model the agent as deciding what to consume every year over their lifespan.

Just as when we modeled the consumer buying just x_1 and x_2 instead of many goods and services, we make a simplifying assumption that collapses many time periods into two: present and future. In the present, right now, the agent works and in the future, one year later, she does not (she retires).

In addition, there is another implied simplifying assumption: the agent knows with certainty how long she will live. She is born and works as one-year old, is retired as a two-year old and dies on the last day of her second year. She decides, as soon as she is born, how much she will consume in year 1 (the present) and year 2 (the future).

Instead of having two goods x_1 and x_2, we have consumption of a single good in the present, c_1, and the future, c_2. The price of the single good is \$1/unit so if you have, say, \$40, you can buy 40 units. There is no inflation so the price is the same in both time periods.

Notice the usual modeling technique at work here—realistic details are simply assumed away. Most people's lives unfold as follows: Childhood becomes teen-aged years, and then a long period of working adult life eventually turns to retirement years and death. The Intertemporal Choice Model collapses all of that into two time periods. It also assumes away complications from not knowing exactly when we die.

Faced with criticisms about the unrealistic nature of the model, economists respond by saying that we are not interested in realism. We reduce the complex real world to a model that can be analyzed with comparative statics to produce testable predictions. For economists, the goal is not to describe reality, but to predict via comparative statics. We strip away all complications to create an unreal, incredibly simple model that contains the kernel of the problem so we can work out how the agent responds to shocks.

Modeling is not easy. There is science (and math) and art involved. Users and consumers of these models need sharp critical thinking skills—sometimes important elements are assumed away.

We continue building the model by defining the initial endowment as the amount of present and future income you start with. The initial endowment in the first year is m_1 and in the second year m_2. The first year's initial endowment is income from working and the second year's initial endowment is income from sources like Social Security. Thus, it makes sense that $m_1 > m_2$, which says that income is higher during the working than the retired year. Since the price is $1/unit, the initial endowment incomes are also initial endowment consumption in the two periods.

We are ready to work on the optimization problem itself. We follow the usual approach, modeling the budget constraint, then satisfaction, then putting the two together to find the initial solution. Of course, after finding the initial optimum we will do comparative statics analysis, where we will answer the question: What causes the interest rate elasticity of savings to be so small?

The Budget Constraint

STEP Open the Excel workbook *IntertemporalChoice.xls* and read the *Intro* sheet, then go to the *MovingAround* sheet.

The consumer begins at the initial endowment point, 80,20, where 80 represents her income and consumption in time period 1 (remember that the price of the good is $1/unit). Income and consumption of 20 in time period 2 is lower (given that she is not working). These numbers are arbitrary and do not have any special meaning.

A critical concept for the Endowment Model is that the agent does not have to stay at the initial position. In this application, she can move by saving or borrowing. Saving means you consume less in the present and carry over the unconsumed portion into the future. Saving is like selling present consumption and buying future consumption.

Suppose she saves 30 units of consumption in year 1 by saving $30. What would be her position in the second year?

STEP Change cell B19 to 50. This implements the plan to increase future consumption, but look at cells B21 and B22. Instead of simply reallocating from 80,20 to 50,50, by saving 30 units, she got an extra 6 units in interest on her savings.

If you save \$30 for one year at 20%, you end up with \$56. The \$30 you saved (called the *principal*) and *interest* earned of \$30 x 20% = \$6 makes your savings worth \$36 in the future and we add this to the \$20 of initial future income to get the grand total of \$56.

There is an equation that gives us the value of c_2 for any chosen value of c_1.

$$c_2 = m_2 + (m_1 - c_1) + r(m_1 - c_1)$$

The equation says that the amount of consumption in time period 2 equals the initial endowment amount in time period 2, m_2, plus the principal saved, $m_1 - c_1$, plus the interest earned on the amount saved, $r(m_1 - c_1)$. We can rewrite this in a simpler form by collecting the savings term.

$$c_2 = m_2 + (1 + r)(m_1 - c_1)$$

This is the equation of the budget constraint in this model. It shows that the intercept is $m_2 + (1 + r)m_1$ and the slope is $-(1 + r)$ (just multiply through by $(1 + r)$). The slope tells us that saving \$1 will yield $1 + r$ dollars in time period 2.

What would be the maximum consumption possible in time period 2? We have two ways to answer this question.

STEP Change cell B19 to 0. She consumes nothing now and ends up with 116 units in the future.

"But she will starve if she consumes nothing in period 1." That would be another constraint that is not being modeled. We are not saying she will consume nothing in the present time period, we are merely exploring the consumption possibilities.

Saving everything (the same as consuming nothing in the present) can also be found by computing the value of the y intercept. We can evaluate $m_2+(1+r)m_1$ at $m_1 = 80, m_2 = 20$, and $r = 20\%$, yielding $20+(1+0.2)80 = 116$. This is the same answer that we got with Excel.

The y intercept tells us the *future value* of the agent's initial endowment, measuring income in both periods in terms of time period 2.

Instead of saving, the agent can borrow. Suppose the agent decided to consume more than 80 units in time period 1. How could she do this? Easy: use her time period 2 income to borrow from it. As before, however, we have to be careful. The interest rate plays a role.

STEP Change cell B19 to 90. She borrows \$10 from her future income.

Does she end up with 90,10—subtracting 10 from c_2 and adding it to c_1? No way. As Excel shows, she has to pay interest on the borrowed funds. If she borrows \$10, she ends up with only \$8 in the future because she has to pay back the principal (\$10) and the interest (\$2).

What is the most she could consume in time period 1?

STEP Change cell B19 to 100. What happens?

She cannot do this. She cannot choose negative x_2. She does not have enough future income to enable 100 units of time period 1 consumption.

STEP Continue entering numbers in cell B19 until you drive c_2 (in cells B23 and B24) to zero.

The x intercept is $96\frac{2}{3}$. It is the *present value* of her endowment, measuring income in both periods from the standpoint of time period 1.

STEP Proceed to the *Properties* sheet.

Our work in the *MovingAround* sheet makes it easy to understand the budget line displayed in the *Properties* sheet. Clearly, given an initial endowment, movement up the budget line is saving and down is borrowing.

These are just consumption possibilities. We do not know what this person will do until we incorporate her preferences. We do know she can be anywhere on the constraint (including the initial endowment point). It all depends on her indifference map and where the highest attainable indifference curves lie.

STEP Proceed to the *Changes* sheet. Change the interest rate, cell L8, to 50%. Your screen will look like Figure 5.5.

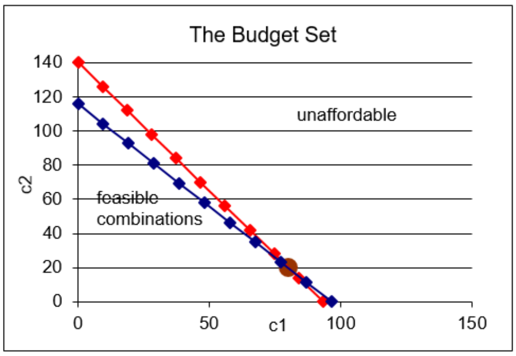


Figure 5.5: Increasing r.
Source: IntertemporalChoice.xls!Changes

Our work with the Endowment Model in the previous section enables us to easily interpret the result. As before, the budget constraint swivels around the initial endowment point.

Above the initial endowment point, the increase in r is a good thing, increasing consumption possibilities. If the agent is a saver, the shock is welcome.

Borrowers, however, would not be happy with an increase in r. This is a price increase to present consumption and reduces consumption possibilities for borrowers.

STEP Click the Reset button. Change m_1 and m_2 to see how these shocks are like an income shock. It maintains the slope, but shifts the budget constraint.

Now that we understand how the budget constraint works, we are ready to turn to the agent's goal, maximizing utility.

Preferences

The agent has preferences over present and future consumption that can be captured by the indifference map.

We use the usual Cobb-Douglas function form to express preferences as a utility function.

STEP Proceed to the *Preferences* sheet. Compare the utility functions with $d = 0.5$ and $d = 0.1$. The utility function allows us to model different preferences.

Figure 5.6 shows two different agents with different rates of time preference for future consumption. The person on the right exhibits a strong preference for present consumption, while the person on the left is more willing to wait.

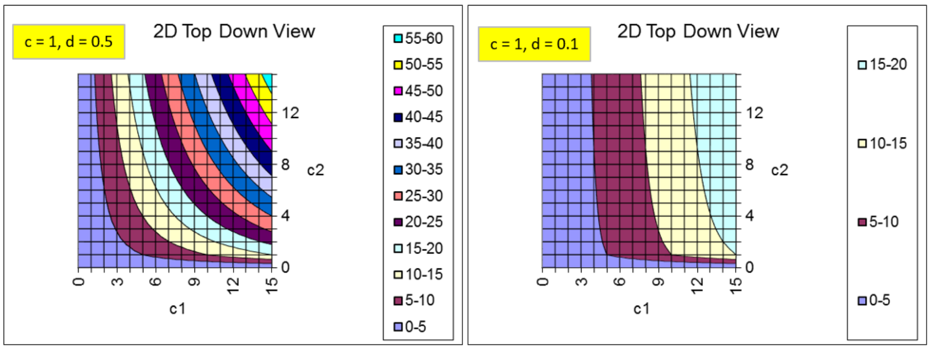


Figure 5.6: Modeling rates of time preference.
Source: IntertemporalChoice.xls!Preferences

A more immediate gratification personality is represented on the right side of Figure 5.6. We would say this person is more impatient—he likes present much more than future consumption. The exponent d is much smaller than c, which means inputs into the utility function through c_2 provide much less utility than via c_1.

The steep indifference curves reveal that he is willing to trade a great deal of future consumption for a just a little more present consumption. His MRS at a given point (for example, 6,6) is higher (in absolute value) than the MRS of the person on the left.

We do not say the person on the right has "bad preferences" (although the language used in this example, such as *impatience* does seem to connote disapproval). Economists take preferences as given. We are not supposed to judge them as right or wrong. A person with preferences that substantially ignore the future is treated the same as someone who does not like broccoli or likes the color blue.

There is a complication here, however, in that a person's rate of time preference almost certainly changes over time. A young person may not save much because she does not value the future, but she may regret her decision when she gets older. Deciding whose preferences should rule, young or old you, is a difficult philosophical problem.

With the budget line and preferences, we can now solve the constrained utility maximization problem.

Finding the Initial Solution

STEP Proceed to the *OptimalChoice* sheet. Figure 5.7 shows the initial display. The current bundle is 80,20—the initial endowment point. The agent is not maximizing satisfaction subject to the budget constraint. The indifference curve is clearly cutting the budget line and, therefore, the agent should move northwest up the budget line to maximize utility.

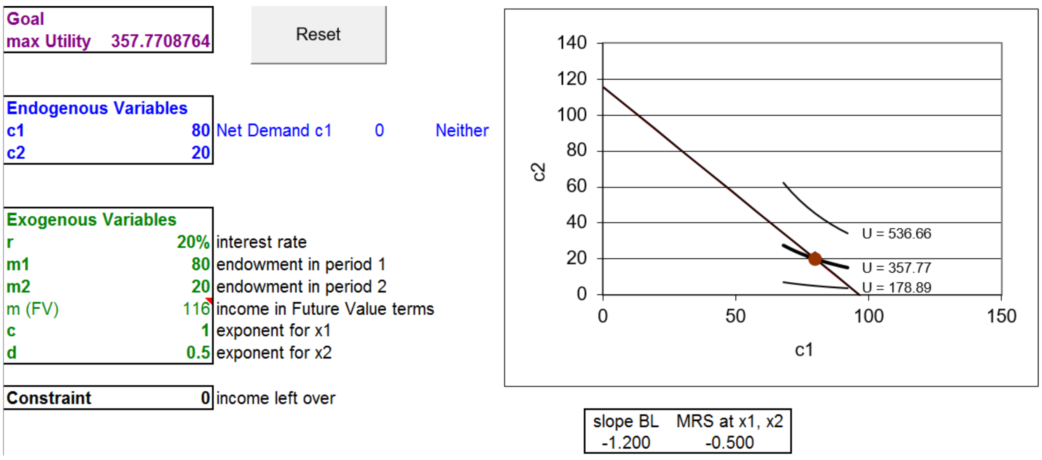


Figure 5.7: An inefficient position.
Source: IntertemporalChoice.xls!OptimalChoice

STEP Run Solver to find the initial solution.

The agent opts for the point $64\frac{4}{9}, 38\frac{2}{3}$. This means she has decided to save $15\frac{5}{9}$ of her present consumption. She chooses this present and future combination, implying this level of saving, because this maximizes utility subject to the budget constraint.

Notice that the negative net demand is interpreted as saving. It is computed as optimal c_1 minus the initial endowment of present consumption. As mentioned earlier, saving is like selling present consumption to buy greater future consumption. We often drop the minus sign so we do not get confused by increases and decreases in saving.

Comparative Statics

We focus on r. We want to know how savings will respond when r changes. Remember our question: Why is the interest rate elasticity of savings so low?

Before we begin our comparative statics analysis, we need to be clear about the language used. Since the shock variable, r, is measured as a percent, things can get confusing once we start working on responses and elasticities. We need to keep clear the difference between a *percentage point change* and *percent change*. They sound the same, but the former is a difference (Δ), $new - inital$, and the latter is a percent computation, $\frac{new-initial}{initial}$.

So, if r increases from 20% to 30%, that is a 10 percentage point change since we compute 30 - 20, but a 50 percent change: $\frac{30-20}{20}$. The same language would be used if we were working with unemployment rates. An increase from 5% to 6% is a one percentage point increase and a 20% increase.

The finance literature uses basis points for differences in variables measured in percents. There are 100 basis points in one percentage point. If a bond yield rises from 3.25% to 3.35%, that is an increase of 10 basis points.

STEP Run the Comparative Statics Wizard, changing the interest rate by 10 percentage points (0.1) increments. Keep track of c_1, c_2, net demand, and whether the person is a saver or borrower (cells D11 and E11).

Your results should be similar to those in the *CSr* sheet.

STEP Use your CSWiz results to compute the interest rate elasticity of savings from r = 20% to 30%.

We find that the interest rate elasticity of savings from r = 20% to 30% is about 0.11. (Check the formula in cell I15 in the *CSr* sheet if needed.) That is quite low. A 50 percent increase in r only increased savings by a little over 5 percent.

This elasticity is similar to the 0.15 elasticity at the beginning of this chapter. Why is this happening? Why is saving so unresponsive to changes in the interest rate?

The answer lies in the income and substitution effects. For savings, the income and substitution effects from a change in r work in opposite directions (when c_1 is a normal good). Thus, they tend to cancel each other out and the total effect ends up being small.

To head off serious misunderstanding, you need to know right now that this does not mean that we are dealing with a Giffen good. We will see that we are dealing with cross effects when r rises for a saver and Giffen goods are defined in terms of own effects. Also, c_1 and c_2 are both normal goods in a Cobb-Douglas utility function so we know we can't get Giffenness.

STEP To see how the income and substitution effects apply to this problem, return to the *OptimalChoice* sheet. Suppose r increases to 300%. Change B16 to this absurdly high interest rate.

This huge change enables us to see clearly what is happening on the graph. The budget line swivels in a clockwise direction, getting much steeper. Remember that the slope is $-(1+r)$ so an increase in r makes the line steeper. This is good for savers and bad for borrowers.

STEP After changing cell B16 to 300%, run Solver to find the new initial solution.

Solver gives the new optimal solution, $c_1{}^* = 56\frac{2}{3}$ and $c_2{}^* = 113\frac{1}{3}$, when $r = 300\%$. Optimal savings has increased from \$15.56 to \$23.33, so that is good news, but this is a pretty weak response to the massive increase in the interest rate from 20% to 300%.

Figure 5.8 shows the initial solution (point A) and the new optimal solution (point C). It also includes a dashed line that is parallel to point C's budget line, but goes through point A. This, of course, is the line that is used to separate the total effect into income and substitution effects using point B.

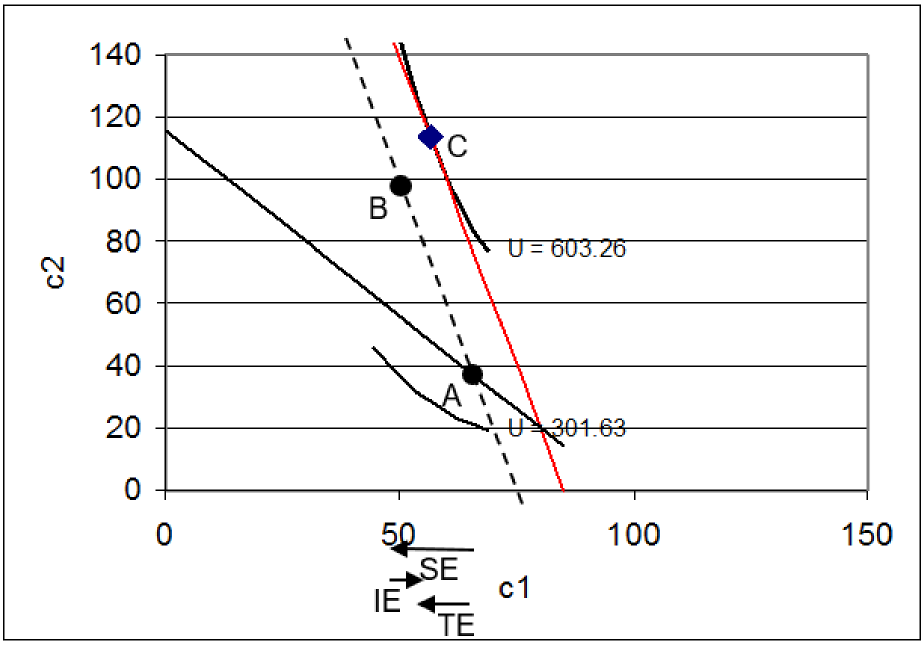


Figure 5.8: Income and substitution effects.
Source: IntertemporalChoice.xls!OptimalChoice: cell F52

How much income (m_1) did we have to take away (hypothetically, of course) to cancel out the income effect of the higher interest rate? We can use Excel to answer this question.

STEP With $r = 300\%$, enter the initial solution (point A). To minimize rounding error, use a formula with fractions. So, enter "= 64 + 4/9" in B11 and "= 38 + 2/3" in B12. Now, start decreasing m_1 (in cell B17). Your goal is to find that value of m_1 so that the initial solution is on the budget line—i.e., the constraint cell is zero.

A little experimentation should convince you that $m_1 = 69\frac{1}{9}$ is the value that puts the dashed budget line through the initial solution.

If you want to be daring, you could use Solver. Call Solver, then click the Reset All button. The objective is the constraint cell (B23) and you want to make the value of it zero by changing m_1 (B17). Solver gives the same answer as above.

Or, you could use the budget constraint to find the m_1 needed to buy the original optimal bundle with $r = 300\%$. Simply plug in the initial optimal solution along with the new value of r (and initial m_2) and solve for m_1. You are finding the value of m_1 that would enable you to buy the initial optimal combination with the higher interest rate. The analytical answer agrees with the numerical approach.

STEP Now, with r = 300% and $m_1 = 69\frac{1}{9}$, run Solver to find point B.

Be careful with the interpretation of savings for point B. Remember that income is not really $m_1 = 69\frac{1}{9}$, but 80. This means that at point B, the agent would save $30.59, not $19.07 as displayed in cell D11.

Figure 5.9 shows the results in a table. You can see Figures 5.8 and 5.9 side by side by scrolling down to row 50 or so in the *OptimalChoice* sheet. Look at how the substitution effect leads to a large increase in savings, but the income effect cancels out part of this increase.

Point	Description	c1*	c2*	Savings*
A	Initial solution	64.44	38.67	15.56
B	Imaginary	49.41	98.81	30.59
C	New solution	56.67	113.33	23.33

Effect	Movement	Amount
SE	A to B	15.03
IE	B to C	-7.26
TE	A to C	7.77

Figure 5.9: Total, income, and substitution effects.
Source: IntertemporalChoice.xls!OptimalChoice: cell M51

The income and substitution effects provide an explanation for the low interest rate elasticity of savings. What is happening is that the two effects are working against each other when r rises and the agent is a saver.

Does this mean c_1 is an inferior good? No. The reason why the effects are opposing each other is because, for savers, an increase in the interest rate is like a decrease in the price of future consumption so the effects on c_1 and savings are actually cross effects. Look carefully at Figure 5.8. In the region of the graph with points A, B, and C, it is as if we decreased p_2,and rotated the budget line up clockwise (with a steeper slope).

Saving and Borrowing Explained

The Intertemporal Choice Model is an application of the Endowment Model in the Theory of Consumer Behavior. The model says that the agent chooses the amount to consume in time periods 1 and 2 in order to maximize satisfaction given a budget constraint.

The model explains saving (or borrowing) as an optimizing move on the part of an agent who is trading off present and future consumption.

The model can also explain why the interest rate elasticity of savings is often estimated as a positive, but small number, which means that saving is quite unresponsive to the interest rate. The explanation rests on the fact that the income effect opposes the substitution effect for c_1 and savings (for those with negative net demand for c_1).

Exercises

1. Solve the problem in the *OptimalChoice* sheet using analytical methods. In other words, find the reduced form expressions for optimal c_1, c_2, and saving from

$$\max_{c_1, c_2} u(c_1, c_2) = c_1^c c_2^d$$
$$\text{s.t. } c_2 = m_2 + (1 + \text{r})(m_1 - c_1)$$

   Show your work.

2. Use the parameter values in the *OptimalChoice* sheet (with $r = 20\%$) to evaluate your answers for question 1. Provide numerical answers for the optimal combination of consumption in time periods 1 and 2 and for optimal saving.

3. Do your answers from question 2 agree with Excel's Solver results? Is this surprising? Explain.

4. Use your reduced form solution from question 1 to compute the interest rate elasticity of savings at r = 20%.

5. In working through this chapter, you found the interest rate elasticity of savings from r = 20% to 30%. Why is the elasticity computed at a point (in question 4 above) different from this elasticity?

References

The epigraph is on page 66 of Irving Fisher, *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It* (first edition, 1930; reprinted 1977 by Porcupine Press).

Joseph Schumpeter had high praise for Fisher: "[S]ome future historian may well consider Fisher as the greatest of America's scientific economists up to our own day" (*History of Economic Analysis*, 1954, p. 872). Schumpeter chose to ignore Fisher's "propagandist activities (temperance, eugenics, hygiene, and others)," but he did point out that Fisher's reputation as an economist was negatively affected: "Fisher, a reformer of the highest and purest type, never counted costs—even those most intensive pain costs which consist in being looked upon as something of a crank—and his fame as a scientist suffered correspondingly" (*History of Economic Analysis*, 1954, p. 873).

For a recent biography of Fisher, who seems to be enjoying a rehabilitation of sorts, see Robert W. Dimand (2019), *Irving Fisher*.

The empirical evidence on the interest rate elasticity of savings is mixed (which is actually evidence that it is not large). For a dated, but perhaps comprehensible example, see Irwin Friend and Joel Hasbrouck, "Saving and After-Tax Rates of Return," *The Review of Economics and Statistics*, Vol. 65, No. 4. (November, 1983), pp. 537–543, www.jstor.org/stable/1935921.

The literature on the effect of Individual Retirement Accounts and other plans (such as 401(k)) on saving is truly vast. A Google Scholar search on "individual retirement accounts saving" produces hundreds of thousands of hits. This topic would make an excellent paper or undergraduate senior thesis.

> The Prophet said: "Charity is a necessity for every Muslim." He was asked: "What if a person has nothing?" The Prophet replied: "He should work with his own hands for his benefit and then give something out of such earnings in charity."
>
> Prophet Muhammed

5.3 An Economic Analysis of Charity

The phrase "an economic analysis of" is code for "using the framework of optimization and comparative statics to study observed behavior." In this case, we use the Endowment Model from the Theory of Consumer Behavior to study charitable giving.

How can economics have anything to say about giving away money? Isn't charity something really nice people do, not the selfish, rational maximizers that inhabit economics? Doesn't this mean that thinking like an economist is useless for studying charity?

These questions are based on a common misunderstanding that economics applies only to a subset of the world. So, the mistaken thinking goes, you can use economics to study certain things like banking or unemployment, but not war or marriage. This is wrong because modern economics is not defined by content, but by method. Anything involving choice, like going to war or getting married or brushing your teeth or joining a church can be analyzed with the tools of economics.

We will see that the economic approach offers a different view of charitable giving. By casting the problem as a choice—how much to give is the key endogenous variable—we can apply the optimizing and comparative statics framework of economics. We do not claim this is the only or even the best perspective, but it does provide another way to understand charity.

Basic Facts about Giving

Each year, people all around the world give away a lot of money, goods, and time (as volunteers). Humans are sympathetic when people close to them are in distress. All religions encourage charity and caring for people less fortunate.

Giving USA provides data on philanthropy in the United States. Figure 5.10, from the *2018 Annual Report*, shows the breakdown of the $410 billion that were contributed to charities in 2017. To help understand what this number means, we can compare total contributions to the size of the economy and we find a giving rate of about 2.1% of GDP.

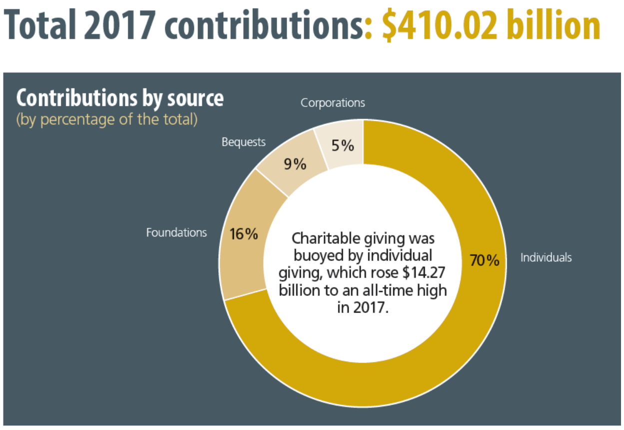


Figure 5.10: Charitable giving by source of contribution.
Source: Giving USA 2018 Annual Report

The *2018 Annual Report* contextualizes total giving by tracking giving over time, shown in Figure 5.11. Total giving jumped in the mid 1990s and reached its highest level in 2017. That is good news.

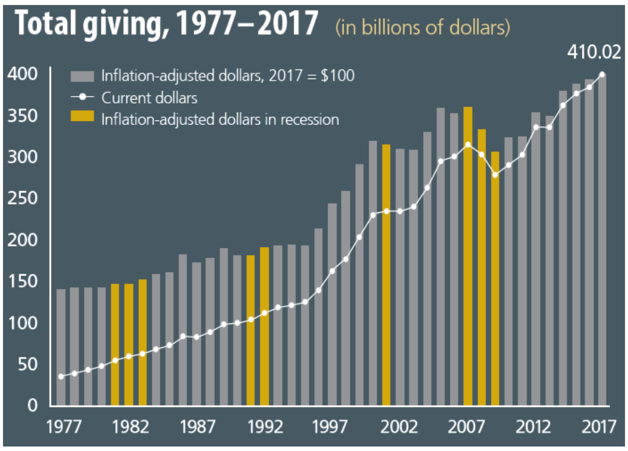


Figure 5.11: Charitable giving over time.
Source: Giving USA 2018 Annual Report

The Internal Revenue Service is another source of data on charitable giving because taxpayers claim deductions when they give to charity to lower the tax owed. The IRS also collects data on non-profit organizations which do not pay tax, but they have to file Form 990. IRS data can be found at www.irs.gov/statistics.

Charitable giving data shows that it not only varies over time, there is also tremendous individual variation. Many people give nothing, others give a little, and a few people donate a lot. Religions encourage members to tithe, giving 10% of their income. Upon death, some people give substantial fractions of their estates to charity, while others hand it all to their heirs.

There are many questions we can ask about charitable giving, but our top three are:

1. Why do people give to charity?
2. What determines how much they give?
3. How can charitable giving be stimulated?

Because this is an economic analysis of charity, we are going to answer these questions by using the method of economics. We will set up and solve an optimization problem. This will provide the economic explanation for why people give and what determines how much they give. We will see that charitable giving can be stimulated by changing exogenous variables, ceteris paribus.

Our model will do the usual stripping away of realistic details, making incredible simplifying assumptions, to enable us to solve the model and play comparative statics games. Keep your eye on the procedure as we set up, solve, and compute our key measure—the tax break elasticity of giving.

An Endowment Model of Giving

As usual, we begin with the budget constraint, then we model preferences, and we use both to find the initial solution to the problem of maximizing satisfaction subject to the budget constraint.

The optimization problem is entirely from the donor's point of view. It is the donor, the giver, who decides how much, if any, to grant to the beneficiary, the recipient.

Figure 5.12 depicts the donor's budget constraint in this application. The initial endowment is the coordinate pair that represents the donor's consumption (on the y axis) and the beneficiary's consumption (on the x axis). There is only one good (which represents consumption of all goods) and its price is \$1/unit. So, if the donor has \$100 and the beneficiary only \$10, we know the initial endowment is at the point 10,100.

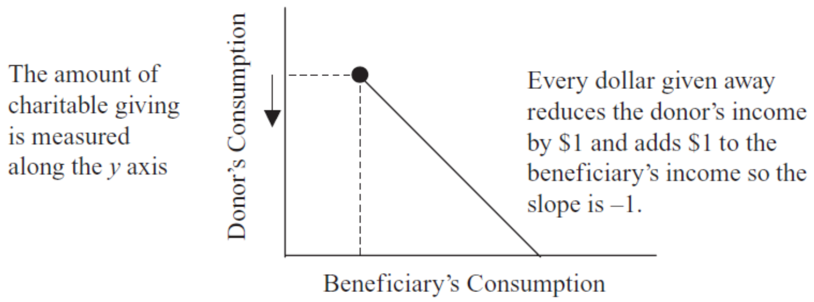


Figure 5.12: The budget constraint.

Giving is modeled as moving down the budget line in Figure 5.12. If the donor gives \$20 away, then she will have \$80 and the beneficiary will have \$30. Of course, the donor could give all of her money away, choosing to be at the x intercept. It is easy to see that the donor decides how much, if any, to give, by choosing a point on the budget line which determines both the donor's own consumption and the beneficiary's consumption.

Thus, at any point on the budget line, we can compute the amount of giving as simply the vertical distance (along the y axis) from the initial endowment to the point on the budget line. If the donor decides to stay on the initial endowment point, then they give nothing to the beneficiary.

The slope of the budget line is -1 because there is a dollar-for-dollar exchange from the donor to the beneficiary.

Notice that this budget line does not extend left or northwest from the initial endowment because that would imply taking money from the beneficiary. The donor cannot do that.

Finally, because we will (of course) be doing comparative statics analysis, we point out that a tax break for those who donate money means that the budget line will have a shallower slope. If the donor gives \$1 and is rewarded, for example, with a 30¢decrease in taxes, then the recipient gets \$1, but the donor actually gave only 70¢. The slope is not -1, but $-(1 - TaxBreak)$. By adjusting the tax break, we can see how the agent responds.

This is too abstract. It is time to go to Excel to understand how the tax break really works.

STEP Open the Excel workbook *Charity.xls* and read the *Intro* sheet, then go to the *MovingAround* sheet.

All you see is a single point at 20,80—this is the initial endowment. The donor gives nothing and there is no tax break.

STEP Change cell C5, the amount the donor gives, to 20. The beneficiary gets the 20, adding it to his initial 20, and new red dot is at 40,60. The slope of the constraint is -1, displayed in I5.

Without a tax break, every dollar given is subtracted from the donor and added to the beneficiary. But the tax code incentivizes giving by lowering the donor's tax liability.

STEP Change E5, the amount of the tax break, to 40%. The red dot jumped up. Hit *ctrl-z* a few times to move back forth between zero and a 40% tax break.

With or without the tax break, the beneficiary still gets 20, but a tax break on charitable donations affects how much the donor actually gave up. With a 40% tax break, the sheet shows that the donor really gave up only 12 because taxes are lowered by 8 (40% of 20). Thus, the slope of the constraint is -0.6.

Wait, if the donor gives 12 and the recipient gets 20, who makes up the difference? The government. The beneficiary gets the full donation, but the donor pays less tax to the government. Clearly, by manipulating the tax break, the government can make charitable giving less expensive to donors.

So, if the tax break increases, what happens to the budget line? Think it through. You can check yourself when we get to the *OptimalChoice* sheet.

But before we get there, we have to consider the donor's preferences. The constraint is only about possibilities. To know what the donor will do, we need to know the donor's utility function.

The neat trick here is to enable the beneficiary's consumption to affect the donor's satisfaction. The way we model giving is to have the self-interested agent care about others.

The usual Cobb-Douglas functional form will represent the donor's satisfaction derived from her own consumption and the beneficiary's consumption.

$$U = BeneficiaryCon^c DonorCon^d$$

As usual, the exponents allow us to model different preferences. If c and d are equal, the donor gets as much satisfaction from her own consumption as the beneficiary's consumption. She is a saint. Although possible, this is unlikely. Most people get more satisfaction from their own consumption and, thus, d is greater than c.

We will use the *OptimalChoice* sheet with different exponent values to see the effect on the graph, but it is worth thinking through two scenarios. What would happen to the indifference curves, starting from $c = d$ as we lowered c? What would happen to the indifference curves if c fell all the way to zero? Again, thinking this through and testing yourself is good way to learn—you can check your answer in the *OptimalChoice* sheet.

It is worth remembering that preferences are not right or wrong. We take them as given and we model the agent as maximizing based on given preferences. It can be difficult to do this—we naturally disapprove of someone who doesn't care about others.

Another source of confusion is that preferences can and do change, but that is not to say that they are chosen by the agent. Changes to preferences are like shocks to other exogenous variables—they are imposed by forces outside the agent's control and then the agent re-optimizes in the new environment.

STEP Proceed to the *OptimalChoice* sheet to see how the donor's optimization problem can be implemented in Excel.

The sheet shows a mathematical expression of the constrained utility maximization problem. The constraint is different than usual. If we write the constraint as an equation, we need to compute the y intercept and incorporate the fact that the donor cannot take from the recipient (the empty space in the northwest corner of Figure 5.12).

We cannot use the usual Lagrangean method to deal with this complicated constraint because it only works with equality constraints. There is an analytical method called Kuhn-Tucker that can be used, but it is beyond the scope of this book.

Fortunately, the numerical method is still available. For Excel and Solver, the complicated constraint is easily handled by adding a second constraint (cell B26) and incorporating it as an inequality—this allows the donor to choose m_1 or greater for the beneficiary. The usual budget line constraint is in cell B25. Applying both constraints gives Solver the equivalent of Figure 5.12 and it has no trouble finding the optimal solution.

Figure 5.13 shows the starting position. The endogenous variables are consumption by beneficiary and donor. These are chosen by the donor to maximize utility subject to the budget constraint.

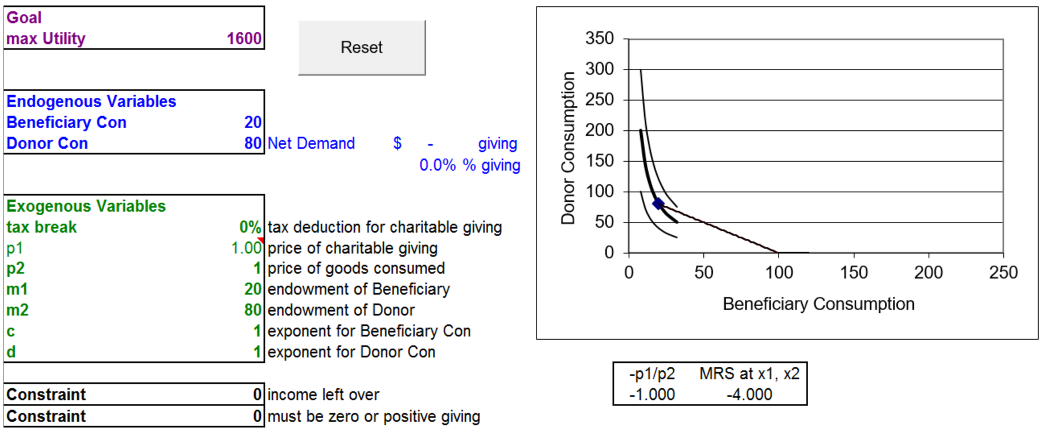


Figure 5.13: Donor with $c = d$ opening position.
Source: Charity.xls!OptimalChoice

The exogenous variables include the amount of the tax break (initially set at zero so the slope of the budget constraint is -1), prices normalized to one, the initial endowment, and the impact of donor and beneficiary consumption on the donor's utility.

With $c = d$, the donor cares as much about the beneficiary as herself and the MRS $> \frac{p_1}{p_2}$ at the initial endowment. We know the donor can increase her satisfaction by traveling down the budget line. For example, suppose the agent decided to donate \$10. How would this affect the chart?

STEP Change cell B11 to 30 and B12 to 70.

The MRS is now closer to the price ratio and utility has risen (from 1600 to 2100). The agent has moved down the budget line and is on a higher indifference curve.

STEP Run Solver to find the initial optimal solution.

The agent chooses the point 50,50 to maximize utility (at 2500), which means she donates $30 to the beneficiary. The net demand is the amount of giving and we express it as a dollar amount and as a percentage of the donor's income (cell D13).

This is one mighty nice donor. She has an incredibly high giving rate of 37.5%. Because $c = d$, she cares as much about the beneficiary as she does herself. It makes common sense that she picks an equal 50,50 split as her optimal solution.

Comparative Statics

There are several shocks to consider. We start with preferences.

STEP Change the exponent for the beneficiary's consumption to 0.2.

This answers the earlier question about the effect of c on the indifference curves: they become much flatter as c falls, ceteris paribus. With $c = 0.2$, the donor does not care as much about the beneficiary as before.

The shape of the indifference curve is tied to the MRS. With $c = 0.2$, the MRS at 50,50 has fallen to 0.2 (in absolute value). The low MRS and flat indifference curve mean that the donor is willing to trade only a little of her consumption for a lot of additional beneficiary consumption.

The culmination of lowering c is a donor who does not care about the beneficiary at all. With $c = 0$, the indifference curves became horizontal, MRS is zero, and beneficiary consumption is a neutral good.

It is obvious that the donor with $c = 0.2$ is not going to be as generous as before when $c = 1$, but how much will they give?

STEP Run Solver. Figure 5.14 displays the result.

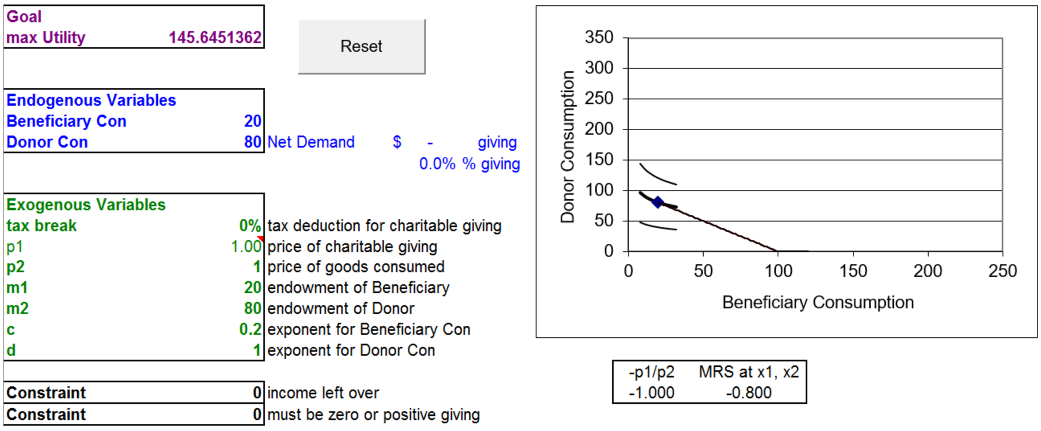


Figure 5.14: Donor with $c = 0.2$ corner solution.
Source: Charity.xls!OptimalChoice

The result is a surprise. The best the agent can do is to donate nothing so that is what she does. Even though the MRS does not equal the price ratio, this donor is optimizing. This is a corner solution.

Our work thus far provides answers to two of the three questions we initially asked.

1. Why do people give to charity? To maximize satisfaction. A donor gives because the consumption of others affects his or her utility. Notice that giving is perfectly compatible with self-interest. The economic model says that the donor feels good when she gives and that is why she gives.

2. What determines how much they give? Clearly preferences matter. How much the donor cares about others (the exponent c in the donor's utility function) plays a major role. Of course, the constraint also matters. Donor's income, beneficiary's income, and the slope of the constraint affect the amount of giving.

3. How can charitable giving be stimulated?

Let's work on the third question. We could try to convince people to care more about others, increasing c (certainly this is a primary goal of religion), but a way to stimulate giving is to lower the price of giving.

As we saw earlier, dollars given to charity reduce the donor's taxable income and reduce tax owed. If the donor is in a 30% tax bracket, every dollar donated to charity saves the donor 30 cents in taxes. Thus, the beneficiary receives the dollar, but the donor is actually paying only 70 cents—with Uncle Sam picking up the remaining 30 cents.

What effect will a 30% tax break have on the budget constraint and charitable giving of a donor with $c = 0.2$? Apply the shock in Excel and find out.

STEP Change the *tax_break* variable (B16) to 30% and note that p_1 becomes 0.70 and the budget line swings out.

The new red budget line is flatter than the original because of the tax break. This answers the earlier question about the effect of a tax break on the budget constraint: the bigger the tax break, the more the line swings and flattens out. This is just like lowering p_1 in the Standard Model.

Notice that the MRS is greater than the slope of the new budget line. This agent can improve her utility by traveling down the constraint. This means she will donate to the beneficiary, as shown in Figure 15.15.

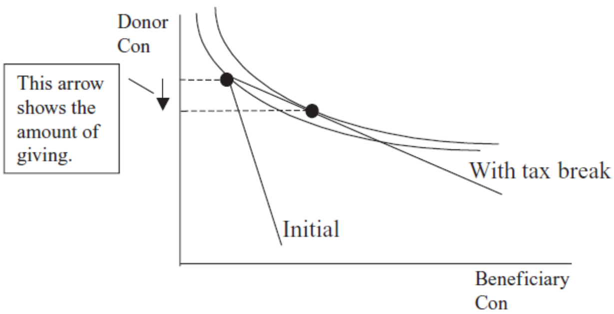


Figure 5.15: The effect of a tax break on giving.

But exactly how much giving does the tax break generate? Let's find out.

STEP With $c = 0.2$ and *tax_break* = 30%, run Solver.

In this case, the tax break has induced charitable giving. It is hard to see on the graph, but the MRS = $\frac{p_1}{p_2}$ condition (under the chart) tells you the indifference curve is now tangent to the budget line. Figure 15.5 shows what happened.

With a tax break of 30%, we get $1.67 of giving which is 2.1% of the donor's income (the American giving rate in 2018).

We can also explore how responsive our donor would be to further shocks in the *tax_break*. We will compute the tax break elasticity of giving.

STEP Change the *tax_break* cell to 40%.

That's a 10 percentage point change in the tax break and a rather hefty 33% change. The budget line swings out a little bit more, but it is hard to see the change in the chart. We know, however, since MRS does not equal $\frac{p_1}{p_2}$, that we need to re-optimize.

STEP Run Solver.

Charity increased from $1.67 to $3.33. That is a big response—a doubling or 100% increase in giving was generated from a 33% increase in the tax break. That is a tax break elasticity of giving of 3.

STEP Proceed to the *CS1* sheet to see a more detailed comparative statics analysis.

Notice that the shock was 1% point, not 10. Notice also that the elasticity from a tax break of 30% to 31% is about 2.87 (H17), not 3. Even though we do not have a reduced form expression, the fact that the measured elasticity depends on the size of the shock tells us that giving is a non-linear function of the tax break.

But regardless of whether it is 3 or 2.87, that high an elasticity is really good news, right? If giving is super-responsive to a tax break, little tweaks in the tax break will generate big increases in giving.

But we need to be careful in how we interpret our result. We do not know whether these preferences and other exogenous variables are representative of many donors. That is an empirical question that requires real-world data. For example, with $c = 0.5$, tax break increases are much less effective in stimulating more giving.

STEP Click the Reset button, change c to 0.5 and the tax break to 30%, and run Solver. Charitable giving is at $17.33.

This makes sense since giving is much higher than it was when $c = 0.2$ and $tax\_break = 30\%$. But what is the tax break elasticity of giving?

STEP Change the tax break cell to 40% and run Solver. Charitable giving rises to \$18.67.

Ponder the computation for a moment. There are a lot of numbers floating around. How would you compute the tax break elasticity of giving?

It is the percentage change in giving divided by the percentage change in the tax break. The numerator is $\frac{18.67-17.33}{17.33} \approx 7.7\%$. The denominator is 33% ($\frac{0.4-0.3}{0.3}$—notice that it doesn't matter if you use the percents version, $\frac{40\%-30\%}{30\%}$). Thus, the tax break elasticity is $\frac{7.7\%}{33\%} = 0.23$.

This result is much less favorable for a policymaker looking to increase charitable giving by manipulating the tax break. For this donor, giving is insensitive to tax break increases.

The Theory of Consumer Behavior can explain a wide variety of giving outcomes. Unfortunately, theory alone does not tell us about the magnitude of a particular effect in the real world. By changing c, we see that the tax break elasticity of giving is drastically affected, ranging from extremely elastic (3) to quite inelastic (0.23). We must gather data and employ econometric techniques to estimate the responsiveness of giving as the tax break changes in the real world. Theory does, however, give us a framework for analyzing the problem.

The Economic Approach Is Widely Applicable

Charitable giving can be viewed through the lens of an Endowment Model using the Theory of Consumer Behavior. The initial endowment is the consumption of the donor and the beneficiary. The donor can choose to give part, all, or none of her endowment to the beneficiary. The amount she gives is determined by that point that maximizes her satisfaction subject to the budget constraint.

We can stimulate giving by lowering the price of giving. This rotates the budget line and yields a new optimal solution. The amount of the increase in giving is an empirical question that cannot be answered by theory alone.

If we view giving as the solution to an optimization problem, we are doing an economic analysis of giving. "An economic analysis" is a phrase often used to communicate that the behavior under consideration will be cast in the framework of optimization and comparative statics.

Many people think economics is about stocks, business, and money. This content-based definition of economics is too limited. Economics is a method of analysis and it can be applied to such "non-economic" issues as charity and many, many other areas.

Seeing charitable giving through the lens of economics does not mean that this is the only way to study charity. The hope is that it provides insight and furthers understanding of what is surely a multifaceted, complex process.

Exercises

1. The total change in charitable giving can be explained via the income and substitution effects for giving. For $c = 0.5$, compute the income and substitution effects when the tax break changes from 30% to 40%. Describe your procedure.

2. Use Word's Drawing Tools to draw a rough sketch of the income and substitution effects for giving, labeling points A, B, and C and using arrows to show the income, substitution, and total effects. Do not include the indifference curves to reduce clutter.

3. Income and substitution effects were originally used to explain Giffen goods. If the tax break increase leads to a decrease in charitable giving, is this Giffen behavior? Why or why not?

References

The epigraph is a *hadith*, which the website islam.uga.edu/hadith.html explains is "a saying of Muhammad or a report about something he did." It would have been easy to find a quotation on charity from any religion because a primary purpose of religion is to encourage us to treat each other with kindness.

If you are thinking of giving to a charitable organization, you can do some background research at www.guidestar.org/ (free registration required to access basic reports) and www.givewell.org/.

Kiva.org is a microcredit organization that allows you to make loans to low-income entrepreneurs all around the world.

If you liked the food stamps application and understand the concept that cash is as good as or better than in-kind (the Carte Blanche Principle), check out www.givedirectly.org.

> During the early 1960s, Kenneth Arrow and Karl Borch published several important articles that can be viewed as the beginning of modern economic analysis of insurance activity.
>
> Georges Dionne and Scott E. Harrington

5.4 An Economic Analysis of Insurance

Why do people buy insurance?

If you are an economist, the answer is easy: because it makes them better off. According to economists, people solve an optimization problem and it turns out that those who buy insurance end up with greater satisfaction, on a higher indifference curve, than if they did not buy insurance.

We will use an Endowment Model to explain how and why insurance is an optimal choice. We will see yet another application of how to solve a constrained utility maximization problem and perform comparative statics analyses.

But the really deep lesson is that the Theory of Consumer Behavior is amazingly flexible and can answer questions from a wide range of problems. In this chapter, we have explored why people save and borrow, give to charity, and, now, buy insurance.

First, we will set up the problem with the usual constraint, indifference curves, and initial optimal solution (with MRS equal to the slope condition). The presence of *risk*, a probability that an event occurs, throws a curveball into the analysis, but we will convert things into our usual framework.

Second, we will do comparative statics. For example, we derive a demand curve for insurance. We can explore the effects of a higher *premium*, the price of insurance, on the quantity of insurance demanded. We are on the lookout for the premium elasticity of insurance.

An Endowment Model of Insurance

There are three parts to every optimization problem. In this case, we have the following:

1. *Goal*: maximize satisfaction (as represented by the utility function).

2. *Endogenous variables*: consumption in two states of nature, good and bad; by choosing the amount of insurance, we control two choice variables at once.

3. *Exogenous variables*: initial assets, potential loss, probability of loss, insurance premium, and preferences over the states of nature.

As usual, we start with the constraint, then turn to preferences, and finally use the constraint and utility function to find the initial solution.

STEP Open the Excel workbook *Insurance.xls* and read the *Intro* sheet, then go to the *Constraint* sheet.

The idea is that you have an asset, say your car or house, which may suffer a given amount of damage from an accident, called the *PotentialLoss*, with a known probability, π (the Greek letter, pi) that the damage occurs. Initially, the *PotentialLoss* is $10,000, which is only a fraction of the value of the house.

You can buy K dollars of insurance, this is the *InsuredAmount*, by paying a price (called a premium) of γ (the Greek letter, gamma) per $100 of insurance coverage. On opening, you are not buying any insurance.

If you buy insurance, then if the accident occurs, you get reimbursed for the loss. You can buy insurance in $100 increments, up to the *PotentialLoss*, in which case you would be fully insured. The trade-off is that you have to pay for insurance up front, before you know if the accident will happen or not.

After you decide how much insurance to buy, there are two possible outcomes, known as *states of nature*: the bad and good outcomes.

STEP Click on cell B18 to see the formula for your consumption in the bad outcome.

The *ConsumptionBad* outcome means the accident actually occurred, leaving your consumption as $InitialAssets - PotentialLoss + K - \gamma K$. You subtract the loss that occurred and the amount you paid for insurance (γK), but you add the amount K that the insurance company pays you because you suffered the accident. You could be fully covered, but you do not have to be. You decide how much insurance to buy.

Your consumption in the good state of nature is simply $InitialAssets - \gamma K$. You do not suffer the accident, but you still have to pay for the insurance.

STEP Click on cell B19 to see the formula for the good outcome.

Cells B23:B25 display in which state of nature you end up. Cell B23 has the formula "=RAND()." This draws a number from a uniform distribution on the interval [0,1].

STEP Hit the F9 key on your keyboard repeatedly to understand Excel's RAND() function works.

Each time you hit the F9 key, Excel draws a random number from 0 to 1 in cell B23. The number drawn is never smaller than zero or bigger than one.

Cell B24 converts the random draw in the cell above it into a zero or a one—zero means the accident did not happen (good outcome) and one means it did (bad outcome). It uses an IF statement to display a "1" (the accident happened) when the random draw is less than 0.01 (the value of π in cell B8).

It is hard to see that anything is really happening in cell B24 because the probability of the accident occurring is so small.

STEP Change π to 50%, then hit F9 a few times. You should be able to see cell B24 flip from 0 to 1 and back again as the random draw is less than 0.5 and greater than 0.5.

Notice that the *FinalAssets* variable, cell B25, depends on whether or not the accident occurred.

Next, let's buy some insurance to see what that does to the spreadsheet.

STEP Click the Reset button and set cell B13 to \$1000. This will cost you \$10.

Notice the values for the good and bad states of nature. You have altered both. If the accident occurs, your consumption is \$25,990, which is \$990 better than the \$25,000 for the bad outcome when you did not buy insurance. Of course, the good outcome is \$10 lower (at \$34,990) in the good outcome because you have to pay for the insurance even when the accident does not occur.

STEP Click the Graph the Constraint button. Click OK to the "4" points default option and read each text box as it appears. At the end, the budget line is displayed (see Figure 5.16).

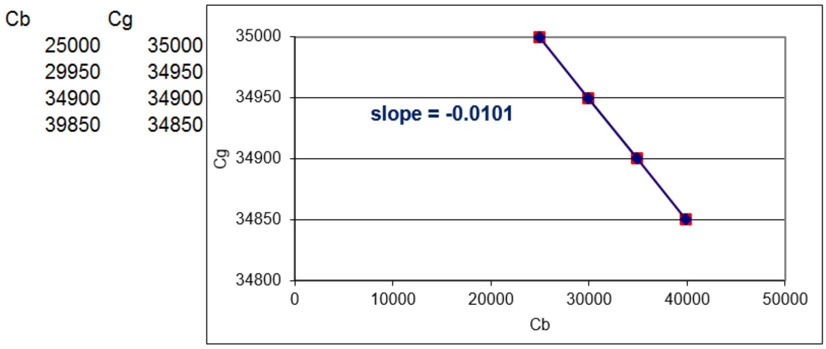


Figure 5.16: The budget line.
Source: Insurance.xls!Constraint

From the initial endowment (C_b, C_g without insurance), you can move down the budget line by buying insurance. You lower your consumption in the good state of nature (C_g is on the y axis), but raise it in the bad state of nature (C_b is on the x axis).

The terms of trade (the slope of the budget line) are determined by gamma (the insurance premium). The slope of the budget line is $-\frac{\gamma}{1-\gamma}$, which with $\gamma = 0.01$ is $\frac{-1}{99} = 0.01$ (the "01" keeps repeating forever). The graph rounds the slope to five decimal places.

Changes in initial assets or potential loss shift the budget constraint. We are interested, however, in deriving a demand curve for insurance so we will shock the insurance premium (the price of insurance). This will pivot or rotate the budget line.

STEP Change the insurance premium to \$1.20 per \$100 of insurance coverage.

You see the familiar swinging in (clockwise rotation) from a p_1 increase. A buyer of insurance would be disappointed in this shock because her consumption possibilities are diminished.

Now that we understand the constraint, we turn to the agent's tastes. We model utility as preferences over the two states of nature. The fact that there is risk involved in which state of nature occurs complicates things.

Instead of having utility simply depend on the amount of consumption in the good and bad outcomes, we include the agent's expectations about the chances of each outcome occurring. Fortunately, our usual Cobb-Douglas functional form can incorporate this new information.

We use the exponents in the Cobb-Douglas functional form to represent the agent's beliefs about the probability of the accident occurring. There are two simplifying assumptions. The first is that the agent accurately gauges the probability of loss, which means we can use π as the exponent in the utility function. The second assumption uses the fact that there are only two mutually exclusive outcomes so the bad outcome occurs with probability π and the good outcome has likelihood $1 - \pi$. The possibility of a partial loss is assumed away.

The utility function is then

$$U = C_b^{\pi} C_g^{1-\pi}$$

The idea behind the utility function is simple: The higher the probability of loss, the more the agent will care about the bad outcome. In terms of the indifference map, the higher π, the steeper the indifference curves. This means the agent cares more about consumption in the bad state of nature as risk rises.

Unlike the Standard Model where the exponents in the Cobb-Douglas utility function can be used to represent changes in preferences, changes in the exponents do not indicate a change in preferences for the utility function with risk. To get a change in preferences, we need an entirely different utility function.

It is beyond the scope of this book, but there is a great deal of research on choosing with random outcomes. The field of behavioral economics was born with the discovery of paradoxes, violations of transitivity and other inconsistencies, when people made choices involving randomness. Our Cobb-Douglas utility function can be written as an expected utility function by simply taking the natural log:

$$lnU = \pi C_b + (1 - \pi) C_g$$

This function reflects risk averse preferences. It is a starting point for modeling attitudes and feelings toward risk and randomness.

STEP Proceed to the *Preferences* sheet to see an implementation of the Cobb-Douglas utility function.

The sheet tries to give a new way of understanding how constrained utility maximization works. It shows consumption in the bad and good states of nature, \$25,000 and \$35,000, respectively, without insurance. This is the initial endowment point.

With $\pi = 1\%$, we can compute the level of utility for the initial endowment combination of consumption in the bad and good states of nature. This is shown in cells D13 and E13. We can also compute the MRS at the initial endowment, displayed in cells G13 and H13.

The *Dead* and *Live* utility and MRS are the same because we are at the initial endowment. The *Dead* cells are numbers. They will not change when we change the cells in column B. The *Live* cells contain formulas. They will update when you change the values of $C_b, C + g$, and π.

STEP Ponder and answer the question in cell A6. Click on the Answer when you are ready. Do the same for B10.

The Live utility and MRS cells change when you change cells B13 and B14. As you moved down from the initial endowment, utility rose and the MRS fell. It got closer to the slope which means we are closer to the optimal solution.

We are ready to find the initial optimal solution.

STEP Proceed to the *OptimalChoice* sheet.

The *OptimalChoice* sheet reproduces the *Constraint* sheet, but it adds the indifference map to the chart and displays the slope of the budget line and the MRS at the bottom of the chart. It also displays the utility in cell B20 from the chosen consumption in the two states of nature.

It is really hard to see what is happening with the indifference curve at the initial endowment and the slope of the budget line.

STEP Zoom in—double-click the y axis and make the minimum bound 34800 and the maximum bound 35200.

You can now see clearly that when MRS > slope of the budget line, the budget line cuts the indifference curve. By moving down the budget line, you can reach higher levels of satisfaction.

STEP Enter 5000 in cell B13 to see where the agent stands when buying $5000 of insurance.

The chart shows movement down the budget line to a higher level of utility. We are closer to the optimal solution, but not there yet because MRS is not equal to the slope of the budget line.

STEP Run Solver to find the optimal solution.

The Solver dialog box is notable for the fact that there are no constraints. The way we implemented the problem in Excel enabled us to directly maximize the utility cell by choosing a single variable, the amount of insurance purchased. We can still use, however, the canonical Theory of Consumer Behavior graph to show the result.

At the optimal solution, the consumer decides to buy $10,000 of insurance. In the bad state, if the accident occurs, the agent is fully covered, so is consumption $35,000? No, because the agent has to pay $100 for the insurance, so consumption would be $34,900 in the bad state.

In the good state, where there is no accident, consumption is also $34,900. This is surprising. Insurance has removed the effect of risk. Consumption is the same in both states. This is an extreme example of diversification.

Diversification is a strategy to lower risk by spreading your wealth over different states of nature. By moving $100 from the good state of nature (buying insurance), the agent has a guaranteed level of utility regardless of whether the accident happens. Without insurance, the expected return is $34,900 since 99% x $35,000 + 1% x $25,000 = $34,900. But the agent has to put up with the risk of every 1 in 100 times getting $25,000. By diversifying, the expected return is the same, $34,900, with absolutely no risk.

Such a perfect result—the complete elimination of risk—relies on the fact that the two states of nature are perfectly correlated. In the real world, when states of nature are not perfectly correlated (such as the stock market), diversification can lower risk while maintaining the same expected return, but it cannot completely eliminate it.

We know that people buy insurance because it increases satisfaction. This application models choosing the amount of insurance that maximizes utility subject to the budget constraint. Next, we use the model to derive a demand curve for insurance.

Comparative Statics

The procedure is straightforward: we vary the insurance premium (the price of insurance), γ, ceteris paribus, and track the optimal amount of insurance purchased (K) to derive a demand curve for insurance.

We use numerical methods and leave the analytical work for the exercises.

STEP In the OptimalChoice sheet, change γ to \$1.30 per \$100 of insurance. What happens?

The budget line (displayed in red on your screen) gets steeper. The agent needs to re-optimize.

STEP Run Solver to find the new optimal solution.

If you did not zoom in on the y axis as instructed earlier, it is hard to see on the chart, but the cells below the chart confirm that the MRS equals the slope of the budget line when the agent buys \$1847 of insurance.

We can conclude that demand for insurance is downward sloping when the premium rises from \$1.00 to \$1.30 since the amount of insurance purchased fell from \$10,000 to \$1847. That is extremely responsive.

STEP Compute the price elasticity of demand. Proceed to the *CSgamma* sheet to check your answer. Notice that Excel tries to help when you enter the formula by formatting the result as dollars. This is incorrect. Elasticity is unitless.

The *CSgamma* sheet shows that the CSWiz add-in was used to explore the effect of the insurance premium on the amount of insurance purchased. Gamma was incremented by 0.1 (10 cents) with 10 shocks. Optimal $K, \gamma K, C_b$, and C_g were tracked as γ changed. The sheet includes a chart of $K^* = f(\gamma)$, the demand curve for insurance.

Notice the curious behavior of the model as γ rises: at \$1.40, optimal K becomes negative. This is an Endowment Model. When premium prices get high enough, the agent switches from buying insurance to selling insurance!

If this option is not allowed, you can impose the constraint in Excel that K be greater than or equal to zero. Then, with high premiums, the consumer is at a corner solution and buys no insurance.

Modeling Insurance via the Endowment Model

Insurance is another application of an Endowment Model in the Theory of Consumer Behavior. The usual ideas were applied: the budget constraint, preferences, and MRS equals slope of budget line at the optimal solution. In addition, the usual recipe of the economic approach, finding the initial optimum and then comparative statics, was followed.

But this application does have its own twists and novelties. We used a Cobb-Douglas functional form to model satisfaction where the exponents reflect the probabilities of the states of nature. We also used Excel's Solver without a budget constraint because of the way we implemented the problem in Excel. To be clear, this problem can be solved via the Lagrangean method (see the first exercise question) and we could have implemented a "max U subject to a constraint" model in Excel. We would get, of course, the same answer.

Exercises

1. Use analytical methods to derive a general reduced form solution for K^*. Show your work.

   Although you can use the Lagrangean method, it is easier to maximize the utility directly, substituting in the values for each state of nature.

$$\max_K U = C_b^{\pi} C_g^{1-\pi}$$

The key is that consumption in the good and bad states of nature depends on K:

$$C_b = InitialAssets - PotentialLoss + K - \gamma K$$

$$C_g = InitialAssets - \gamma K$$

We can simply substitute these equations into the utility function and maximize this:

$$\max_K U = [InitialAssets - PotentialLoss + K - \gamma K]^{\pi}[InitialAssets - \gamma K]^{1-\pi}$$

2. Compare the analytical versus numerical approaches by evaluating your answer to question 1 at the initial parameter values in the *OptimalChoice* sheet. (Click the Reset button if needed.) Do you find that $K^* = \$10,000$?

3. Use your reduced form for K^* to find the probability of loss elasticity of insurance demand at $\pi = 1\%$. Show your work. If you cannot find the reduced form, use

$$K^* = \frac{[\pi - \gamma]\,InitialAssets + [1-\pi][\gamma]PotentialLoss}{[\gamma][1-\gamma]}$$

4. Use the Comparative Statics Wizard to find the probability of loss elasticity of insurance demand from $\pi = 1\%$ to 1.1%. Take a picture of your results, including the elasticity calculation.

5. Compare your answers in question 3 and 4. Do these elasticities differ? Why or why not?

References

The epigraph is from the first page of *Foundations of Insurance Economics* by Georges Dionne and Scott E. Harrington, editors, published in 1990. Insurance economics as an organized subfield is quite young, but rapidly growing. It focuses economics, probability, and computer science on applied problems in the world of risk and insurance.

In their wildly popular *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything* (2005), Steven D. Levitt and Stephen J. Dubner include this example from the world of insurance markets:

> In the late 1990s, the price of term life insurance fell dramatically. This posed something of a mystery, for the decline had no obvious cause. Other types of insurance, including health and automobile and homeowners' coverage, were certainly not falling in price. Nor had there been any radical changes among insurance companies, insurance brokers, or the people who buy term life insurance. So what happened? The Internet happened. In the spring of 1996, Quotesmith.com became the first of several websites that enabled a customer to compare, within seconds, the price of term life insurance sold by dozens of different companies. (p. 66)

The freakonomics.com website has podcasts and other resources.

Chapter 6

Bads

Risk Versus Return

Automobile Safety Regulation

Labor Supply

One of the best-documented propositions in the field of finance is that, on average, investors have received higher rates of return for bearing greater risk.

Burton Malkiel

6.1 Risk Versus Return

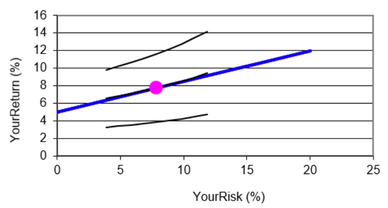
In finance, a *portfolio* means the total holdings of stocks, bonds, and other securities of an individual (or other entity, such as a trust or foundation).

Because the investor can decide which securities to include in her portfolio, in other words, because choices are made, we can apply the method of economics. Optimal Portfolio Theory is the name given to the application of the Theory of Consumer Behavior to analyze decisions about which assets to hold.

An important stop on our journey is shown in Figure 6.1, the initial solution to the constrained optimization problem.

Exogenous Variables		
Risk-free Return	5	r_f
Average Risky Return	12	r_m
Market Risk	20	σ
a	1	**exponent on Risk**
b	1	**exponent on Return**
Endogenous Variables		
YourMix	39.2857143	% of risky asset
YourRisk	7.85714286	
YourReturn	7.75	
Goal		
Utility	171.607143	

Risk	Return
0	5
2	5.7
4	6.4
6	7.1
8	7.8
10	8.5
12	9.2
14	9.9
16	10.6
18	11.3
20	12



slope	MRS
0.3500	0.3500

Figure 6.1: The initial solution.
Source: RiskReturn.xls!OptimalChoice

There are some strange features in Figure 6.1 and you are not expected to understand it right away. Perhaps the weirdest thing is that the budget constraint and indifference curves are upward sloping. Because risk (on the x axis) is a bad (not a good), the agent substitutes *more* of the bad for *more* of the good (return, on the y axis) on an indifference curve.

There are also, however, elements that are familiar and comfortable in Figure 6.1. There are exogenous (green) and endogenous (blue) variables with a goal. There is a constraint and a few curves with a tangency highlighted that is obviously the optimal solution. And we can see the usual MRS = slope condition below the chart.

Of course, Figure 6.1 is just the initial optimal solution. There is more to do than simply finding the initial solution. That is why Figure 6.1 is an important stop on our journey, but we have more to travel. We want to explore how the optimal solution changes as one of the exogenous variables changes, ceteris paribus. This is called comparative statics analysis.

The procedure that defines the Theory of Consumer Behavior is clear: constraint, preferences, find initial solution, then comparative statics to make statements about how a shock variable affects an optimal choice variable. We will do an elasticity computation and interpretation of the shock. The short way of saying all of this is to just say that we are going to do an economic analysis of portfolio choice.

But since we will be talking about returns from assets, volatility, and the stock market, let's look at some data to make sure we understand some basic facts.

Stock Market Returns

STEP Open the Excel workbook *RiskReturn.xls* and read the *Intro* sheet, then go to the *Data* sheet.

The sheet has returns from the S&P 500 index, a group of 500 large companies, downloaded from www.moneychimp.com/features/market_cagr.htm.

These data are used to show that returns are quite volatile. The sheet also explains the difference between the arithmetic and geometric mean.

STEP Read the explanation in the *Data* sheet, scroll down to see the data (all the way down to 1871), and then click the Next button.

This reveals more material. Keep reading and clicking the Next buttons until you get to the end and then be sure to click the More button.

Of utmost importance is that you understand the volatility in the S&P 500 returns. They swing wildly and unexpectedly, from incredible spurts of 50% to staggering losses of almost negative 50%.

STEP Look in columns A and B of the *Data* sheet at the 1930s, during the Great Depression. Scroll slowly back up, looking at the data.

The volatility in the stock market, measured by the standard deviation, SD, of almost 20%, is unwelcome and unsatisfying. The fear of financial disaster and the risk of losing money lowers utility.

Then why do people put their money in assets like the S&P 500? Because the overall annual return is high—much higher than safer, less volatile assets. For the S&P 500, the overall annual return (as you now know, measured by the geometric mean, GM, or compound annual growth rate, CAGR) is about 9% per year.

The stock market's 9% annual return is much higher than that available from a safe, stable asset that produces consistent annual returns like US Treasury Bills. Cell H10 in the *More* sheet shows that the SD is a mere three percentage points. The variability arises because the yield changes over time, but once you buy a US Treasury note for a particular length of time, you can be quite sure that you will be paid. But right below the SD we see that the overall annual return is one-third of the stock market's return.

The key point is that financial markets offer the investor a menu of options, from low risk, low return to high risk, high return, and the investor chooses. All we need to do is model that choice as an optimization problem.

Optimal Portfolio Theory

The *Compare*, *Mix*, and *Constraint* sheets in *RiskReturn.xls* demonstrate that an investor can mix two assets, a risk-free and a risky asset, to create a portfolio that has a particular combination of risk and return.

The investor is not free to pick any combination of risk and return. They must stay within the constraint imposed by the market. The idea is that you have a fixed amount of money, say $10,000, to allocate across two assets.

The *risk-free asset*, say a US Treasury Bill, has a certain (practically speaking) rate of return, say 5% per year, which is unrealistically high for the current climate. Thus, you are sure to get 5% of $10,000, or $500, along with your initial investment of $10,000 at the end of the year. Each year, a $10,000 investment is guaranteed to produce $500 of return.

The *risky asset*, say a mutual fund of stocks, has a greater return, but also volatility in the actual realized return. We will suppose that the actual return will be drawn from a normal distribution centered on 12%, with a spread of 20%. Both of these values are a little higher than the historical experience of the S&P 500 (in the *Data* sheet). Our parameter values mean that the typical realized value in our hypothetical world will be around 12% ± 20% points. It also means you will actually lose money (suffering a negative return) about a quarter of the time.

But this is way too abstract. To understand the meaning of these parameters, let's work on a concrete problem with actual numbers and a clear display of what is going on.

STEP Go to the *Compare* sheet.

The bell-shaped curve is the normal distribution from which each year's return will be drawn. The center and spread are controlled in cells A2 and C2.

The sheet allows you to run the two investments against each other and shows how volatility impacts the annual returns.

STEP Click the Invest One Year button.

For the risk-free asset, cells I3 and L3 show 5% and $500. In other words, if you place $10,000 in the risk-free asset, these are the returns on that investment.

The risky asset is different. Cells J4 and M4 show a number that is taken from the normal distribution on the left of your screen, centered on 12 with an SD of 20. Thus, the number in J4 is likely to be around 12, but could easily be in the range −8 to 32 (±1 SD from the average) and roughly 95% of the time will be between −28 and 52 (±2 SDs from 12).

STEP Click the Invest One Year button a few times.

You can clearly see what is happening here. The return from the risk-free asset is always the same, but the risky asset bounces around.

Once you have more than one year of returns, the display shows more information in columns P:S. You can see the arithmetic mean of the returns, SD, the exact geometric mean, and its approximation.

STEP Click the Invest One Year button many times, at least 20.

Notice what is happening to the average of the returns of the risky asset as you keep adding years: The average return is converging to 12% (the average return from the normal distribution in A2). In other words, over the long haul, the risky asset will outperform the risk-free asset. However, in any one year, the risky asset can do pretty badly. Look at your screen to confirm that this is true. You will see some whopper losses (and gains)—just like the real-world S&P 500 data.

STEP Click the Reset button and set the dispersion to 6% (in C2). Repeatedly (many times) click the Invest One Year button.

The SD of the normal distribution controls the variability. The lower SD makes the normal distribution much more spiked. In other words, the draws from the distribution are much more concentrated at the average and it is much less likely that you will see values far from the center of the distribution.

As you get one yearly return after another (keep drawing more returns), it is easy to see that the returns are much closer to 12%. You will rarely lose money with an average of 12% and an SD of 6%.

In finance, risk is denoted by the Greek letter sigma, σ. The SD and σ are the same thing. Both represent risk as volatility and bounce in returns, including the possibility of negative returns. Risk is bad and undesirable. The lower the risk, the better.

What determines the amount of risk in the risky asset? That depends on the asset. We have seen that the S&P 500 has a lot of volatility. From 1871 to 2019, it has experienced an overall annual return of about 9% with an SD of 18%. The *More* sheet showed that other assets have different volatility. So, the investor is given the average and SD parameters of various assets and chooses what to invest in.

Although we ran risk-free and risky assets in the *Compare* sheet, in fact, the choice is not simply between a risk-free and a risky asset. You can combine the two in varying proportions.

For example, you could split your investment and put $5000 in the risk-free and $5000 in the risky asset. In this case, your return would be halfway between the risk-free and risky assets:

$$\frac{r_f + r_m}{2} = 8.5\%$$

Although the return is lower than using the risky asset alone, your risk, the variability in returns, would be cut in half also.

STEP Proceed to the *Mix* sheet to see this idea in action.

The *Mix* sheet is the same as the *Compare* sheet, except it has a scroll bar in H1 to control the allocation of your $10,000 across the two assets.

STEP After you set the scroll bar value (any value will do; pick the one you think makes the most sense for you), click the Invest One Year button many times.

You should be able to see that the average return for your mix (or portfolio) converges on a return that is in between the risk-free and risky assets. In other words, you can choose the return and risk that you get. You must, however, trade them off—more return requires accepting more risk.

STEP Experiment. Use the Reset button to try different mixes and parameter values (yellow-backgrounded cells A2, C2, and F2).

You can copy the *Mix* sheet (right-click the sheet tab, select *Move or Copy*, and check Create a Copy) if you want to compare different scenarios. The more you experiment, the more you learn.

Your work in the *Compare* and *Mix* sheets makes understanding the constraint much easier because you have seen that there are two assets that can be mixed to form a portfolio with a continuous range of risk and return possibilities. This constitutes the constraint for the investor. He or she is free to choose combinations of risk and return, trading higher risk for greater return.

STEP Proceed to the *Constraint* sheet.

There are two endogenous variables, *YourRisk* and *YourReturn*, in cells B14 and B15. These are the risk and return you have chosen, in other words, a single point on the budget line. However, we can create a single variable, *YourMix* (just like in the *Mix* sheet) that controls the proportion of your investment in the two assets and the values of risk and return you select.

Clearly, you can mix the risk-free and risky assets in any combination from 0 to 100%. Zero means you buy just the risk-free asset and 100% means you buy only the stock market.

Do not confuse the exogenous variable *Market Risk* with the endogenous variable *YourRisk*. The riskiness of the risky asset, sigma, is exogenous to the agent. But the agent determines how much risk to take and, therefore, the chosen amount of risk is endogenous.

STEP Change B13 to 20%, 50%, and 90%.

As you change B13, the red dot moves on the constraint. You can put the red dot wherever you like along the line. At 50%, you are setting *YourRisk* to 10% (this is the variability in the 50/50 portfolio) and *YourReturn* to 8.5% (halfway between r_f and r_m).

The equation of the budget line (derived in the *Constraint* sheet) is

$$YourReturn = r_f + \frac{r_m - r_f}{\sigma} YourRisk$$

Clearly, if you choose a risk of zero, then your return is the risk-free return. This is the y intercept. As you accept more risk, your return grows with a slope given by $\frac{r_m - r_f}{\sigma}$

Notice that combinations under the budget constraint are feasible, but will not be selected because more return can always be obtained at the same risk by going straight up. Points to the northwest of the line are more desirable, but are unattainable.

Which mix is the best, the optimal choice? We cannot answer this question with the constraint alone. It tells us only the choices we can make. To answer the question, we need to model preferences.

But before we leave the constraint, let's explore the effect of a change in sigma, Market Risk. This will be our shock variable when we do comparative statics analysis.

Remember when you lowered the SD to 6% and that made the variability in the risky asset go way down? That was a welcome shock. What would happen to the constraint if we applied that shock? Before we do it, ponder the question. Do you have an answer? Let's see how you did.

STEP Change Market Risk, cell B10, to 6.

The budget line rotates up (counterclockwise) around the y intercept. This gives the investor access to higher returns with the same risk or the same return with less risk. Mathematically, it also makes sense since we lowered the denominator in the slope, so the slope term increased, making the line steeper.

STEP Proceed to the *Preferences* sheet to see how we handle risk as a bad.

Our usual Cobb-Douglas functional form can be modified to reflect a bad with a simple tweak:

$$U(YourRisk, YourReturn) = (30 - YourRisk)^a YourReturn^b$$

The clever trick here is subtracting a variable from a constant, which has been chosen to be bigger than the possible values of the variable. By having a constant, 30, which is a bigger number than the relevant range for *Risk* (from zero to 20), as we increase the chosen amount of *YourRisk*, 30–$YourRisk$ falls. This gives us a bad because utility falls as *YourRisk* rises (for $YourRisk < 30$). *YourReturn* is a good—as *YourReturn* rises, so does utility.

The chart shows three representative, upward sloping indifference curves. The investor gets equal satisfaction by the combinations of risk and return on a single indifference curve. If the investor takes on more risk, she must be given more return to compensate.

STEP The agent is free to choose any combination of risk and return that is on the budget line. Change B12 to 50.

Figure 6.2 shows the result. In addition to the three original indifference curves with a black dot, three new curves are displayed along with a red dot. The black dot is the initial 75% mix choice and it produced *Dead Utility* of 153.75 and a *Dead MRS* of about 0.6833.

Dead Utility	Live Utility	Dead MRS	Live MRS
153.75	170	0.683333	0.425

slope
0.3500

Price of risk $=(r_m - r_f)/\sigma$

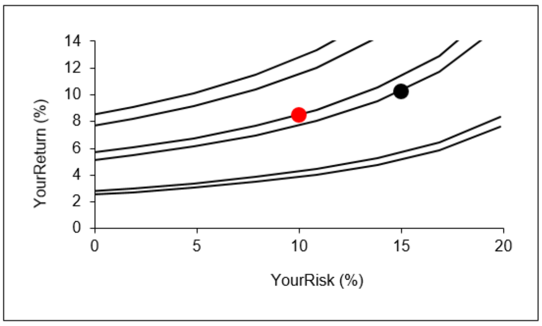


Figure 6.2: The investor's indifference map, $a = b = 1$.
Source: RiskReturn.xls!Preferences

The red dot is *live* in the sense that it depends on the value of B10. The chart displays the indifference curve that goes through the mix value in B10, along with an indifference curve and another below it.

A mix of 50% risky is better than 75% for this investor because utility went up. The red dot is on a higher indifference curve. Notice also that the MRS fell, getting closer to the slope of the budget line. That means the investor is getting closer to the optimal solution.

STEP Change B12 to 90.

Now the reverse is true. The red dot is on a lower indifference curve and the MRS is farther away from the slope.

STEP Change the exponent on *YourReturn* in B19 to 4 and click the Set button.

The indifference curves are now much flatter. What does this mean?

STEP Change B12 to 50 and 90.

We are getting different results than before? What is going on?

If $b > a$, the investor cares more about return than risk. The flat indifference curves (with low MRS) mean that they are willing to accept a lot of risk for a little more return. These preferences mean that this investor will find an optimal solution with a high risk, high return combination.

STEP Change B19 to 0.4 and click the Set button. Explore the satisfaction produced by mixes of 50% and 90%. What do you learn?

With a low b (lower than a), this investor is more concerned with risk. They are conservative and their optimal solution will lie on a low mix value. In fact, these preferences produce a corner solution, with the investor putting all $10,000 into the risk-free asset.

Preferences are not right or wrong. If you are young and saving for retirement, it makes sense that $a < b$, but even then, if a person does not like risk, that is not a defect. An aggressive investor is not in any sense better than a conservative investor. Some people like risk and others do not in the same way that some people like broccoli or the color blue and others do not.

Preferences are not set in stone. They can be affected by the environment. A short time horizon, such as needing funds for college in a year, will rotate the indifference map, reflecting an investor who is more conservative. Likewise, retired people, typically, become more conservative and less willing to accept risk.

With the constraint and preferences modeled, we are ready to find the optimal solution.

STEP Proceed to the *OptimalChoice* sheet to see the numerical method in action.

The *OptimalChoice* sheet opens with an inefficient solution. The MRS is greater than the slope of the budget line so the indifference curve cuts the line. The agent should move down the line, accepting less return for less risk. This increases satisfaction. But how far down to travel?

STEP Run Solver to find the answer to this question.

At the optimal solution, the MRS equals the slope of the budget line and the agent is on the highest attainable indifference curve.

For this agent (with these attitudes toward risk and return) and the given market trade-off between risk and return (captured by the equation of the budget constraint), the optimal solution is found with a mix of about 39% of funds invested in the risky asset. Thus, the optimal risk to accept is $7\frac{6}{7}$ and the optimal return is $7\frac{3}{4}$.

Via analytical methods, we can use this Lagrangean to find optimal *YourRisk* (x_1) and *YourReturn* (x_2).

$$\max_{x_1,x_2,\lambda} L = (30 - x_1)x_2 + \lambda\left(x_2 - \left(\frac{r_m - r_f}{\sigma}\right)x_1 - r_f\right)$$

Try doing this problem and if you get stuck, the solution for a similar problem in the *Q&A* sheet is in the *Answers* folder.

Comparative Statics

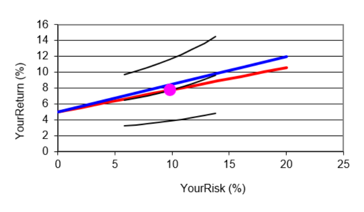
As usual, there are a number of comparative statics exercises to consider and they can be done via numerical or analytical methods. Let's explore the effect of an increase in sigma, the amount of risk the market forces you to bear in return for better performance.

STEP In the *OptimalChoice* sheet, increase σ from 20 to 25. What happens?

Figure 6.3 and your screen show a new, red budget line that has rotated clockwise and down.

Exogenous Variables		
Risk-free Return	5	r_f
Average Risky Return	12	r_m
Market Risk	25	σ
a	1	exponent on Risk
b	1	exponent on Return
Endogenous Variables		
YourMix	39.2857136	% of risky asset
YourRisk	9.8214284	
YourReturn	7.74999995	
Goal		
Utility	156.383929	

Risk	Return
0	5
2	5.56
4	6.12
6	6.68
8	7.24
10	7.8
12	8.36
14	8.92
16	9.48
18	10.04
20	10.6



slope	MRS
0.2800	0.3841

Figure 6.3: Increasing sigma, Solver yet to be run.
Source: RiskReturn.xls!OptimalChoice

The flatter slope is bad for the investor because consumption possibilities have been reduced. The market says that for a given amount of return, you must accept more risk. How will the investor respond to this shock?

STEP Run Solver to find out.

You will see that the agent chooses less risk and less return. What elasticity is under consideration here? There are several. There is the *sigma* elasticity of *YourRisk*, the *sigma* elasticity of *YourReturn*, and the *sigma* elasticity of *YourMix*.

Of course, these elasticities can also be computed at a point, using the derivative. One of the exercises asks you to do exactly that.

STEP Try your hand at computing the *sigma* elasticity of *YourRisk* from $\sigma = 20\%$ to 25%. Check your answer in the *CSsigma* sheet.

Of course, these elasticities can also be computed at a point, using the derivative. One of the exercises asks you to do exactly that.

Because the change in sigma is a change in the slope of the budget line, we can use the Slutsky decomposition approach to break down the total effect into income and substitution effects. This work is left for you as an exercise.

Asset Allocation is an Optimization Problem

Optimal Portfolio Theory is yet another application of the Theory of Consumer Behavior. The twist here is that one of the choices, risk, is a bad. The agent cannot ignore risk. She is forced to accept more risk to secure greater return.

The core concepts of the Theory of Consumer Behavior remain easily visible: a budget constraint describing consumption possibilities, preferences translated into an indifference map, maximization of utility given a budget constraint, and MRS equals slope of budget line at the optimal solution.

Perhaps most importantly, once we cast the problem as a choice, how to allocate assets among stocks, bonds, and other financial instruments, we are firmly in the land of Economics. This particular optimization problem is different from previous applications in that individuals are keenly interested in getting the optimal solution right. There is often a lot of money at stake and mistakes can prove costly (for example, with a retirement portfolio).

As economists, we remain interested in comparative statics. Changing preferences are an important shock variable in this application. We do not shake our heads at the conservative investor who finds an optimal solution (given conservative preferences) at a low risk, low return point.

Exercises

1. Use the equation that follows to solve for *YourRisk** (x_1) and *YourReturn** (x_2) in terms of the exogenous variables. Show your work.

$$\max_{x_1, x_2, \lambda} L = (30 - x_1)x_2 + \lambda \left(x_2 - \left(\frac{r_m - r_f}{\sigma} \right) x_1 - r_f \right)$$

2. Use your reduced form solution to find the *sigma* elasticity of *YourRisk* at $\sigma = 20\%$ (and the values of the other exogenous variables from the initial position of the OptimalChoice sheet—click the Reset button if needed). Show your work.

3. Use Word's Drawing Tools to draw a well-labeled graph that depicts the total, income, and substitution effects for *YourRisk*. Make the substitution effect greater than an opposing income effect.

4. Compute the total, income, and substitution effects for *YourRisk* for the change in sigma from 20% to 25%. Show your work and describe your procedure.

References

The epigraph is from page 184 (9th edition) of a classic, excellent book on personal finance and the stock market. *A Random Walk Down Wall Street* by Burton Malkiel was originally published in 1973 by W. W. Norton & Company and the 12^{th} edition came out in 2020. This is not one of those silly books with a scheme to beat the market. Malkiel is sober and reliable. On page 26, he says,

> Let me make it quite clear that this is not a book for speculators; I am not going to promise you overnight riches. I am not promising you stock-market miracles. Indeed, a subtitle for this book might well have been *The Get Rich Slowly but Surely Book*.

For a much deeper analysis of finance with an Excel-based presentation style, see *Principles of Finance with Excel* by Simon Benninga (New York: Oxford University Press, 2017. 3rd edition).

> Minivans have the lowest fraction of driver fatalities that are men under 26 years old (4 percent); sports cars have the highest (39 percent). So we suspect that differences in the behavior of their drivers account in large measure for why these two classes of vehicles pose such different risks to the people who operate them.
>
> Thomas P. Wenzel and Marc Ross

6.2 Automobile Safety Regulation

Cars are much, much safer today than in the past. Everyone knows that seat belts, airbags, and anti-lock brakes have made cars safer. The future holds great promise: guidance and avoidance systems, fly-by-wire technology that will eliminate steering columns, and much more; culminating in self-driving vehicles that communicate with each other.

But cars remain dangerous, both to vehicle occupants and others, such as cyclists and pedestrians. The United States uses the Fatal Accident Reporting System (FARS) to gather information about every motor vehicle crash in which someone dies. Such an event requires sending detailed information to FARS. Police record many variables, including time, weather conditions, demographic data, and whether drugs or alcohol were involved.

STEP To see the data, open the Excel workbook *SafetyRegulation.xls* and read the *Intro* sheet, then go to the *Data* sheet.

You can see that 36,650 people died in 2018 in a traffic accident. About half of the fatalities were drivers, almost 5,000 were motorcyclists, and 7,354 were non-motorists.

While FARS has data on the total number of deaths back to 1994 (36,254), simply comparing total fatalities over time is not a good way to measure driving safety. Under *Other National Statistics*, the data show that, year after year, there are many more people driving cars many more miles. So, we need to adjust the total number of fatalities to account for these increases.

We need a *fatality rate*, not the total number of fatalities. By dividing total deaths by the number of miles traveled, we get a measure of fatalities per mile traveled. This results in a tiny number so, to make it easier to read, the fatality rate is reported per 100 million miles traveled.

Adjusting with miles traveled is not the only way to create a fatality rate. The *Data* sheet shows rates based on population, registered vehicles, and licensed drivers. They all tell the same story.

Figure 6.4 shows the United States traffic fatality rate. The number of fatalities per 100 million miles traveled has fallen from 1.73 in 1994 to 1.17 in 2017, which is about a 30% decrease during this time period. That is welcome news.

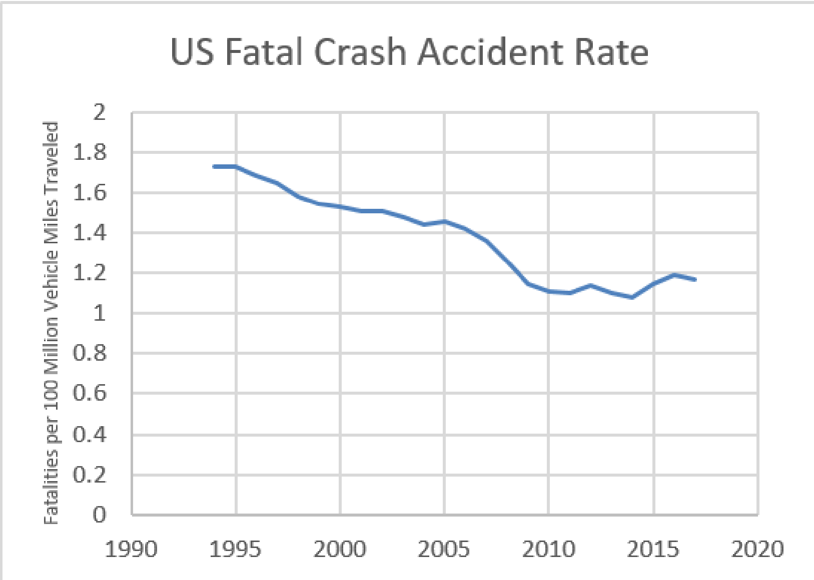


Figure 6.4: Traffic fatalities per 100 million miles traveled.
Source: SafetyRegulation.xls!Data
www-fars.nhtsa.dot.gov/

Less encouraging in Figure 6.4 is the leveling off since 2009 and the increase from 2014 to 2016. Distracted driving because of phone use and texting are suspected contributors.

The data in FARS only track fatalities and, thus, say nothing about nonfatal accidents. It turns out we are doing better here also—injury rates and severity of injury have also declined.

So, all is well? Actually, not exactly.

Although it may seem greedy, fatalities and injuries should have fallen by a lot more. We are doing better because fatal accident and injury rates have fallen, but we should be doing much, much better. After all, the car you drive today is much, much safer than a car from 20 or 30 years ago. If the vehicle you drive today is much safer than vehicles from 20 or 30 years ago, then fatal accident and injury rates should have fallen more to reflect these improvements. So, what is going on?

Economics can help answer this question. We will apply the remarkably flexible Theory of Consumer Behavior to driving a car. Any problem that can be framed as a choice given a set of exogenous variables can be analyzed via the economic approach. There are certainly choices to be made while driving: what route to take, how fast to drive, and what car to drive are three of many choices drivers make. We will focus on a subset of choices that involve how carefully to drive.

Theoretical Intuition

The key article that spawned a great deal of further work in this area was written in 1975 by University of Chicago economist Sam Peltzman. The abstract for "The Effects of Automobile Safety Regulation" (p. 677) says,

> Technological studies imply that annual highway deaths would be 20 percent greater without legally mandated installation of various safety devices on automobiles. However, this literature ignores offsetting effects of nonregulatory demand for safety and driver response to the devices. This article indicates that these offsets are virtually complete, so that regulation has not decreased highway deaths. Time-series (but not cross-section) data imply some saving of auto occupants' lives at the expense of more pedestrian deaths and more nonfatal accidents, a pattern consistent with optimal driver response to regulation.

This requires some translation. By technological studies, Peltzman is referring to estimates by engineers that are based on extrapolation. Cars with seat belts, airbags, anti-lock brakes, and so on are assumed to be driven in exactly the same way as cars without these safety features. This will give maximum bang for our safety buck.

Economics, however, tells us that we won't get this maximum return on improved safety features because there is a driver response to being in a safer car. By offsetting effects, Peltzman means that the gains from the safety devices are countered, offset, by more aggressive driving.

Peltzman's key insight, which separates an economist from the way an engineer considers the problem, is to incorporate driver response. He says on page 681:

> The typical driver may thus be thought of as facing a choice, not unlike that between leisure and money income, involving the probability of death from accident and what for convenience I will call "driving intensity." More speed, thrills, etc., can be obtained only by forgoing some safety.

This claim sounds rather outrageous at first. Do I suddenly turn into an Indy 500 race car driver upon hearing that my car has airbags? No, but consider some practical examples in your own life:

- Do you drive differently in the rain or snow than on a clear day?
- Do speed bumps, if you can't swerve around them, lead you to reduce your speed?
- Would you drive faster on a road in Montana with no cars for miles around versus on the Dan Ryan Expressway in Chicago? In which case, Montana or Chicago (presuming you are actually moving on the Dan Ryan), would you pay more attention to the road and your driving?
- If your car had some magic repulsion system that prevented you from hitting another car (we almost have this), would you drive faster and more aggressively?

Economists believe that agents change their behavior to find a new optimal solution when conditions change. In fact, many believe this is the hallmark of economics as a discipline. Many non-economists either do not believe this or are not aware of how this affects us in many different ways.

If you do not believe that safer cars lead to more aggressive driving, consider the converse: Do more dangerous cars lead to more careful driving? Here is how Steven Landsburg puts it:

> If the seat belts were removed from your car, wouldn't you be more cautious in driving? Carrying this observation to the extreme, Armen Alchian of the University of California at Los Angeles has suggested a way to bring about a major reduction in the accident rate: Require every car to have a spear mounted on the steering wheel, pointing directly at the driver's heart. Alchian confidently predicts that we would see a lot less tailgating. (Landsburg, p. 5)

The idea at work here is only obvious once you are made aware of it. Consider the tax on cars over $30,000 passed by Congress in 1990. By adding a 10% tax to such luxury cars, staffers computed that the government would earn 10% of the sales revenue (price x quantity) generated by the number of luxury cars sold the year before the tax was imposed. They were sadly mistaken. Why?

People bought fewer luxury cars! This is a response to a changed environment. You cannot take for granted that everyone will keep doing the same thing when there is a shock.

This idea has far-reaching application. Consider, for example, its relevance to the field of macroeconomics. Robert Lucas won the Nobel Prize in Economics in 1995. His citation reads, "for having developed and applied the hypothesis of rational expectations, and thereby having transformed macroeconomic analysis and deepened our understanding of economic policy." (See www.nobelprize.org/prizes/economic-sciences/1995/press-release/)

What exactly did Lucas do to win the Nobel? One key contribution was pointing out that if policy makers fail to take into account how people will respond to a proposed new policy, then the projections of what will happen will be wrong. This is called the *Lucas Critique*.

The Lucas Critique is exactly what is happening in the case of safety features on cars. Economists argue that you should not assume that drivers are going to continue to behave in exactly the same way before and after the advent of automobile safety improvements.

What we need is a model of how drivers decide how to drive. The Theory of Consumer Behavior gives us that model. You know what will happen next: we will figure out the constraint. And after that? Preferences. That will be followed by the initial solution and, then, comparative statics. We will find the effect of safer cars on accident risk. This is the economic approach.

The Initial Solution

The driver chooses how *intensively* to drive, which means how aggressively to drive. Faster starts, not coming to a complete stop, changing lanes, and passing slower cars are all more intensive types of driving, as are searching for a song or talking on your phone while driving. More intensive driving saves time and it is more fun. Driving intensity is a good and more is better.

Unfortunately, it isn't free. As you drive more intensively, your chances of having an accident rise. No one wants to crash, damaging property and injuring themselves or others. Your accident risk, the probability that you have an accident, is a function of how you drive.

The driver chooses a combination of two variables, *Driving Intensity* and *Accident Risk*, that maximize utility, subject to the constraint.

The equation of the constraint ties the two choice variables together in a simple way.

$$DrivingIntensity = SafetyFeatures * AccidentRisk$$

Safety Features represents the exogenous variable, safety technology, and provides a relative price at which the driver can trade risk for intensity.

On the Initial line in Figure 6.5, the driver is forced to accept a great deal of additional *Accident Risk* for a little more *Driving Intensity* because the line is so flat.

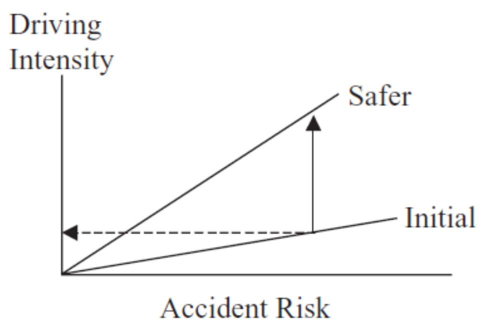


Figure 6.5: The driver's constraint.

When cars get safer, the constraint line gets steeper, rotating counterclockwise from the origin, as shown in Figure 6.5. There are two ways to understand the improvement made available by better safety technology. The horizontal, dashed arrow shows that you can get the same *Driving Intensity* at a much lower *Accident Risk*. You can also read the graph vertically. For a given *Accident Risk*, a safer car gives you a lot more *Driving Intensity* (follow the vertical, solid arrow).

Figure 6.5 shows that safer technology can be interpreted as a decrease in the price of *Driving Intensity*. It affects the graph just like a decrease in p_2 in the Standard Model.

The constraint is only half of the story. We need preferences to find out how a driver will decide to maximize satisfaction.

We use a Cobb-Douglas functional form to model the driver's preferences for *Accident Risk* (x_1) and *Driving Intensity* (x_2), subtracting *Accident Risk* from a constant so that increases in x_1 lead to less utility.

$$U(x_1, x_2) = (1 - x_1)^c x_2^d$$

Risk is measured between zero and 100 percent so $0 \leq x_1 \leq 1$. As x_1 increases in this interval, utility falls. The indifference curves will be upward sloping because x_1, *Accident Risk*, is a bad.

We can solve this model via numerical and analytical methods. We begin with Excel's Solver.

STEP Proceed to the *OptimalChoice* sheet.

The sheet shows the goal, endogenous variables, and exogenous variables. Initially, the driver is at 25%,0.25, which is a point on the budget line (because the constraint cell shows zero). We will use % notation for *Accident Risk* because it is a probability. The unrealistically high chances of an accident were chosen to maximize visibility on the graph. We use decimal points (such as 0.5) for the driving intensity variable, which we interpret as an index number on a scale from 0 to 1.

We know the opening point is feasible, but is it an optimal solution?

In previous Excel files, the graph is immediately displayed so you can instantly see if there is a tangency. The missing graph gives you a chance to exercise your analytical powers. Can you create a mental image of the chart even though it is not there? Remember, comparing the slope of the budget line to the MRS at any point tells us what is going on.

The slope is simply the *Safety Features* exogenous variable, which is +1. So now the graph looks like Figure 6.5 with a 45 degree line from the origin.

But what about the indifference curves? The MRS is minus the ratio of marginal utilities. With $c = d = 1$, we have

$$MRS = -\frac{\frac{dU}{dx_1}}{\frac{dU}{dx_2}} = -\frac{-x_2}{1-x_1} = \frac{x_2}{1-x_1}$$

We evaluate this expression at the chosen point, 25%, 0.25, and get

$$\frac{x_2}{1-x_1} = \frac{[0.25]}{1-[25\%]} = \frac{1}{3}$$

We immediately know the driver is not optimizing.

In addition, we know he can increase satisfaction by taking more risk and more intensity, traveling up the budget line because the indifference curve is flatter ($\frac{1}{3}$) than the budget line (+1) at the opening point of 25%,0.25.

Do you have a picture in your mind's eye of this situation? Think about it. Remember, the MRS is smaller than the slope so the indifference curve has to be flatter where it cuts the line.

STEP When you are ready (after you have formed the mental picture of the situation), click the Show Chart button to see what is going on at the 25%,0.25 point.

The canonical graph (with a bad) appears and the cells below the chart show the slope and MRS at the chosen point.

STEP Next, run Excel's Solver to find the optimal solution.

With $c = d = 1$ and a *Safety Features* value of 1, it is not surprising that the optimal solution is at 50%,0.50. Of course, at this point, the slope = MRS.

To implement the analytical approach, the Lagrangean looks like this:

$$\max_{x_1,x_2,\lambda} L = (1-x_1)x_2 + \lambda(x_2 - Sx_1)$$

An exercise asks you to find the reduced form solution.

Comparative Statics

Suppose we get safer cars so the terms of trade between *Driving Intensity* and *Accident Risk* improve. What happens to the optimal solution?

STEP Change cell B16 to 2.

How does the engineer view the problem? To her, the driver keeps acting the same way, driving just like before. There will be a great gain in safety with much lower risk of an accident. This is shown by the left-pointing arrow in Figure 6.6. Intensity stays the same and risk falls by a great deal.

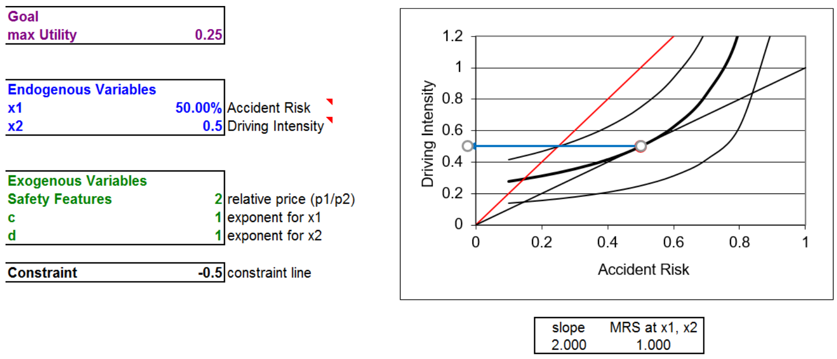


Figure 6.6: Improved safety features shock.
Source: SafetyRegulation.xls!OptimalChoice

For the engineer, because *Driving Intensity* remains constant, if it was 0.5, then improving Safety Features to 2 makes the accident risk fall to 25%. We simply travel horizontally along a given driving intensity to the new constraint.

The economist doesn't see it this way at all. She sees *Driving Intensity* as a choice variable and as the solution to an optimization problem. Change the parameters and you change the optimizing agent's behavior. It is clear from Figure 6.6 that the driver is not optimizing because the slope does not equal the MRS.

STEP With new safety technology rotating the constraint line, we must run Solver to find the new optimal solution.

The result is quite surprising. The *Accident Risk* has remained exactly the same! What is going on? In Peltzman's language, this is *completely offset-*

ting behavior. The optimal response to the safer car is to drive much more aggressively and this has completely offset the gain from the improved safety equipment.

How can this be? By decomposing the zero total effect on *Accident Risk* into its income and substitution effects, we can better understand this curious result.

Figure 6.7 shows what is happening. The improved safety features lower the price of driving intensity, so the driver buys more of it. On the y axis, the substitution and income effects work together to increase the driver's speed, lane changes, and other ways to drive more intensively. On the x axis, which measures risk taken while driving, the effects oppose each other, canceling each other out and leaving no gain in accident safety.

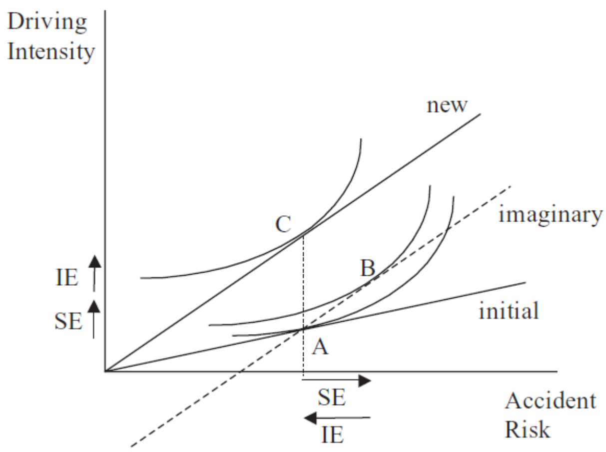


Figure 6.7: Income and substitution effects.

As driving intensity gets cheaper, the substitution effect (the move from A to B in Figure 6.7) leads the driver to choose more intensity and pay for it with more risk. The income effect leads the driver to buy yet more intensity and (because risk is a normal bad) less risk. The end result, for this utility function, is completely offsetting behavior.

Of course, this is not necessarily what we would see in the real world. We do not know how many drivers are represented by these preferences. The income effect for risk could outweigh the substitution effect, leaving point C to the left of A in Figure 6.7.

Theory alone cannot answer the question of what we will see in the real world. Empirical work in this area does confirm that offsetting behavior exists, but there is disagreement as to its extent.

An Economic Analysis of Driving

Choices abound when it comes to cars and driving. Should I take the highway or stay on a surface street? Change the oil now or wait a while longer? Pass this slow car or just take it easy and get there a few minutes later? Because there are choices, we can apply economics. This chapter focused on applying the Theory of Consumer Behavior to the choice of how intensively to drive. The agent is forced to trade off a bad (the risk of having an accident) for getting there faster and greater driving enjoyment.

Yes, teenagers make different choices than older drivers and everyone drives differently on a congested, icy road than on a sunny day with no traffic, but our comparative statics question focused on how improved automobile technology impacts the optimal way to drive.

Offsetting behavior is an application of the *Lucas Critique*: do not extrapolate. Instead, we should recognize that agents change their behavior when the environment changes. Theory cannot tell us how much offsetting behavior we will get. Only data and econometric analysis can tell us that.

Economists believe that we have not had as great a reduction in automobile fatalities and injuries as our much, much safer cars would enable because drivers have chosen to maximize satisfaction by trading some safety for driving intensity. Offsetting behavior explains why we aren't doing much, much better in traffic fatalities. But do not despair—we are maximizing satisfaction given our new technology.

Exercises

1. Use the equation that follows to solve for $x_1{}^*$ and $x_2{}^*$ in terms of S (safety features). Show your work.

$$\max_{x_1, x_2, \lambda} L = (1 - x_1)x_2 + \lambda(x_2 - Sx_1)$$

2. Use your reduced form solution to find the S elasticity of $x_1{}^*$ at $S = 1$. Show your work.

3. If the utility function was such that *Driving Intensity* was a Giffen good, describe where point C would be located on Figure 6.7.

4. If the utility function was such that *Driving Intensity* was a Giffen good, would this raise or lower traffic fatalities? Explain.

References

The epigraph is from page 125 of Thomas P. Wenzel and Marc Ross, “Safer Vehicles for People and the Planet,” *American Scientist*, Vol. 96, No. 2 (March–April, 2008), pp. 122–128, www.americanscientist.org/article/safer-vehicles-for-people-and-the-planet. They claim that the conventional wisdom that we need cars to be heavy to be safe is wrong. Heavier cars waste more fuel. How much more? “If a typical car could somehow drop 10% of its mass, its fuel economy would increase by anywhere from 3% to 8%. (The larger value applies if the size of the engine is also reduced to keep acceleration performance the same.)” (p. 124). The authors are not economists, but notice how they frame the result with percentage changes.

For an excellent review of empirical work on traffic safety, see *Traffic Safety* by Leonhard Evans, online at www.scienceservingsociety.com/.

This original idea is from Sam Peltzman, “The Effects of Automobile Safety Regulation,” *The Journal of Political Economy*, Vol. 83, No. 4 (August, 1975), pp. 677–726, www.jstor.org/stable/1830396

For a simple (no math or graphs) explanation of the idea behind offsetting behavior, see “The Power of Incentives: How Seat Belts Kill,” in Steven E. Landsburg, *The Armchair Economist* (New York: The Free Press, 1993).

Russell S. Sobel and Todd M. Nesbit point out that aggregated traffic fatality data is a poor way to test for a Peltzman effect. They find strong support for offsetting behavior from improved safety in professional auto racing in “Automobile Safety Regulation and the Incentive to Drive Recklessly: Evidence from NASCAR,” *Southern Economic Journal*, Vol. 74, No. 1 (Jul., 2007), pp. 71-84, www.jstor.org/stable/20111953.

Tom Vanderbilt’s *Traffic: Why We Drive the Way We Do (and What It Says About Us)* (New York: Alfred A. Knopf), 2008, touches on a variety of issues about cars and driving.

In the past it was futile to double the wages of an agricultural worker in Silesia who mowed a certain tract of land on a contract, in the hope of inducing him to increase his exertion. He would simply have reduced by half the work expended.

Max Weber

6.3 Labor Supply

We began the Theory of Consumer Behavior with the Standard Model where cash income (m) is given. The Endowment Model replaced given cash income with an initial endowment of two goods so the budget constraint became $p_1x_1 + p_2x_2 = p_1\omega_1 + p_2\omega_2$. We then focused on choices with bads—risky assets and accidents.

The application in this section is another example using a bad. As always, our eventual goal is comparative statics and elasticity. In this case, we will derive a supply curve for labor and concentrate on the wage elasticity of labor supply.

An innovation in this section is that the accompanying Excel workbook is less finished than usual. This enables you to practice implementing the model in Excel.

Setting Up the Problem

Instead of a mere consumer, the agent in this application is a consumer and worker.

Although an initial amount of non-labor income is assumed, total income can be increased by working. More hours at work means more income and greater consumption of goods and services. Consumption is good, but work is bad. Therein lies the problem.

Our consumer/worker can buy a single good, G, representing all consumer goods, at price p. Utility increases as she consumes more G.

The 24 hours in a day are divided into two types: work and leisure. The number of hours spent working in one day, H, is chosen by the agent. Earned income is simply wH,where w is the wage rate in \$/hr. Although work generates income, our agent does not like to work. H is a bad in the utility function.

With this background, we are ready to organize the information into the three areas that comprise an optimization problem:

1. *Goal*: maximize utility, which is a function of goods consumed, G, and work, H, where H is a bad.

2. *Endogenous variables*: G, the amount of goods consumed, and H, the number of hours worked.

3. *Exogenous variables*: p, the price of the composite good; w, the wage rate; m, unearned, non-labor income; and parameters in the utility function.

The solution to this constrained optimization problem is depicted on a graph with a budget constraint and set of indifference curves. We consider each of these elements separately and then combine them.

Budget Constraint

The budget constraint is $m + wH \geq pG$. This equation says that total income is composed of unearned income (m) and earned income (wH). The inequality means that the consumer/worker cannot spend more on goods and services (pG) than the total income available.

Because no time elapses in this optimization problem, there is no reason for the agent to save (i.e., spend less than available) and we can make the constraint a strict equality, $m+wH = pG$. This allows us to use the Lagrangean method to solve the problem analytically.

In terms of a graph, it is easy to see that we can write the constraint as the equation of a line (with G on the y axis and H on the x axis) by dividing by p:

$$m + wH = pG$$

$$G = \frac{m}{p} + \frac{wH}{p}$$

Suppose $w = \$10/\text{hr}$, $m = \$40$, and $p = \$1/\text{unit}$. What would the constraint look like?

STEP Open the Excel workbook *LaborSupply.xls* and read the *Intro* sheet, then go to the *YourConstraint* sheet.

Your task is to fill in the G column and create a chart of the constraint. There are three steps.

STEP Click on B12 and enter a formula equal to the equation for G. The cells w, p, and m are not named so you should use absolute references ($ in front of column letters and row numbers) to enable easy filling down of the formula.

When finished, the formula in B12 should look like this: = $B$4/$B$3 + ($B$2/$B$3)*A12.

STEP The next step is to fill down the formula.

STEP Finally, create a chart with H and G as the source data. Be sure to label the axes of your chart.

The chart is based on hour intervals of work, but fractions of hours are possible. Thus, your chart should be a scatter chart with points connected by lines.

STEP Click the [Reveal the Constraint] button to see a finished version of the budget constraint.

The agent is free to choose any point on the constraint. The y intercept, 40 (equal to $\frac{m}{p}$), yields a small value of consumption, but the agent does not have to work. Movement up the line yields more G, but requires more H.

Points to the northwest of the line are unattainable. For example, the consumer/worker cannot afford the 10,200 combination. Working 10 hours adds $100 to the $40 non-labor income. This is not enough to buy $200 worth of goods.

What shock would enable our consumer/worker to buy the 10,200 combination?

There are three possibilities, one for each exogenous variable in the constraint.

STEP From the *Constraint* sheet (click the Reveal the Constraint button from the *YourConstraint* sheet if needed), change the wage to 16 in B2.

The constraint rotates up, counterclockwise, with a steeper slope and the same intercept, and the combination 10,200 is now feasible, which is easily confirmed by looking at the chart and row 22.

Changes in wages, ceteris paribus, rotate the constraint around the unearned income intercept.

STEP Return the wage to 10 in B2 (the constraint returns to its initial position when you hit the Enter key) and set p (in B3) to 0.7.

Instead of raising the wage, we have made the composite good cheaper. As with a wage increase, this is welcome news since there are more consumption possibilities.

The constraint appears to simply rotate up again, but look more carefully at the chart and underlying data. The slope is steeper, but the intercept has also changed. The $40 of unearned income now buys a little more than 57 units of G. As before, it is easy to see that the combination 10,200 is now feasible.

Changes in price (p), ceteris paribus, rotate and shift the constraint.

STEP Return the price to 1 in B3 (the constraint returns to its initial position when you hit the Enter key) and set m (in B4) to 100.

This time, the constraint shifts vertically up. With $100 of unearned and $100 of earned income (from working 10 hours), the combination 10,200 is now feasible.

Changes in unearned income (m), ceteris paribus, shift the constraint.

Changes in w, p, and m affect the constraint. The initially unattainable combination of 10,200 can be made feasible by appropriately changing any of one of these three exogenous variables.

Preferences

In previous applications with bads, we used a Cobb-Douglas utility function and subtracted the bad from a constant. The same approach is adopted here.

Because the time period under consideration is a day, which has 24 hours, preferences can be represented by $U(H, G) = (24 - H)^c G^d$.

With $H = 0$, the agent gets the maximum value from the first term of the utility function, but remember that earned income will then be low and, therefore, G will be small.

Like the budget constraint, we need a visual representation of the utility function.

STEP Proceed to the *YourIndiffCurve* sheet to implement the utility function in Excel.

The sheet is unfinished. You need to fill in column B and draw a graph of the indifference curve. The indifference curve is initially based on $c = d = 1$ and a level of utility of 1960.

To fill in column B, you need to solve for the value of G that yields a utility level of 1960, given H. In other words, rewrite the utility function in terms of G, like this:

$$U(H, G) = (24 - H)^c G^d$$

$$G^d = \frac{U(H, G)}{(24 - H)^c}$$

$$G = \left[\frac{U(H, G)}{(24 - H)^c}\right]^{1/d}$$

STEP Use the expression above to enter a formula in B12 that computes the value of G necessary to produce a utility of 1960 when H = 2.

Your formula should look like this: = ($B$5/((24 - A12)\^($B$3)))\^(1/$B$4). It evaluates to a value of $G = 89.09$. This result makes sense because when $H = 2$, then $24 - 2 = 22$ and 22 x 89.09 (since $c = d = 1$) equals a utility value of 1960.

Notice again the use of absolute references.

STEP Fill down the formula and draw a chart with H and G as the source data. Label the axes.

Your chart is a graph of a single indifference curve. In fact, the entire quadrant is full of these upward sloping indifference curves and utility increases as you move in a northwesterly direction (taking less of the bad, H, and more of the good, G). This is the usual indifference map when we have a bad on the x axis.

Click the Reveal the Indiff Curve button to check your work or if you need help.

Finally, remember that changes in the exponents make the indifference curves flatter or steeper. A Q&A question explores this point.

Finding the Initial Optimal Solution

Having modeled the constraint and preferences, we are ready to find the initial solution.

The numerical approach is covered here; the analytical method is an exercise question.

STEP Proceed to the *YourOptimalChoice* sheet.

It is blank! You need to implement the problem in this sheet and run Solver to find the initial solution.

Organize the problem into the usual components: goal (maximize utility), endogenous variables (H and G), exogenous variables (w, p, m, c, and d), and a cell for the constraint.

The utility function is $U(H, G) = (24 - H)^c G^d$. The wage rate is \$10/hr, the price of G is \$1/unit, unearned income is \$40, and $c = d = 1$.

Click the Reveal the Optimal Choice button once you are finished or if you get stuck and need help.

Figure 6.8 shows the canonical graph of the initial optimal solution for the consumer/worker's constrained utility maximization problem.

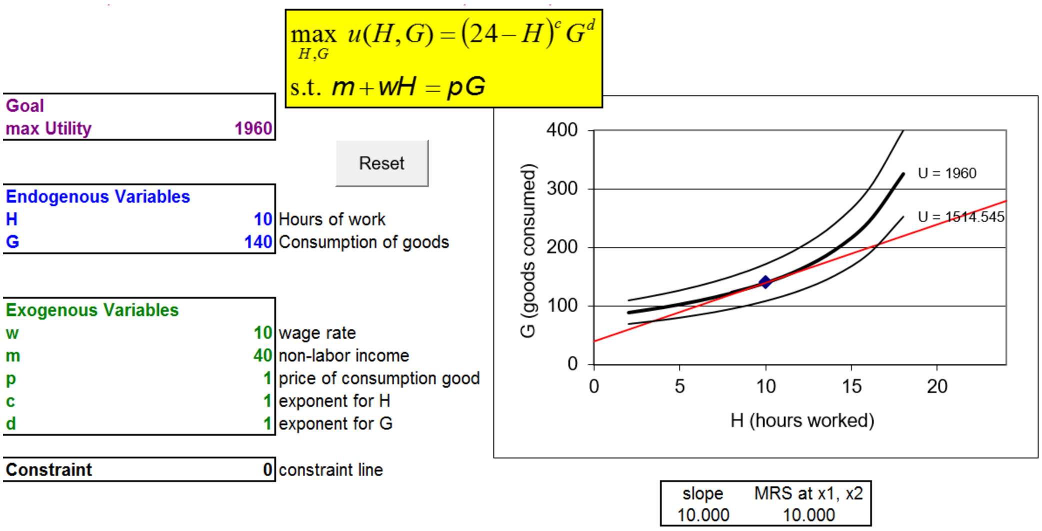


Figure 6.8: The initial solution.
Source: LaborSupply.xls!OptimalChoice

This consumer/worker maximizes utility by working 10 hours, thereby earning \$100 and then buying 140 units of G. There is no better solution. Traveling up or down the budget constraint is guaranteed to lower utility because the indifference curve is just touching the constraint at 10,140. The mathematical way of saying this is that the MRS $= \frac{w}{p}$ at 10,140.

Comparative Statics: Deriving Labor Supply

How does $H*$ respond as the wage rate changes, ceteris paribus? This comparative statics question yields the labor supply curve.

We concentrate on the numerical approach and leave the analytical method for an exercise question.

STEP Proceed to the *OptimalChoice* sheet (in the *YourOptimalChoice* sheet, click the [Reveal the Optimal Choice] button if needed). Use the Comparative Statics Wizard to pick a few points off of the labor supply curve. Make the size of the change in the wage rate 10 and apply the default five shocks.

Use the CSWiz data to compute the wage elasticity of hours worked from $w = \$10$ to \$20/hr. Create a graph the supply and inverse supply of labor curves.

STEP Proceed to the *CS1* sheet and scroll down (if needed) to check your work.

Notice the labor supply and inverse labor supply curves (scroll down if needed). The shape of the curve is intriguing. As wage rises, optimal H seems to level off—it continues to increase, but ever more slowly.

Notice also that the computed wage elasticity of labor supply from $w = 10$ to 20 in E14 is quite small at 0.1. This means that hours worked is unresponsive to changes in wages.

Labor supply has been extensively studied and extremely small elasticities with respect to wage are commonly found (see McClelland and Mok (2012) for a review of the literature). Income and substitution effects explain this result.

STEP Return to the *OptimalChoice* sheet and click the Reset button, then change the wage rate (in B16) from 10 to 20.

The budget constraint rotates up (counterclockwise) in the chart—a welcome change in consumption possibilities. The initial optimal solution, 10,140, is no longer optimal. The consumer/worker needs to re-optimize.

STEP Run Solver (with $w = 20$).

The new optimal solution is at $H = 11$. A 100% increase in the wage (from 10 to 20) has produced a total effect of a 1 hour, or 10%, increase in hours worked.

We can decompose this total effect into income and substitution effects by shifting down the budget line to cancel out the increased purchasing power of the wage increase. In other words, we need to draw in an imaginary, dashed line that goes through the initial solution, with a steeper slope caused by the higher wage.

We can use a modified version of the Income Adjuster Equation to determine the amount of income we need to take away. Recall that we determine how much income to change via $\Delta m = x_1 \Delta p_1$. In the labor supply model, x_1 is obviously H, and the price is now the wage, but we also need a sign change. An increase in the wage increases consumption possibilities in the labor supply model so we need a minus sign to show that wage increases must be offset by income decreases. Below is our modified Income Adjuster Equation with values substituted in:

$$\Delta m = (\Delta H^*)(-\Delta w)$$

$$\Delta m = (10)(-10) = -100$$

This says that we must lower unearned income by \$100 to cancel out the increased purchasing power from the \$10/hr wage increase.

STEP Confirm that $w = 20$ (in B16) and change m to -60 (in B17).

Notice that the budget line goes through the initial combination, 10,140. The line is not dashed, but it should be. Remember that this budget line does not actually exist. No one is going to take \$100 from the agent. We are doing this to decompose the total effect of the wage increase into the income and substitution effects.

STEP Run Solver with $w = 20$ and $m = -60$.

$H^* = 13.5$ hours of work and Figure 6.9 shows the three effects.

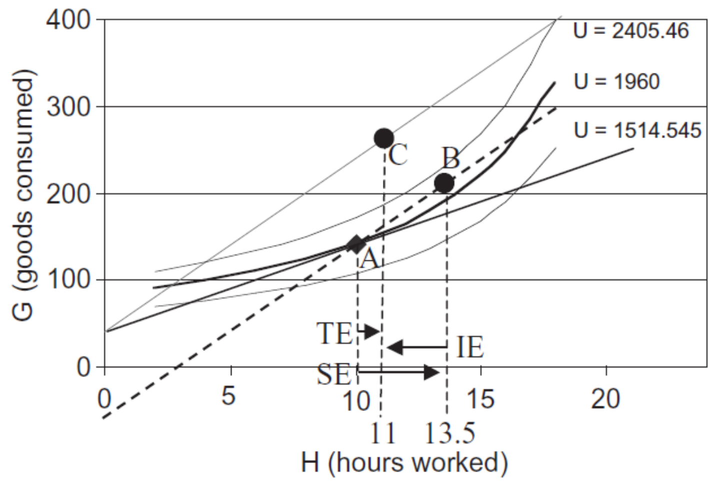


Figure 6.9: Total, income, and substitution effects.
Source: LaborSupply.xls!OptimalChoice

The substitution effect is +3.5, the movement from $H = 10$ (the initial optimal solution) to 13.5 (the optimal solution with the higher wage, but lower m). It is the horizontal movement from point A to B.

The income effect is −2.5, the movement from $H = 13.5$ (point B) to $H = 11$ (point C). The negative sign is important. It says that when income rises, the agent buys less of the bad.

The total effect is, of course, the observed movement from point A to point C, a 1-hour increase in hours worked. This is what would actually be observed as the wage rose from \$10/hr to \$20/hr.

Figure 6.9 makes clear why the response of hours worked to a wage increase is inelastic—the income and substitution effects are working against each other. The fact that the relative price of goods for an hour of work is cheaper drives the agent to work and consume more (this is the substitution effect, from A to B). But the increase in purchasing power encourages the agent to work less (from B to C, the income effect). The total effect on hours worked is small when the two effects are added together.

In fact, the income and substitution effects can explain an even more curious phenomenon that has been observed in the real world—hours worked actually falling as wage rises. Figure 6.10 shows the underlying graph and derived labor supply curve for an unknown utility function. Unlike the labor supply derived from the Cobb-Douglas utility function, which was always positively sloped, the labor supply curve in Figure 6.10 is said to be *backward bending*. At low wages, increases in wage lead to more hours worked (such as from point 1 to 2), but the supply curve becomes negatively sloped when wages rise from point 2 to 3.

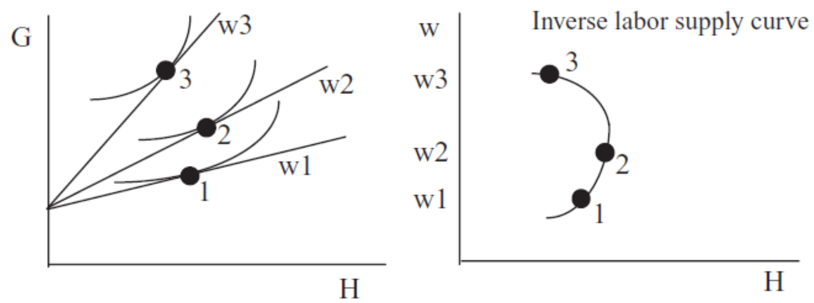


Figure 6.10: A backward bending supply curve.

We have already seen that the small wage elasticity from point 1 to 2 is caused by the income effect's working against the substitution effect. The same explanation underlies the negative response in hours worked as wages rise from point 2 to 3. In this case, not only does the income effect oppose the substitution effect, it actually swamps it.

Figure 6.11 shows what happens when we are on the backward bending portion of the labor supply curve. The substitution effect always induces more hours worked as wages rise. This is the movement from A to B. The income effect, however, counters some of this increase in hours worked. We can afford to work less (from B to C) because the wage is higher. When we are on the backward bending portion of the labor supply curve, the income effect actually overcomes the substitution effect so that the total effect (A to C) is a reduction in hours worked as the wage rises. In Figure 6.11, any point C to the left of A yields a point on the backward bending portion of the labor supply curve.

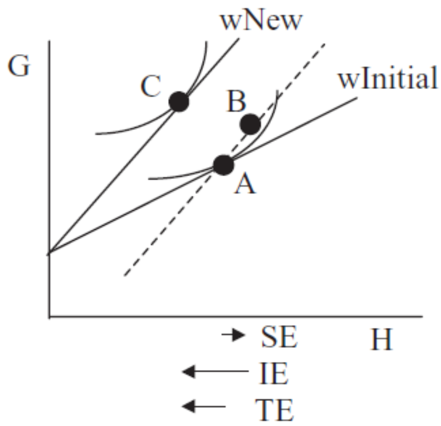


Figure 6.11: Income and substitution effects when H^* falls as w rises.

Wage rises and I work less sounds just about as weird as price rises and I buy more. Is this Giffen behavior?

No because the wage change is not an own price effect. Figure 6.12 shows p_1 and p_2 changes in the Standard Model where two goods are purchased given fixed income. On the left, the change in p_1 produces an own effect on x_1 and a cross effect on x_2.If x_1 rises as p_1 rises, then x_1 is Giffen. If x_2 rises as p_1 rises (notice the cross effect), however, that does not make x_2 a Giffen good. We use the cross effect to say that the goods are substitutes (instead of complements). To determine whether x_2 is Giffen, we have to use the graph on the right of Figure 6.12. If x_2 rises as p_2 rises (notice the own effect), then x_2 is Giffen. In other words, we need an own price change to determine Giffenness.

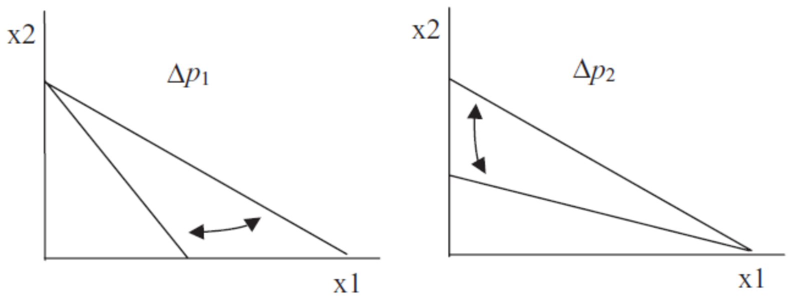


Figure 6.12: Understanding own and cross effects.

Figure 6.12 makes clear that a change in the wage in the labor supply optimization problem is like a change in the price of x_2 in the Standard Model. The wage change is like the graph on the right, with an upward sloping budget constraint. The rotation is around a fixed value—the x intercept in the Standard Model and unearned income in the labor supply model. Thus, the change in wage is an own price effect for G (on the y axis) and a cross price effect for H (on the x axis).

Because a change in the wage exerts a cross effect on hours worked, we cannot say anything about Giffenness for hours worked. We could, however, say that G was Giffen if it fell when wage rose. That would really be weird. Look at the figures of income and substitution effects in this chapter and you will never find a final point C that lies below an initial point A. In fact, leisure (work's counterpart) is usually treated as a normal good: higher income leads to more leisure (and less work).

Deriving the Labor Supply Curve

Labor Economics is a major field within Economics. As a course, it is usually offered as an upper-level elective, with Intermediate Microeconomics as a prerequisite. Labor supply and demand are fundamental concepts. The former is based on a model in which work is a bad (the opposite of leisure, which is a good) and a consumer/worker maximizes satisfaction subject to a budget constraint.

By changing the wage, ceteris paribus, we can derive a labor supply curve. Economists are well aware that labor supply is often quite insensitive to changes in wages. This is explained by the opposing substitution and income effects. The backward bending portion of the labor supply curve is observed when the income effect swamps the substitution effect. This is not Giffen behavior, however, because we are dealing with a cross (not own) price effect.

Exercises

1. Use the Lagrangean method to solve this consumer/worker's constrained optimization problem:

$$\max_{H,G} U = (24 - H)G$$
$$\text{s.t. } 40 + wH = G$$

   Show all of your work.

2. Do your results for H^* and G^* agree with the numerical approach in the text? Is this surprising?

3. Using the Comparative Statics Wizard, the wage elasticity of labor supply from \$10/hr to \$20/hr is 0.1. Use your reduced form solution for H^* to find the wage elasticity of labor supply at $w = \$10$/hr. Show your work.

4. Does your point wage elasticity from the previous question equal 0.1 (the wage elasticity based on a \$10 wage increase)? Why or why not?

5. Whether the labor supply curve is upward sloping or backward bending has nothing to do with the Giffenness of work. If labor supply is positively sloped, G and H are substitutes or complements, but which one? Draw a graph that helps you explain your answer.

References

The epigraph comes from page 355 of Max Weber's classic, *General Economic History*, originally published in German in 1923 and translated to English by Frank H. Knight in 1927. If you are unfamiliar with Weber (pronounced vay-ber), he was interested in the way capitalism changed people's minds and values, especially how it made people more rational and calculating.

With respect to labor supply, the consumer/worker's goals and attitudes are a critical issue. In this chapter, labor supply was derived as the solution to an optimization problem. The agent, however, might not be an optimizer, but a target earner, working only enough hours to make a certain amount of money. If wages double, hours worked are cut in half. If everyone was a target earner, the typical way to attract more workers—pay more—would not work.

Consider this abstract from Henry Farber's 2003 NBER working paper, "Is Tomorrow Another Day? The Labor Supply of New York Cab Drivers":

> I model the labor supply of taxi drivers as the result of optimization based on an inter-temporal utility function. Since income effects in response to temporary fluctuations in daily earnings opportunities are likely to be small, cumulative hours will be much more important than cumulative income in the decision to stop work on a given day. However, if these income effects are large due to very high discount and interest rates, then labor supply functions could be backward bending, and, in the extreme case where the wage elasticity of daily labor supply is minus one, drivers could be target earners. Indeed, Camerer, Babcock, Lowenstein, and Thaler (1997) and Chou (2000) find that the daily wage elasticity of labor supply of New York City cab drivers is substantially negative and conclude that it is likely that cab drivers are target earners. I conclude from my empirical analysis, based on new data, of the stopping behavior of New York City cab drivers that, when accounting for earnings opportunities in a reduced form with measures of clock hours, day of the week, weather, and geographic location, cumulative hours worked on the shift is a primary determinant of the likelihood of stopping work while cumulative income earned on the shift is weakly related, at best, to the likelihood of stopping work. This is consistent with there being inter-temporal substitution and inconsistent with the hypothesis that taxi drivers are target earners.

See http://www.nber.org/papers/w9706.

Google Scholar has tens of thousands of papers on Uber and how drivers decide how many hours to work.

Robert McClelland and Shannon Mok's 2012 working paper that summarizes the wage elasticity literature, "A Review of Recent Research on Labor Supply Elasticities," is freely available from the Congressional Budget Office at www.cbo.gov/publication/43675. A remarkable finding is that men's much larger substitution effect than women's has all but disappeared so that men and women today respond similarly to wage shocks.

Chapter 7

Search Theory

Fixed Sample Search

Sequential Search

Price dispersion is a manifestation—and, indeed, it is the measure—of ignorance in the market.

George Stigler

7.1 Fixed Sample Search

The Theory of Consumer Behavior is based on the idea that buyers choose how much to buy based on preferences, income, and given prices. We know, however, that buyers do not face a single price—there is a distribution of prices and sellers change their prices frequently.

You would think consumers would be unable to choose in such an environment. After all, how can they know the budget constraint without prices? The answer is that they search or, in other words, they go shopping, and then use the lowest prices found to solve their constrained utility maximization problem.

Search Theory is an application of the economic approach to the problem of how long to shop in a world of many prices. Search is a productive activity because it enables one to find lower prices, but it is costly. One can search too little, ending up paying a high price, or search too much—spending hours to find a price that is a few pennies lower does not make much sense.

This chapter introduces the consumer's search optimization problem and is based on the idea that consumers decide in advance how many price quotes to obtain, according to an optimal search rule. This type of search procedure is known as a *fixed sample search*.

Describing the Search Optimization Problem

We assume that consumers do not know the prices charged by each firm. We simplify the problem by assuming that the product in different stores is identical (i.e., homogeneous) so the consumer just wants to buy at the lowest price. Unfortunately, finding that lowest price is costly so the buyer has to decide how long to search.

STEP Open the *FixedSampleSearch.xls* workbook and read the *Intro* sheet, then proceed to the *Setup* sheet.

The first task is to create the distribution of prices faced by the consumer. We assume that prices remain fixed during the search process.

STEP Click the Create the Population button.

You will be asked a series of questions that will establish the prices charged by all of the sellers. This is the population. The idea is that consumer will sample (draw) from the population. This is shopping.

STEP Hit OK when asked the number of stores selling the product to accept the default number of 1000 (no comma separator when entering numbers in Excel). Choose *Uniform* for the distribution and then press OK to accept 5 when prompted for the number of stores. Accept the default values of 0 and 1 for the minimum and maximum prices.

After you hit Enter, you will see a column of red numbers in column A that represent the prices charged by each of the 1,000 stores selling the product. The consumer knows that stores charge different prices, but cannot immediately see each individual store price. They cannot see the lowest and highest price stores in cells B2 and B3.

STEP Scroll down to see the prices charged at each store and confirm that the minimum price store, displayed in cell F2, is correct.

It is difficult to see by simply scrolling down and looking at the prices, but the uniform distribution you used means that prices are scattered equally from zero to one. The normal distribution, on the other hand, would concentrate prices near the average, with fewer low and high prices (like a bell-shaped curve). The log-normal is the most realistic of the three—prices have a long right-hand tail (with a few stores charging very high prices). The primary advantage of the uniform distribution is that it is the easiest to work with analytically.

Figure 7.1 shows a histogram of 1,000 prices from U[0,1]. This notation means that we include the endpoints so we have a uniform distribution with a zero minimum and a maximum of one (giving an average of 0.5 and an SD of 0.2887).

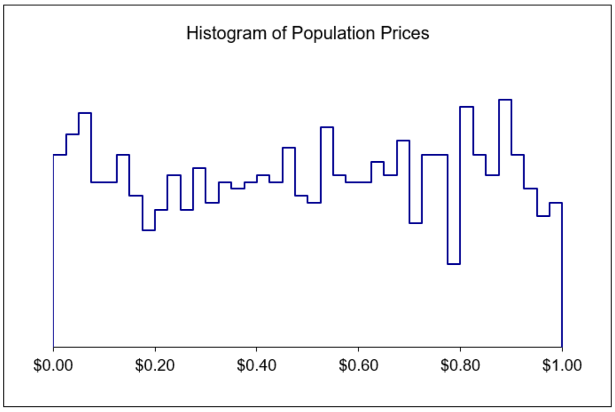


Figure 7.1: An example uniform distribution of prices.

The prices are not exactly evenly distributed on the interval from zero to one. They are drawn from a uniform distribution on the interval 0 to 1, but each realization of 1,000 prices deviates from a purely rectangular distribution due to randomness in sampling from the uniform distribution. The more stores you include in the population, the closer Figure 7.1 will get to a smooth, rectangular distribution. You can see a histogram of your population prices by scrolling over to column AA of the *Setup* sheet.

Consumers know the distribution of prices, but they do not know which firm is charging which price, so they cannot immediately go to the firm that has the lowest price. Instead, the fixed sample search model says that the consumer chooses a number of prices to sample (which you set as 5) and then chooses the lowest of the observed prices.

STEP Click the Draw a Sample One Price at a Time button. A price will appear in the sample column, and a pop-up box tells you where that price came from. Hit OK each time the display comes up. You will hit OK five times because you chose to sample from five stores.

The consumer chooses among the 1,000 stores randomly and ends up with five observed prices. Column L reports the sample average price, the SD of the sampled prices, and the minimum price in the sample (in cell L7). The consumer will purchase the product at the minimum price observed in the sample.

Why doesn't the consumer visit every store and then pick the lowest price? Because it is costly to obtain price information, as shown in cell L11. Each shopping trip (to collect a price) costs 4 cents. To sample all 1,000 stores would cost the consumer an exorbitant $40. On average, the consumer would pay $0.54 (the average of the price distribution plus the cost of obtaining one price) by buying the product at the very first store visited. Clearly, it is better to buy immediately, $n = 1$, than to sample every store, $n = 1{,}000$, but what about other fixed sample sizes? How much will the consumer pay, on average, when sampling five stores?

STEP Hit the [Draw a Full Sample] button repeatedly to draw more samples of size five. Keep your eye on the total price paid in cell L22.

Every time you get a new sample, you get a new total price (composed of the minimum price in sample plus 20 cents). There is no doubt about it—the total price the consumer ends up paying is a random variable. This makes this problem difficult because we need to figure out what the consumer can expect to pay usually or typically. We want to know the average total price. The next section shows how.

Monte Carlo Simulation

The plan is to alter the spreadsheet so a new sample can be drawn simply by recalculating the sheet, which is done by hitting the F9 key. We can then install the Monte Carlo simulation add-in and use it to repeatedly draw new samples, tracking the lowest price in each sample.

STEP Select cell range J2:J6. You should have five cells highlighted. In the formula bar, enter the following formula:

=DRAWSAMPLEARRAY()

and then press *Ctrl* + *Shift* + *Enter* (hold down and continuing holding down the *Ctrl* key, then hold down and continue holding down the *Shift* key, and then hit the *Enter* key). Your sample of five prices will appear in the sample column.

After you select the cells, do not simply hit the *Enter* key. This will put the formula only in the first cell. You want the formula in all five cells that you selected. You have to press *Ctrl* + *Shift* + *Enter* simultaneously.

You have used an *array function* (built into the workbook) that spans the five cells you selected. You cannot individually edit the cells. If you mistakenly try to do so and get stuck, hit the *esc* (escape) key to return to the spreadsheet.

When using this array function, it may display *#VALUE*. Simply hit the F9 key when this happens to refresh the function. If that does not work, recreate the population.

When using the DRAWSAMPLEARRAY() function, you must be sure to set the number of draws in cell C15 to correspond to the number of cells selected and used by the function. If there is a discrepancy, a warning will be displayed.

STEP Hit F9 a few times and keep your eye on cells L7, the minimum price, and L22, the total price paid.

These cells update each time you hit F9. A new sample of five prices is drawn and the minimum price and total price paid are recalculated for the new sample.

The DRAWSAMPLEARRAY() function enables Excel to display the minimum (best) price random variable, but we need to figure out the average minimum price when five price quotes are obtained. This can be done by repeatedly resampling and tracking each outcome – this is called Monte Carlo simulation.

STEP Install the Monte Carlo simulation Excel add-in, *MCSim.xla*, available freely from www3.wabash.edu/econometrics and the *MicroExcel* archive (in the same folder as the Excel workbook for this section). Full documentation is available at this web site. This powerful add-in enables sophisticated simulations with the click of a button.

Remember that installing an add-in requires use of the Add-ins Manager. Do not simply open the *MCSim.xla* file.

Once installed, you can use the add-in to determine the average minimum price and total price paid for the product when five prices are sampled.

STEP Run the Monte Carlo simulation add-in on cells L7 and L22 with 10,000 repetitions.

Your MCSim add-in dialog box should look like Figure 7.2. Click the Proceed button to run the simulation.

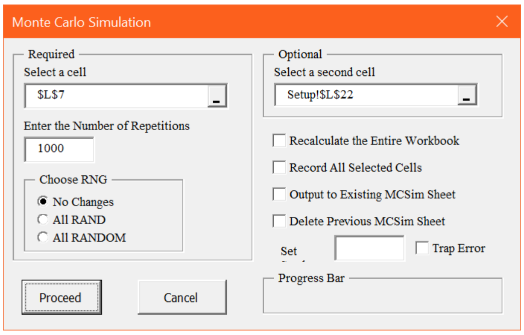


Figure 7.2: Configuring the MCSim dialog box.

Your simulation results will look something like Figure 7.3, but your results will be slightly different. The average of the minimum price distribution should be near 0.17 (1/6). Thus, the consumer will usually pay around $0.37 (adding the 20 cents in search cost) for the product. The total price paid is a shifted version of the best price.

BestPrice		TotalPrice		Notes
Average	0.171	Average	0.371	
SD	0.1410	SD	0.1410	
Max	0.827	Max	1.027	
Min	0.001	Min	0.201	

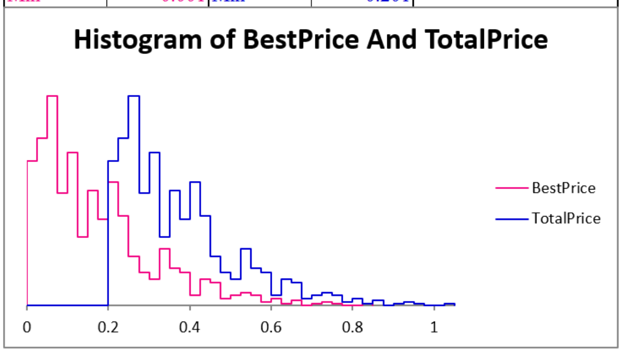


Figure 7.3: Monte Carlo simulation results with $n = 5$.
Source: FixedSampleSearch.xls!MCSim

So now we know that the consumer can expect to pay about $0.37 when searching five stores. This is better than buying at the first store visited, which was $0.54. Compared to the buying at the first store, the expected marginal gain of shopping at five stores, in terms of a lower expected minimum price, is \$0.50 − \$0.17 = \$0.33. The additional cost of searching for five prices instead of one is \$0.16. The additional benefit is greater than the additional cost is another way to know that five stores is better than one store.

But we want to know more than just that searching five stores is better than buying at the first store; we want to find the best sample size—the one that gives the lowest total price paid.

STEP Hit the Clear the Sample button. Change the number of draws in cell C15 to 10. Select cell range J2:J11 and then type in the formula bar: =DRAWSAMPLEARRAY(). Then press the *Ctrl + Shift + Enter* combination to input the array formula. Your sample of 10 prices will appear in column J.

Hit F9 a few times and watch what happens to cell L7, the minimum price. It bounces, but with 10 prices instead of five, it bounces around a different, lower mean.

STEP To find the typical price the consumer can expect to pay, run a Monte Carlo simulation of the minimum and total price when 10 stores are visited.

Figure 7.4 shows the exact average best price and average total price as a function of the sample size for the U[0,1] price distribution. Your simulation results for the best price for $n = 10$ should be close to \$0.0909.

Sample Size	Average Best Price	Search Cost	Total Price Paid
1	$ 0.5000	$ 0.04	$ 0.54
2	$ 0.3333	$ 0.08	$ 0.41
3	$ 0.2500	$ 0.12	$ 0.37
4	$ 0.2000	$ 0.16	$ 0.36
5	$ 0.1667	$ 0.20	$ 0.37
6	$ 0.1429	$ 0.24	$ 0.38
7	$ 0.1250	$ 0.28	$ 0.41
8	$ 0.1111	$ 0.32	$ 0.43
9	$ 0.1000	$ 0.36	$ 0.46
10	$ 0.0909	$ 0.40	$ 0.49

Figure 7.4: Optimal Search with a Uniform Distribution on the interval [0,1].
Source: FixedSampleSearch.xls!Summary

The typical \$0.0909 best price when 10 prices are obtained is lower than when we shopped at five stores, but notice that it isn't worth it. The cost of obtaining 10 prices (\$0.40) is so high that the total price paid is higher than getting just five prices. In fact, getting four prices is the optimal sample size.

Analytical Methods

The optimal search optimization problem can be solved via analytical methods. For the uniform price distribution on the interval from zero to one, the average minimum price in the consumers' hands after visiting n firms is

$$AverageP_{min} = \frac{1}{n+1}$$

The equation for the average minimum price shows that it is decreasing as n rises and it does so at a decreasing rate. In other words, there are diminishing returns to searching for low prices.

The consumer's optimization problem is to minimize the expected total cost of acquiring the product, where $P(n)$ represents the expected minimum price that we know is a function how many prices are collected:

$$\min_n TC = P(n)q + cn$$

We also know that for U[0,1], $P(n) = \frac{1}{n+1}$ so we have:

$$\min_n TC = \frac{1}{n+1}q + cn$$

To find optimal n, take the derivative with respect to n and set it equal to zero:

$$\frac{dTC}{dn} = -\frac{1}{(n+1)^2}q + c = 0$$

$$\frac{1}{(n+1)^2}q = c$$

This equimarginal condition says that the optimal sample size is found where marginal savings from additional search equals marginal cost. As long as the savings from searching an additional store exceeds the cost of collecting one more price, the consumer will continue to search. The marginal savings is just the drop in the expected price, times the number of units that the consumer wants to purchase.

From the equimarginal condition, we can solve for optimal n to get a reduced form solution.

$$\frac{1}{(n+1)^2}q = c \rightarrow q = c(n+1)^2 \rightarrow \sqrt{\frac{q}{c}} = n+1 \rightarrow \sqrt{\frac{q}{c}} - 1 = n*$$

With $q = 1$ and $c = \$0.04$, we have the same solution we found earlier:

$$n* = \sqrt{\frac{q}{c}} - 1 = \sqrt{\frac{1}{0.04}} - 1 = 4$$

Comparative Statics

The reduced form expression makes comparative statics analysis straightforward. It is obvious that higher c, search cost, leads to lower optimal sample size, as shown in Figure 7.5.

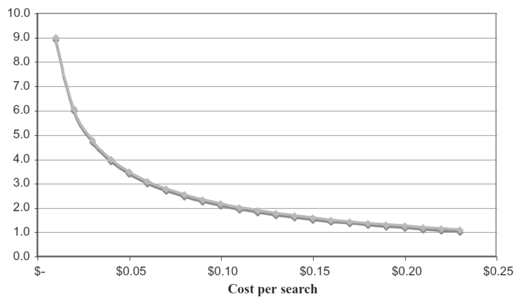


Figure 7.5: Optimal search with changing search cost for $q = 1$.

Search cost is not the same for each consumer. Time is an important element of search cost. Those with more valuable time and, therefore, higher search cost will optimize by obtaining fewer price quotes.

The availability of information is another component of search cost. Informational advertising is how firms let consumers know where they are and what prices are being charged. We can model this type of advertising as a decrease in search costs—today, all the consumer has to do is go online to see what prices are being offered. Search costs are still positive (consumers do not know, for example, whether all firms advertise or just some), but lower than without advertising. Consumers obtain the product for a lower total price when advertising lowers search costs.

If we allow for multiple purchases, that is, a value of $q > 1$, then the returns to search increase and, other things equal, the optimal number of searches increases. The effect of increasing q on the relationship between the cost of search and the optimal number of searches is shown in Figure 7.6.

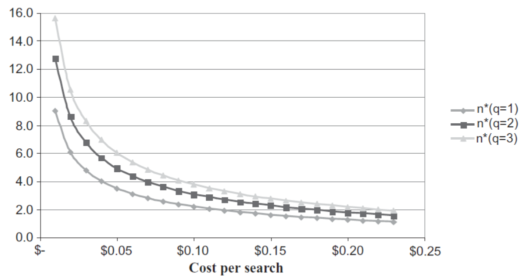


Figure 7.6: Optimal search with changing n and q.

For example, the driver of an 18-wheel truck that carries two 200-gallon diesel tanks is going to search more than someone looking to fill her car with gas. But this example leads to the next chapter, where we introduce a different search model.

Results of Fixed Sample Search

Incomplete price information leads to an entirely new optimization problem. Because consumers will not search every store, since that is too expensive, we see price dispersion. This is a major result of search theory and it deserves further explanation.

You would think that competition would tend to make prices of the same product equal. This is known as the *Law of One Price*. But this only applies to a world where consumers can costlessly gather prices.

In other words, the Law of One Price will fail to hold whenever it is costly to collect price data. This is true in the real world, where some consumers will end up paying higher prices than others because the minimum price in their particular information set is different than the minimum price in another consumer's set.

Because lower search costs induce more search, a reduction in search costs would have the effect of reducing (but not eliminating) price dispersion. Because optimizing consumers will choose not to canvass every store for prices as long as search is costly, price dispersion will exist. This is the key result of the fixed sample search model.

Economists have been interested in search theory for decades. The internet promised a big decrease in search cost and it may well have delivered that, but more recently, technology has really upended search theory. Today, your online search behavior is monitored and your clicks influence the prices you see.

The next level search models do not treat the population of prices as given and do not allow the consumer to randomly sample without changing the price distribution. Consumers still have an optimization problem to solve, but so do firms.

Exercises

Suppose the price distribution of 1,000 firms is uniform, with an average price of \$50 and an SD of \$28.87. Search cost, c, is \$1 per price.

1. On what interval (from the minimum to the maximum) are prices equally likely to fall?

2. Implement this problem in the *Setup* sheet and run a Monte Carlo simulation with a sample size of 20. Take a picture of your results (like Figure 7.3) and paste it in a Word document. What is good about obtaining 20 prices? What is bad?

3. Use the equation for the average minimum price as a function of n for this distribution, $AverageP_{min} = \frac{100}{n+1}$, to find the optimal sample size. Show your work.

4. Find the c elasticity of n at $q = c = 1$. Show your work.

References

The epigraph is from page 214 of George J. Stigler, "The Economics of Information," The *Journal of Political Economy*, Vol. 69, No. 3 (June, 1961), pp. 213–225, www.jstor.org/stable/1829263. This paper is recognized as the beginning of the economics of search.

Stigler was trying to explain price dispersion, but search theory has expanded far beyond this and is especially important in Labor economics. A consumer shopping for a low price product is the same as a worker looking for a high wage job or a firm seeking a high quality employee. See Richard Rogerson,

Robert Shimer, and Randall Wright, “Search-Theoretic Models of the Labor Market: A Survey,” *Journal of Economic Literature*, Vol. 43, No. 4 (December, 2005), pp. 959–988, www.jstor.org/stable/4129380.

Job offers are independent random selections from the distribution of wages. These offers occur periodically and are either accepted or rejected. Under these conditions it is easy to show that the optimal policy for the job searcher is to reject all offers below a single critical number and to accept any offer above this critical number.

J. J. McCall

7.2 Sequential Search

We introduced Search Theory with a Fixed Sample Search Model. A consumer samples from the population of stores and gets a list of n prices for a product, then chooses the minimum price. The bigger n, the lower the minimum price in the list, but the price paid to obtain the price quotes increases as n rises. The consumer has to decide how many prices to obtain.

This section explores the properties of a different situation that is known as the Sequential Search Model. Unlike fixed sample search, where the consumer obtains a set of price quotes and then picks the lowest price, sequential search proceeds one at a time. The consumer samples from the population and gets a single price, then decides whether or not to accept it. If she rejects it, she cannot go back. As the epigraph shows, the sequential search model is easily applied to job offers, but it will be applied in this chapter to another common search problem—buying gas.

Setting Up the Model

Imagine you are driving down the road and you need fuel. As you drive, there are gas stations (say $N = 100$) to the left and right (taking a left does not bother you too much) and you can easily read the price per gallon as you drive up to each station. If you drive past a station, turning around is out of the question (there is traffic and you have a weird phobia about U-turns).

There is a lowest price station and the stations can be ranked from 1 (lowest, best price) to 100 (highest, worst price). You do not know the prices coming up because the stations are randomly distributed on the road. The lowest price station might be 18$^{\text{th}}$ or 72$^{\text{nd}}$ or even the very first one. Figure 7.7 sums it all up.

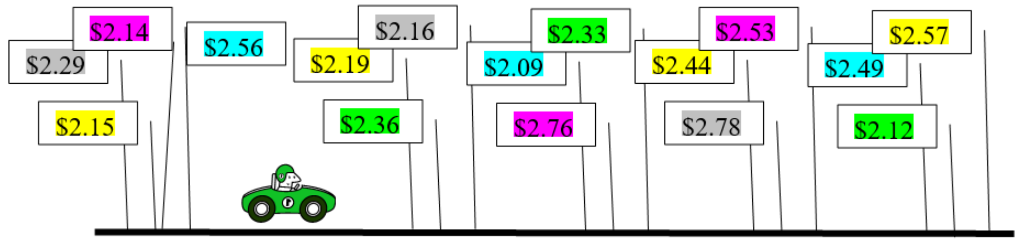


Figure 7.7: Deciding where to buy gas.

Suppose you focus on the following question: How do you maximize the chances of finding the cheapest station?

You might argue that you should drive by all of the stations, and then just pick the best one. This is a terrible idea because you cannot go back (remember, no U-turns). Once you pass a station, you cannot return to it. So, this strategy will only work if the cheapest station is the very last one. The chances of that are 1 in a 100.

A strategy for choosing a station goes like this: Pick some number $K < N$ where you reject (drive by) stations 1 to K, then choose the first station that has a price lower than the lowest of the K stations that you rejected.

Perhaps $K = 50$ is the right answer? That is, drive by stations 1 to 50, then look at the next (51^{st}) station and if it is better than the lowest of the 50 you drove by, pull in. If not, pass it up and consider the 52^{nd} station. If it is cheaper than the previous 51 (or 1 to 50 since we know the 51^{st} station isn't cheaper than the cheapest of the first 50), get gas there.

Continue this process until you get gas somewhere, pulling into the last (100^{th}) station if you get to it (it will have a sign that says, "Last chance gas station").

This strategy will fail if the lowest price is in the group of the K stations you drove by, so you might want to choose K to be small. But if you choose K too small, you will get only a few prices and the first station with a price lower than the lowest of the K stations is unlikely to give you the lowest price.

So, $K = 3$ is probably not going to work well because you probably won't get a super low price in a set of just three so you probably won't end up choosing the lowest price. For example, say the first three stations are ranked 41, 27, and 90. Then as soon as you see a station better than 27, you will pull in there. That might be 1, but with 26 possibilities, that's not likely.

On the other hand, a high value of K, say 98, suffers from the fact that the lowest price station is probably in that group and you've already rejected it! Yes, this problem is certainly tricky.

The Sequential Search Model can be used for much more than buying gas—it has extremely wide applicability and, in math, it is known as *optimal stopping*. In hiring, it is called the *secretary problem*. A firm picks the first K applicants, interviews and rejects them, then picks the next applicant that is better than the best of the K applicants. It also applies to many other areas, including marriage—search online for Kepler optimal stopping to see how the famous astronomer chose his spouse.

STEP Open the Excel workbook *SequentialSearch.xls* and read the *Intro* sheet, then proceed to the *Setup* sheet.

Column A has the 100 stations ranked from 1 to 100. The lowest priced station is 1, and the highest priced station is 100.

STEP Click the Randomly Assign button. It shuffles the stations, randomly distributing them along the road you are traveling in column D.

Cell B7 reports where the lowest priced station (#1) is located. Columns C and D report the location of each station. Column D changes every time you click the Randomly Assign button because the stations are shuffled.

Cell F2 sets the value of K. This is the choice variable in this problem. Our goal is to determine the value of K that maximizes the probability that we get the lowest priced station.

On opening, $K = 10$. We pass up stations 1 to 10, then take the next station that is better than the best of the 10 stations we rejected.

STEP Click the How Did I Do? button. This reshuffles the stations and draws a border in column D for the cell at the K^{th} station.

Cell F5 reports the best of the K stations (that were rejected). Cell F7 displays the station you ended up at.

STEP Scroll down to see why you ended up at that station and read the text on the sheet.

Cell F7 always displays the first station that is better (lower) than the best of the K stations in cell F5.

STEP Repeatedly click the How Did I Do? button. After every click, see how you did. Is 10 a good choice for K?

The definition of a good choice in this case is one that has a high probability of giving us the cheapest station. Our goal is to maximize the chances of getting the cheapest station. We could have a different objective, for example, minimize the average price paid, but this would be a different optimization problem. For the classic version of the optimal stopping problem, we count success only when we find the cheapest station.

STEP Change K to 60 (in F2) and repeatedly click the How Did I Do? button. Is 60 better than 10?

This is difficult to answer with the *Setup* sheet. You would have to repeatedly hit the How Did I Do? button and keep track of the percentage of the time that you got the cheapest station. That would require a lot of patience and time tediously clicking and recording the outcome. Fortunately, there is a better way.

Solving the Problem via Monte Carlo Simulation

The *Setup* sheet is a good way to understand the problem, but it is not helpful for figuring out the optimal value of K. We need a way to quickly, repeatedly sample and record the result. That is what the *MCSim* sheet does.

STEP Proceed to the *MCSim* sheet and look it over.

With $N = 100$ (we can change this parameter later), we set the value of K (in cell D7) and run a Monte Carlo simulation to get the approximate chances of getting the best station (reported in cell H7).

Unlike the MCSim add-in used in the previous section, this Monte Carlo simulation is hard wired into this workbook. Thus, it is extremely fast.

STEP With $N = 100$ and $K = 10$, click the Run Monte Carlo Simulation button. The default number of repetitions is 50,000, which seems high, but a computer can do hundreds of thousands of repetitions in a matter of seconds.

Figure 7.8 shows results. Choosing $K = 10$ gives us the best station about 23.4% of the time. Your results will be slightly different.

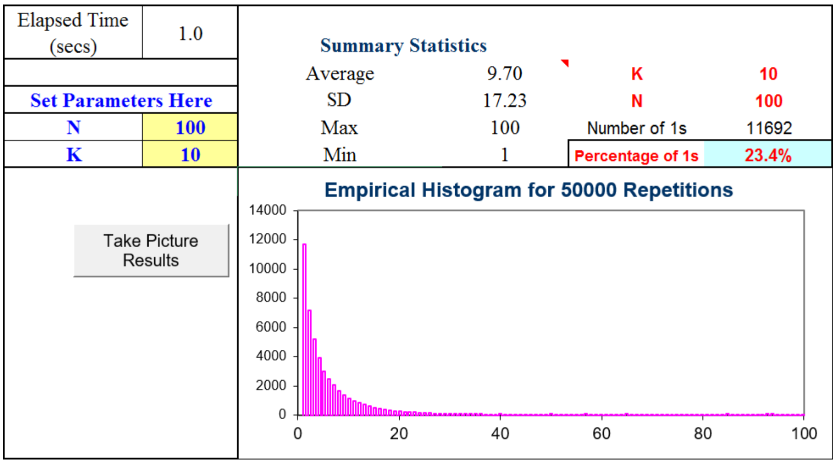


Figure 7.8: Monte Carlo simulation results.
Source: SequentialSearch.xls!MCSim

Notice that we are using Monte Carlo simulation to approximate the exact answer. Monte Carlo simulation cannot give us the exact answer. By increasing the number of repetitions, we improve the approximation, getting closer and closer, but we can never get the exact truth with simulation. The answer it gives depends on the actual outcomes in that particular run. The only way simulation would give the exact answer is if it was based on an infinite number of repetitions.

Can we do better than getting the best station about 23% of the time?

We can answer this question by exploring how the chances of getting the lowest price varies with K. By changing the value of K and running a Monte Carlo simulation, we can evaluate the performance of different values of K.

STEP Explore different values of K and fill in the table in cells J3:M10.

As soon as you do the first entry in the table, $K = 20$, you see that it beats $K = 10$.

STEP Use the data in the filled in table to create a chart of the chances of getting the lowest price station as a function of K. Use the button under the table to check your work.

What do you conclude from this analysis?

One problem with Monte Carlo simulation is the variability in the results. Each run gives different answers since each run is an approximation to the exact answer based on the outcomes realized. Thus, it seems pretty clear that the optimal value of K is between 30 and 40, but using simulation to find the exact answer is difficult.

Figure 7.9 displays results of series of Monte Carlo experiments. Notice that we doubled the number of repetitions to increase the resolution. The best value of K appears to be 36, but the noisiness in the simulation results makes it impossible to determine the answer.

Number of Repetitions	Value of K	Percentage of 1s	Average
100,000	30	35.84%	17.06
100,000	31	36.54%	17.41
100,000	32	36.59%	17.76
100,000	33	36.81%	18.24
100,000	34	36.60%	18.99
100,000	35	36.61%	19.30
100,000	36	36.93%	19.58
100,000	37	36.73%	20.21
100,000	38	36.62%	20.56
100,000	39	36.55%	21.38
100,000	40	36.52%	21.57

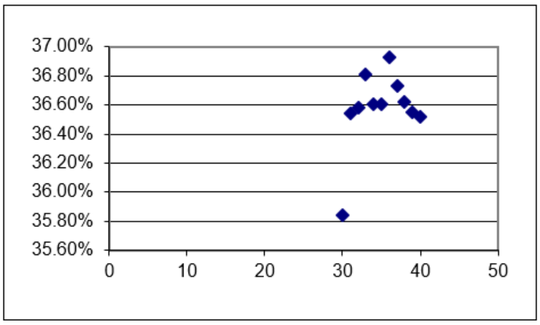


Figure 7.9: Zooming in on the value of optimal K.
Source: SequentialSearch.xls!Answers

With Monte Carlo simulation, we can continue to increase the number of repetitions to improve the approximation.

STEP Proceed to the *Answers* sheet to see more simulation results.

The *Answers* sheet shows that even 1,000,000 repetitions are not enough to definitively give us the correct answer. Simulation is having a difficult time distinguishing between a stopping K value of 36 or 37.

An Exact Solution

This problem can be solved analytically. The solution is implemented in Excel. For the details, see the Ferguson citation at the end of this chapter.

STEP Proceed to the *Analytical* sheet to see the exact probability of getting the cheapest station for a given K-sized sample from N stations from 5 to 100.

For example, cell G10 displays 32.74%. This means you have a 32.74% probability of getting the cheapest station out of 10 stations if you drive by the first six stations and then choose the next station that has a price lower than the cheapest of the K stations you drove by.

For $N = 10$, is $K = 6$ the best solution?

No. The probability of choosing the cheapest station rises if you choose $K = 5$. The 3 and 4 choices are close, but clearly, optimal $K = 3$ (with a 39.87% likelihood of getting the cheapest station) is the best choice.

In the example we have been working on, we had $N = 100$. Monte Carlo simulations showed optimal K around 36 or 37, but we were having trouble locating the exact right answer.

STEP Scroll down to see the probabilities for $N = 100$. Click on cells AL100 and AM100 to see the exact values. The display has been rounded to two decimal (percentage) places, but the computation is precise to more decimal places.

$K^* = 37$ just barely beats out $K = 36$. The fact that they almost give the exact same chances of getting the lowest price explains why we were having so much trouble zooming in on the right answer with Monte Carlo simulation.

It can be shown (see the Ferguson source in the References section) that optimal K is $\frac{N}{e}$, giving a probability of finding the cheapest station of $\frac{1}{e}$. For $N = 100$, $\frac{N}{e} \approx 36.7879$.

If K was a continuous endogenous variable, $\frac{N}{e}$ would be the optimal solution. But it is not, so the exact, correct answer is to pass on the first 37 stations and then take the first one with a lower price than the lowest price of stations 1 to 37.

It is a mystery why the transcendental number e, the base of natural logarithms, plays a role in the solution.

Figure 7.10 shows that as N rises, so does optimal K. What elasticity is under consideration here?

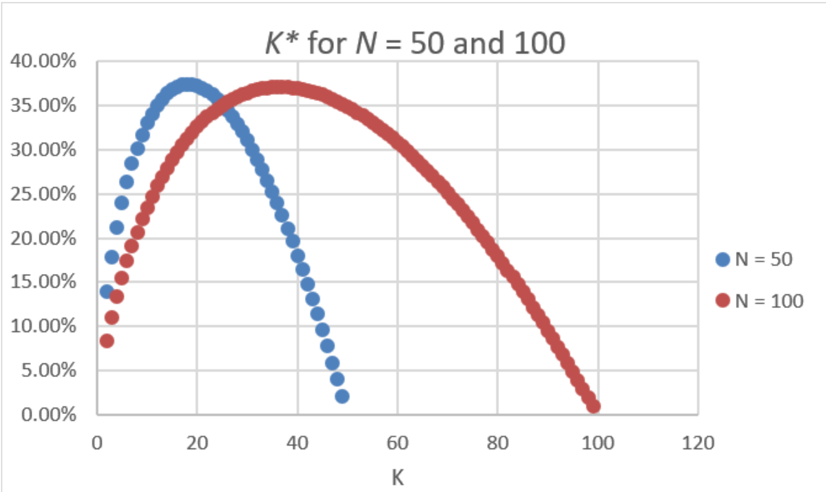


Figure 7.10: Exact probabilities of finding the cheapest station.
Source: SequentialSearch.xls!Analytical

The answer is the N elasticity of K. From $N = 50$ to 100 is a 100% increase. What happens to optimal K? It goes from 18 to 37, so a little more than a 100%. The elasticity is slightly over one. If you use the continuous version of K, then K exactly also doubles and the N elasticity of K is exactly one.

Sequential Search Lessons

Unlike the Fixed Sample Search Model (where you obtain a set of prices and choose the best one), the Sequential Search Model says that you draw sample observations one after the other. This could apply to a decision to choose a gas station. As you drive down the road, you decide whether to turn in and get gas at Station X or pass up that station and proceed to Station Y.

Faced with price dispersion, a driver deciding where to get gas can be modeled as solving a Sequential Search Model. Although there can be other objectives (such as getting lowest average price), the goal could be to maximize the chances of getting the lowest price. We found that as N rises, so does optimal K. The more stations, the more driving you should do before picking a station.

Like the Fixed Sample Search Model, the Sequential Search Model does not have any interaction between firms and consumers. Price dispersion is given and the model is used to analyze how consumers react in the given environment.

In the pre-internet and smartphone days, deciding where to get gas was quite the challenge. A driver passing signs with prices (like Figure 7.7) was a pretty accurate representation of the environment. There was no Google maps or apps displaying prices all around you. Notice, however, that the Law of One Price does not yet apply to gas prices.

Ferguson points out that our Sequential Search Model (which mathematicians call the secretary problem) is part of a class of finite-horizon problems. "There is a large literature on this problem, and one book, *Problems of Best Selection* (in Russian) by Berezovskiy and Gnedin (1984) devoted solely to it" (Ferguson, Chapter 2).

Fixed Sample and Sequential Search Models are merely the tip of the iceberg. There is a vast literature and many applications in the economics of search, economics of information, and economics of uncertainty.

Exercises

1. Use the results in the *Analytical* sheet to compute the N elasticity of K^* from $N = 10$ to 11. Show your work.

2. Use the results in the *Analytical* sheet to draw a chart of K^* as a function of N. Copy and paste your graph in a Word document.

3. Run a Monte Carlo simulation that supports one of the N-K^* combinations in the *Analytical* sheet. Take a picture of your simulation results and paste it in a Word document.

4. Explain why the Monte Carlo simulation was unable to exactly replicate the percentage of times the lowest priced station was found.

References

The epigraph is from pages 115 and 116 of J. J. McCall, "Economics of Information and Job Search," *The Quarterly Journal of Economics*, Vol. 84, No. 1 (February, 1970), pp. 113–126, www.jstor.org/stable/1879403. This paper shows that sequential search (with recall) dominates fixed sample search. For more on this point, see Robert M. Feinberg and William R. Johnson, "The Superiority of Sequential Search: A Calculation," *Southern Economic Journal*, Vol. 43, No. 4 (April, 1977), pp. 1594–1598, www.jstor.org/stable/i243526.

Thomas Ferguson, *Optimal Stopping and Applications* is freely available online at www.math.ucla.edu/ tom/Stopping/Contents.html. Ferguson offers a technical, mathematical presentation of search theory.

C. J. McKenna, *The Economics of Uncertainty* (New York: Oxford University Press, 1986), is a concise, nontechnical introduction to imperfect information models.

John Allen Paulos, *Beyond Numeracy* (New York: Alfred A. Knopf, 1991), p. 64, discusses the optimal interview problem with an easy, intuitive style.

This course surveys research which incorporates psychological evidence into economics. Topics include: prospect theory, biases in probabilistic judgment, self-control and mental accounting with implications for consumption and savings, fairness, altruism, and public goods contributions, financial market anomalies and theories, impact of markets, learning, and incentives, and memory, attention, categorization, and the thinking process.

MITOpenCourseware

Chapter 8

Behavioral Economics

The field of Behavioral Economics (and Behavioral Finance) is a growing research area that focuses on how decisions are actually made. It is closely tied to psychology and neuroscience. Behavioral economists reject the idea of utility maximization as an assumed black box. Both experimental methods and sophisticated procedures (such as MRI brain scans) are used to examine how real-world problems are actually solved. A number of results have emerged that challenge the conventional wisdom in mainstream economics.

One area of long-standing interest in psychology involves repeated choice problems. This chapter focuses on a particular kind of repeated choice in which the satisfaction obtained currently depends on past decisions. This is called distributed choice.

Suppose you are deciding whether to watch TV or play a video game. You face this choice repeatedly. The satisfaction from watching TV or playing a video game depends on how often that choice has been made before. What is the best combination of TV and video games over a period of time and, more importantly, how well do people handle this kind of repeated choice?

Instead of explaining why the repeated choice optimization problem is difficult and presenting results from human trials, it is more fun (and you will learn more) to let you first participate in an experiment.

The Choice Game

STEP Open the Excel workbook *Melioration.xls* and read the Intro sheet, then go to the *Choice Game* sheet to play this simple game.

Your goal is to click the A or B buttons as many times as possible in 10 minutes. When you make a choice, by clicking on one of the buttons, you are forced to wait. Waiting is costly because you cannot click (make another choice) while waiting.

STEP Click the Practice option button (near the top left corner of the screen) to see how the game works.

You get up to 100 practice trials. In practice mode, time is not kept. You can take as long as you want between button clicks. Practice now.

There is definitely something going on that you are trying to figure out and there is an optimal strategy. You can click the same button over and over or switch back and forth.

Are you ready to play? Unlike practice, when you play, a timer will be running. You will not use the buttons on the sheet like you did in practice mode. The buttons will be on a dialog box, right next to each other. You will have 10 minutes to make as many choices as possible. The time remaining will be displayed as you play.

Ten minutes might be too long for you to play so click the Exit Now button if you want to stop playing. As long as you start play and make a few choices, you will be able to continue working and learning about melioration.

STEP Click the Play option button. Good luck!

After you finish the game, a message box displays your score and a *Results* sheet shows a record of your picks. It reports results based on a full ten minutes of play, so if you stopped prematurely, you can ignore your results.

Let's deconstruct this game and see how it works. Figure 8.1 shows the first 10 choices made by a player. The player started with A, then switched to B with his 7^{th} choice, but switched back to A, then ended with B.

	A	B	C	D
1	Choice Number	Pause Time A	Pause Time B	Choice Made
2	1	2.00		A
3	2	2.40		A
4	3	2.80		A
5	4	3.20		A
6	5	3.60		A
7	6	4.00		A
8	7		8.00	B
9	8	4.40		A
10	9	4.80		A
11	10		7.60	B

Figure 8.1: Ten plays of the game.

STEP You can see the full record of yet another player by clicking the Show Example button (near cell G9 in the *Results* sheet, which was revealed when you finished playing the choice game).

This player tried streaks of A and B. Notice how the time paused changed.

These results sheets also compare the number of choices made to the maximum possible and computes a score as a percentage of the maximum. Let's find out how the maximum can be attained and why people are usually so bad at playing this game.

Actual Results

Experimental trials with this game were conducted by Herrnstein and Prelec (1991) and you can compare how you did to the average result (and to the player in the *MoreResults* sheet).

STEP Click the Show Data button in the *MoreResults* sheet.

The *Data* sheet shows how 17 subjects played the choice game that you just played. Each dot in the chart, reproduced in Figure 8.2, shows the fraction of times that a player chose A (on the x axis) and the corresponding average delay endured by that player (on the y axis). The player with the shortest delay, the first one in the table, also has the most choices (number of choices = 600/average delay) and is the winner in this set of players. How did you do?

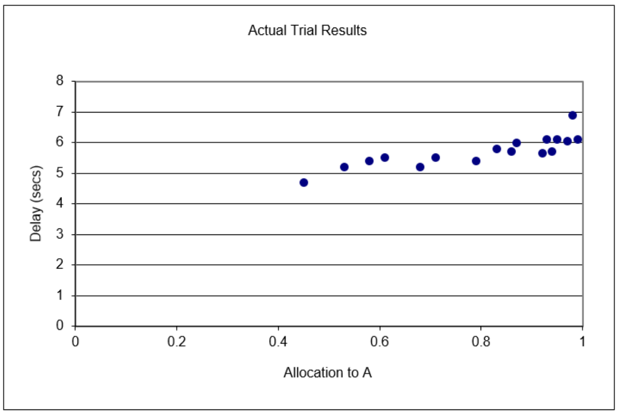


Figure 8.2: Actual results from a single session of the choice game.
Source: Melioration.xls!Data, revealed after game played.

STEP To add your result to the chart in Excel, copy your results from cells J2 and K2 of the *Results* sheet, select cell A23 in the Data sheet, and Paste Special (Values) (or simply type in the two numbers). A red dot will appear in the chart. This shows how you did.

Did you beat the best player out of the 17 in the chart? We know you could have because even the best player in that group of 17 failed to optimize. The explanation for this failure requires that we understand the delay function for each choice.

The heart of the choice game is the wait time between choices. The duration of the pause is a function of the previous 10 choices (including the current choice). For choice A, the wait time, in seconds, is 2 + 0.4 x Proportion of A Choices in the last 10 choices. So, if the last 10 choices had been B, then A would have a very short and satisfying pause time of just 2 seconds. As you click on A, however, the pause time for choice A rises by 0.4 seconds until it hits a maximum of 6 seconds.

Choice B's wait time is determined by 8 - 0.4 x Proportion of B Choices in the last 10 choices. As you click on B, the duration of the pause gets lower and lower until reaching a minimum of 4 seconds.

STEP Confirm that the wait times were determined as described by returning to the three results sheets and examining the pause times in columns B and C.

You can see that the first clicks of A and B had pause times of 2 and 8 seconds, respectively. You can also check that each pause time is following the functions described above. The *MoreResults* sheet with the streaky A and B strategy makes it easy to see the mechanics of the choice game.

Choice A exhibits increasing marginal cost—every time you click on A, you are penalized and forced to wait longer. Choice B rewards you with a decrease in wait time when it is clicked, but the wait time starts very high so you have to be persistent and stick to it. Plus, choice A is always 2 seconds lower than choice B so you are constantly being lured toward choice A.

Most people play this game by being attracted to A's short wait time, until it gets unbearable and they switch to B. But they can't stay with B long because it is painful to wait at first and they do not have the patience and self-discipline to stick with B. Sound familiar? B could be exercise or dieting or studying—you know you should and it gets easier if you stick to it, but it can be hard to start.

Now that you know the rules of the game, how do you actually optimize with this game? Simple—start with choice B and never deviate.

STEP To see this optimal strategy in action, go to the *Solution* sheet by clicking the Show Solution button in the *Data* sheet (below the chart).

Column B shows what happens when you exclusively choose A. It starts well, but you end up with many 6 second pauses.

STEP Scroll down to see that you make 103 choices in 600 seconds, yielding an average delay of 5.8 seconds. This is a poor outcome.

Column F displays what happens when B is exclusively chosen. The first few wait times are long, but each choice of B lowers the wait time until the minimum, 4 seconds, is reached.

STEP Scroll down to see that clicking choice B every time lets you make 144 choices (with an average delay of 4.167 seconds).

The strategy of choosing B exclusively cannot be beat (except for an endgame correction, which is one of the exercise questions). If the player switches from B to A, the temporary gain is swamped by higher wait times when the inevitable switch back to B occurs.

To be sure that this point is clear, consider switching after having reached the 4 second minimum pause time for choice B. What would happen?

STEP Change cell K15 (in the *Solution* sheet) to A.

Five consecutive A choices are made and each one has a pause time less than or equal to four seconds, as shown in column L. Thus, we have saved time. But when we switch back to B (since we know A's pause time will continue to rise and we can get to 4 seconds with B), we have to suffer higher pause times. The trade-off is not worth it. We end up making fewer choices (142 instead of 144) and suffering a longer average delay.

The *Solution* sheet makes clear the following key point: The optimal strategy is to choose B exclusively and never deviate. If you failed to do this, do not worry; you have plenty of company. Very few humans figure this out.

Melioration Explained

Herrnstein and Prelec (1991) designed the experiment to test for the presence of something called melioration (pronounced mee-lee-uh-RAY-shun). To *meliorate* (or *ameliorate*) means to make better or more tolerable. Melioration says that we are drawn to choices that *immediately* reduce pain or give immediate satisfaction. We do a poor job of maximizing when there is a trade-off between short- and long-run returns. We are shortsighted and look to make immediate improvements. In fact, melioration has been found in other animals besides humans.

The attraction of switching to A and having the pause time fall is melioration at work. The immediate pain of waiting is lessened and, thus, players are drawn toward choice A.

In addition to the actual choices from the 17 players, Figure 8.3 shows wait times for choices A and B given the proportion of A choices in the previous 10. It is easy to see, once again, that the optimal solution is to choose B exclusively because that lets you travel down the solid line to the intercept at 4 seconds. If you ever jump on the A train, you are swept upwards toward a 6-second wait time.

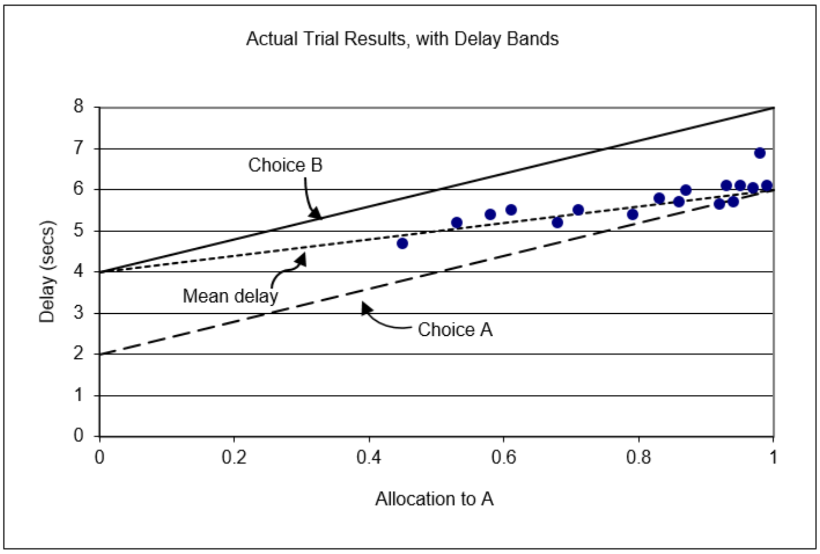


Figure 8.3: Understanding melioration.
Source: Melioration.xls!Data, revealed after game played.

Figure 8.3 shows that if the last 10 choices were B and then A was chosen, the player would immediately gain a reduction in wait time from 4 to 2 seconds (jumping from the higher to the lower line). For a few choices, the player would be better off, but after the 5$^{\text{th}}$ consecutive A choice, the wait time would be greater than 4 seconds. The player would be forced to endure longer wait times than would have been obtained by sticking with B.

Furthermore, it is hard to switch to B because wait time immediately jumps by 2 seconds. The player will have to suffer through the ride down the B line, with choice A promising a 2-second decrease with every click. The immediate attraction of the 2-second decrease is the core of the melioration process that guides subjects to choose A.

Figure 8.3 makes clear that the 17 human subjects who played the choice game failed to optimize. The fraction of allocation to A should be zero, but most players do not do this. This begs the question, so what?

Herrnstein and Prelec (1991) argue that the lack of optimization is a big deal. For them, choice is often not a single, isolated decision, but a series of many decisions, distributed over time. Frequency of athletic exercise, buying lottery tickets, choices of restaurants, and rate of work in freelance occupations are some of the examples offered.

For all of these distributed choice problems, melioration is common and this means people systematically fail to optimize. "This would imply that preferences as revealed by the marketplace may be a distortion of the true underlying preferences" (Herrnstein and Prelec, 1991, p. 137). Melioration helps explain complaints about one's own behavior (such as exercising too little), which is part of a growing literature on self-control. It also may contribute to the study of impulsiveness and addiction.

Of course, this presumes that the laboratory findings carry over to real-world settings. This is often an Achilles' heel of experimental economics. Results are often criticized as having little external validity because they are based on fake scenarios played by college students. Herrnstein and Prelec (1991) acknowledge that little money was at stake (they paid their players based on performance), but they rely on two other motivating factors. "First, delays are genuinely annoying and the difference between two and four seconds is not trivial, as any computer user will appreciate. Second, the 'puzzle' nature of the experiment presents a challenge that is presumably satisfying to solve" (Herrnstein and Prelec, 1991, p. 144).

Others have tried to nail down exactly what causes melioration and how it can be overcome. Neth, Sims, and Gray (2005, p. 357) were surprised:

> We hypothesized that frequent and informative feedback about optimal performance might be the key to enable people to overcome the documented tendency to meliorate when choices are rewarded probabilistically. Much to our surprise, this intuition turned out to be mistaken. Instead of maximizing, 19 out of 22 participants demonstrated a clear bias towards melioration, regardless of feedback condition.

The Future of Behavioral Economics

With faculty, courses, conferences, and specialized journals, there is no doubt that Behavioral Economics is here to stay. In 2002, the Nobel Prize in Economic Sciences was awarded to Daniel Kahneman and Vernon Smith for work incorporating psychology and laboratory methods in the study of decision making. Richard Thaler won the Nobel in 2017 for his contributions to behavioral economics.

Unlike conventional economics, which simply assumes optimizing behavior and rationality, behavioral economists seek to determine under what conditions agents struggle to optimize. They work with psychologists and neuroscientists to devise tests and laboratory experiments. The key result is that they find persistently sub-optimizing behavior.

Melioration is but one simple example of work in this area. Melioration means that decision makers fail to optimize because they focus on the small (immediate, single choice) instead of the large (future, many choices). This can be applied any time that incremental steps lead to an undesirable place:

> A person does not normally make a once-and-for-all decision to become an exercise junkie, a miser, a glutton, a profligate, or a gambler; rather, he slips into the pattern through a myriad of innocent, or almost innocent choices, each of which carries little weight. Indeed, he may be the last one to recognize "how far he has slipped," and may take corrective action only when prompted by others. (Herrnstein and Prelec, 1991, p. 149)

According to the behavioral economists, the list of examples where humans struggle to optimize is actually quite long. Evaluating probabilities (such as risk), choice over time, and misperception of reality are all areas being actively studied.

It remains unclear whether the results being generated by behavioral economists are merely a series of peculiar puzzles that will extend the boundaries of economics or more serious anomalies that will one day bring down the paradigm of rationality and optimizing behavior that is the hallmark of modern, mainstream economics.

Exercises

If you did the Q&A problems and changed the parameters, set them back to the original values (2 and 0.4 for A and 8 and -0.4 for B).

1. With your observation included, copy and paste the chart titled *Actual Trial Results* in a Word document. Comment briefly on how you did.

2. What endgame correction could be implemented to increase the total number of choices? What is the true, exact maximum number of choices? Explain.

   Herrnstein and Prelec (1991), p. 142, point out that, "In fact, the subjects showed no evidence of having been influenced by the endgame contingency."

3. With columns Q:U in the *Solution* sheet, use Solver to find the optimal solution to the choice game. Notice how the choice variables have been constrained. How does Solver do? Explain.

4. Training someone to touch type does not guarantee continued touch typing in the workplace. How would melioration explain this result?

References

The epigraph is from a course available freely at ocw.mit.edu. The course description in the epigraph was from the Spring 2004 version of Behavioral Economics and Finance (see ocw.mit.edu/courses/economics/14-127-behavioral-economics-and-finance-spring-2004/). The readings for this course include introductory and more advanced work.

The repeated choice problem in this chapter is based on two papers: (1) Richard J. Herrnstein and Drazen Prelec, "Melioration: A Theory of Distributed Choice," *The Journal of Economic Perspectives*, Vol. 5, No. 3 (Summer, 1991), pp. 137–156, www.jstor.org/stable/1942800 and (2) Herrnstein and Prelec's "Melioration," pages 235–263 in *Choice Over Time*, edited by George Loewenstein and Jon Elster (1992).

Herrnstein, a psychologist, teamed up with Charles Murray, a political scientist, to write a controversial book titled *The Bell Curve: Intelligence and Class Structure in American Life* (1994). The book argued that nature (IQ) is more important than nurture (socioeconomic status) in explaining a wide range of outcomes.

Another paper specifically focused on melioration is Hansjörg Neth, Chris R. Sims, and Wayne D. Gray, "Melioration Despite More Information: The

Role of Feedback Frequency in Stable Suboptimal Performance," *Proceedings of the Human Factors and Ergonomics Society* 49th Annual Meeting, 2005, doi.org/10.1177/154193120504900330.

There are many books on behavioral economics and finance. A classic is from Nobel Prize winner Richard Thaler, *The Winner's Curse: Paradoxes and Anomalies of Economic Life* (1994). This is a good place to start learning about behavioral economics. Other good reads include the following:

Dan Ariely, *Predictably Irrational: The Hidden Forces that Shape Our Decisions* (2008).

Daniel Kahneman, *Thinking Fast and Slow* (2011).

Michael Lewis, *The Undoing Project: A Friendship that Changed Our Minds* (2016).

Richard Thaler, *Misbehaving: The Making of Behavioral Economics* (2015).

Richard Thaler and Cass Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (2008).

How use doth breed a habit in a man!

William Shakespeare

Chapter 9

Rational Addiction

This chapter is different. It does not have steps that you follow as you work in Excel. It does not have any exercise questions. There is an Excel file that you will open and work on, but it is entirely self-contained. Just open the file and start reading.

Before you begin, however, consider a little of the science behind learning. Once we know how we learn, then we can optimize!

The Neuroscience of Learning

Suppose you want to improve your free throw shooting and you really cared about this so you decided to practice for one hour per day for two weeks. Most people think that standing at the free throw line and shooting free throws would be the best use of your time, but this is wrong. A much better use of your one hour per day is to shoot from all over the court—spend 10 minutes in one spot, then move to another spot, varying distance from say 10 to 20 feet (the free throw line is 15 feet from the basket). This is *interleaved practice* and it works also for learning and studying.

Interleaved practice is counter-intuitive and paradoxical. Many coaches refuse to believe it, but careful controlled experiments in a variety of applications reveal it is a fundamental principle (Brown, et al., 2014). It works for physical skills (don't throw 100 curve balls, interleave with other pitches), memorization (don't repeat one thing, interleave items), and higher learning—reflect on how this book has repeated concepts like elasticity in a variety of applications.

In addition to interleaving, below is a list of best-practice learning strategies that you can apply to every course you take:

1. Interleaved Practice (switching)
2. Spaced Practice (avoid cramming)
3. Elaboration (invent your own how and why questions)
4. Concrete Examples (the more specific, the better)
5. Dual Coding (words and visuals)
6. Retrieval Practice (repeatedly recall what you know)

Unbeknownst to you, this book has been using all of these strategies to help you learn.

To get more information on these six science-based ways to learn more efficiently, visit these two web sites:

- www.learningscientists.org/posters
- www.youtube.com/watch?v=CPxSzxylRCI

And one more thing that you believe about learning that is wrong: you think your ability to learn economics (or math or music) is preordained. Your brain either has a knack for economics or it does not and, if not, you cannot learn economics (or math or music). This is wrong.

Neuroscience makes clear that your brain is *plastic*. It is moldable and flexible. You have already learned a great deal of economics, math, and Excel. Yes, some details are fuzzy and you have not mastered every single thing, but keep trying. As you see more examples and applications, it gets easier to grasp and your understanding deepens.

Rational Addiction

As you work on the Excel file, you will be reviewing concepts and feel comfortable with Solver, charts, and Excel itself. This will reinforce basic material that you already know, but you will also be exposed to some new ideas as you continue to master the economic way of thinking.

This application is controversial and generates passionate debate. Non-economists, especially, find it outrageous. After you finish, you can make up your own mind on what you think about it.

Open *RationalAddiction.xlsm* to begin.

References

The epigraph is from Shakespeare's *Two Gentlemen of Verona*. The *Conversation* sheet in *RationalAddiction.xlsm* explains what it means.

The full story behind the puzzling interleaved practice phenomenon and much more about how we learn is in Peter Brown, Henry Roediger III, and Mark McDaniel (2014), *Make it Stick: The Science of Successful Learning*.

Part II

The Theory of the Firm

For Friedman, lack of realism of assumptions is not a virtue. It is a necessary evil: to base theories on absolutely realistic assumptions is like drawing a map on a one-to-one scale.

Mark Blaug

Overview

Consumer Theory focuses on the buyer. It models a consumer's optimization problem and emphasizes deriving a demand curve as the most important result.

The Theory of the Firm is about the seller. Firm decisions about inputs and outputs are modeled as optimization problems. The key result will be deriving a supply curve.

The chapters are organized as shown in Figure II.1. Notice that the production function is the first idea presented. It plays a role in each of the three optimization problems faced by the firm.

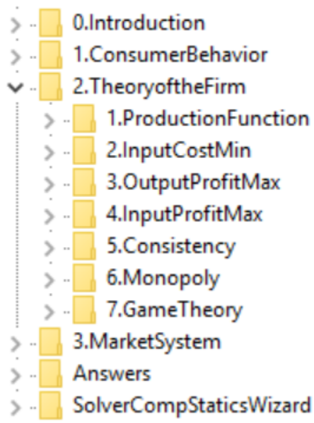


Figure II.1: Content map with focus on the theory of the firm.

Figure II.1 also provides a broad overview of the entire landscape. We have completed the Theory of Consumer Behavior and, once we finish our work in the Theory of the Firm, we will be ready to analyze the behavior of consumers and firms together in part III, the Market System.

Unlike the Theory of Consumer Behavior, the Theory of the Firm is made up of three interrelated optimization problems.

1. Input cost minimization: Choose inputs to minimize the cost of producing a given level of output. Derive the cost function by changing q and tracking the minimum total cost.

2. Output profit maximization: Choose output to maximize profits. Derive the supply curve by changing the price and tracking the optimal output.

3. Input profit maximization: Choose inputs to maximize profits. Derive an input demand curve by changing an input price and tracking optimal input use.

Of course, optimization and comparative statics play a prominent role, but watch out for these three crucial innovations in the Theory of the Firm.

1. Market structure: The Theory of the Firm includes the market environment as an important consideration in the model. The firm can be a *price taker*, a perfectly competitive firm, or a *price maker*, a monopolist. There are many other market structures, for example, oligopoly (where there are a few firms) and monopolistic competition.

2. Time period: The Theory of the Firm distinguishes between long run and short run decision making horizons. In the *long run*, all factors are freely variable and firms may enter or exit the industry. In the *short run*, at least one input (usually capital) is fixed and the firm may cease production (shut down), but it must pay fixed costs whether it produces or not.

3. Output is cardinally measurable: Unlike utility, the output produced by a firm and the resulting revenues, costs, and profits can be directly observed and measured on a cardinal scale. Thus, we will be able to use and interpret the Lagrangean multiplier.

Methodology

The assumptions underlying the Theory of Consumer Behavior are never seen in reality and the Theory of the Firm doubles down on this strategy by making even more outlandish assumptions. The time has come to explain why economists do this.

Each discipline has its own rules for determining truth and acceptable procedures for producing knowledge. These rules and norms are known as the *methodology*. For example, economics utilizes highly abstract models. The assumptions of these models are plainly unrealistic and false. Real-world human beings do not behave like perfectly rational, calculating machines. Then why do economists assume they do?

The classic defense of unrealistic assumptions is "The Methodology of Positive Economics," the first chapter in Milton Friedman's (1953) *Essays in Positive Economics*. Friedman's argument was initially controversial, but it became conventional thinking in economics. Friedman urged economists to ignore how unrealistic the assumptions were and focus on the predictive power of a model. If you want to predict how billiard balls will move when hit by an expert pool player, vectors and complicated mathematics are involved.

> 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. (Friedman, 1953, p. 21)

The Theory of the Firm (like the Theory of Consumer Behavior) is built on the idea of decision makers acting as if they were rationally calculating and optimizing agents. This is plainly unreal, but the point is not to describe how consumers or firms actually make decisions. Instead, we want a model that makes predictions about changes in output, for example, as product price changes (this is a supply curve).

We know firms do not take price as given and there is no such thing as perfect competition in the real world, but we assume this because we are not trying to build an accurate representation of an actual firm. Instead, we want to able to predict how a firm responds to a price shock—just like we want to predict how a billiard ball will move when struck.

It is quite easy to forget the methodology of economics and find oneself wondering how economists can believe such a ridiculously unreal and abstract model of a firm. Remember, economists do not test theories via the assumptions—it is the implications that matter.

The usefulness of abstract models and unrealistic assumptions often drives opinion or evaluation of the work. For example, the economic theory of rational addiction says addicts rationally "choose" harmful addiction or dangerous activities. To some, this is so obviously untrue that they cannot engage with the theory.

The term *homo economicus*, a non-existent version of *homo sapiens*, is used to mock the narrow, calculating humans that inhabit the made-up world of economics.

In France, over a thousand Economics graduate students signed a letter in 2000 attacking the abstract, unrealistic, mathematical training they were receiving. This launched the Post-Autistic Economics movement.

Friedman's view came to dominate mainstream economics, but it did not end the argument in philosophical circles and heterodox economics. The debate about methodology rages on while most economists continue to build and work with highly abstract, completely unrealistic models.

Your Role—A Reminder

As before, mastery of the Theory of the Firm requires your effort, energy, and engagement. Be sure to **experiment**, changing cells and asking "what if" questions as you proceed through the Excel workbooks. Focus on the repeated patterns and continue to add to your stock of knowledge.

Remember that economics has a core logic that has been referred to as "the economic way of thinking" or "the economic approach." Learning to see and think like an economist should be your ultimate goal.

References

The epigraph is from page 703 of the third edition (1978) of Mark Blaug's *Economic Theory in Retrospect* (originally published in 1962). Blaug's concluding chapter, "A Methodological Postscript," is a good review of how theories develop and knowledge grows.

Methodology is part of the *philosophy of science*. Economists pay little attention to methodology, but that does not mean it is unimportant.

Let us choose that function $P' = bL^kC^{k-1}$ and find such numerical values of b and k that P will "best" approximate P [output] in the sense of the Theory of Least Squares. Then relative to the indices and the period we have the norm $P = 1.01L^{\frac{3}{4}}C^{\frac{1}{4}}$.

Charles W. Cobb and Paul H. Douglas

Chapter 10

Production Function

The production function is the backbone of the Theory of the Firm. It describes the current state of technology and how input can be transformed into output.

The production function can be displayed in a variety of ways, including product curves and isoquants. In every optimization problem faced by the firm, the production function is included.

Key Definitions and Assumptions

Inputs, also known as factors of production, are used to make output, sometimes called product. As shown in Figure 10.1, the firm is a highly abstract entity—a black box—that transforms inputs into output.

Figure 10.1: The black box nature of the firm.

The specific details of how the firm is organized and how it actually combines the inputs to make goods and services is ignored by the theory, hidden in the black box.

Inputs are often broken down into large categories, such as land, labor, raw materials, and capital. We will simplify even further by collapsing everything that is not labor into the capital category.

Labor, L, is human toil and effort. It is measured in units of time, usually hours.

Capital has a confusing history in economics. As a factor of production, *capital*, K, means things that produce other things, such as machinery, tools, or equipment. That is different from financial or venture capital that is a fund of money. The title of Karl Marx's famous book, *Das Kapital*, uses capital in the sense of wealth, denominated in money. The Theory of the Firm's K is measured in numbers of machines.

Like labor, capital is rented. The firm does not own any of its machines or buildings. This is extremely unrealistic, but allows us to avoid complicated issues involving depreciation, financing of machinery purchases (debt versus equity, for example), and so on.

Another extreme simplifying assumption is that there is no time involved. Like the consumer maximizing utility subject to a budget constraint, the firm exists only for a nanosecond. It makes decisions about how much to produce to maximize profits with no worries about inventories or the trajectory of future sales. It produces the output in an instant.

We avoid complications arising from the production of more than one good or service by assuming that the firm produces only one product. That makes revenues simply price times quantity sold of the one product.

Without going into detail again about unrealistic assumptions, it seems helpful to point out that we are not trying to build an accurate model of a real-world firm. Our primary goal is to derive a supply curve. We want to know how a firm responds to a change in price, ceteris paribus. By assuming away many real-world complications, we can model the firm's maximization problem, solve it, and do comparative statics to get the supply curve.

Mathematical Representation

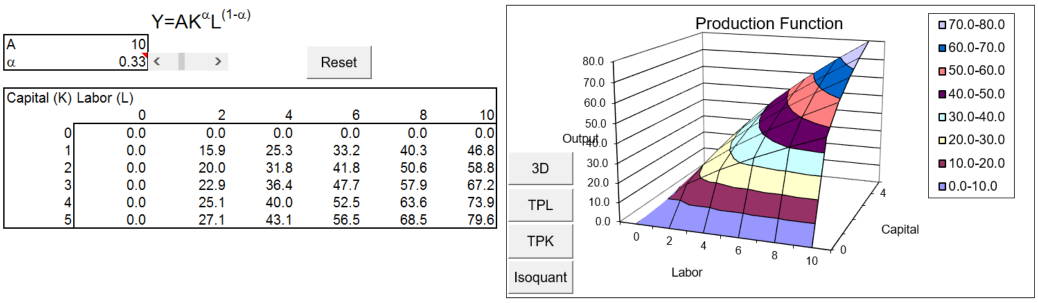
Just like the Theory of Consumer Behavior, which uses a utility function to model tastes and preferences, the Theory of the Firm uses a production function to capture the ability of firm's to transform inputs into outputs. Unlike utility, production is objective and observable. We can count how much output is made from a given number of hours of labor and machines.

The *production set* describes all of the technologically feasible outputs from a given amount of inputs. The *production function* describes the maximum output possible from a given amount of inputs. Notice how the production function assumes the inputs are being used in the best way possible.

The most abstract, general notation for a production function is $y = f(L, K)$. The $f()$ represents the technology available to the firm. A specific, concrete example of a production function is the Cobb-Douglas functional form: $y = AL^{\alpha}K^{\beta}$. Let's see what it looks like in Excel.

STEP Open the Excel workbook *ProductionFunction.xls*, read the *Intro* sheet, then go to the *Technology* sheet to see an example of the production function.

In Figure 10.2, the production set is the surface of the 3D object and everything inside; the production function is just the surface.



Capital (K) \ Labor (L)	0	2	4	6	8	10
0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	15.9	25.3	33.2	40.3	46.8
2	0.0	20.0	31.8	41.8	50.6	58.8
3	0.0	22.9	36.4	47.7	57.9	67.2
4	0.0	25.1	40.0	52.5	63.6	73.9
5	0.0	27.1	43.1	56.5	68.5	79.6

Figure 10.2: The production function.
Source: ProductionFunction.xls!Technology.

The production function implicitly includes an already solved engineering optimization problem—it gives the maximum output from any given combination of inputs. In other words, we are assuming that the inputs are organized in their most productive configuration and nothing is wasted.

Notice that the Cobb-Douglas function on the Technology sheet has been set up so it can be controlled by a single parameter, α (alpha), by making the exponents α and $(1 - \alpha)$. Use the scroll bar to change alpha and notice how the shape of the production function surface changes. Alpha is a parameter that takes values between zero and one.

STEP Click the Reset button to return the sheet to its default, initial position.

Product Curves

In addition to the 3D view, the production function can be displayed in other ways. To graph the production function in two dimensions, we need to suppress an axis. If we keep output and suppress one of the input axes we get a *total product curve*. If we suppress output and keep the two inputs, we get an *isoquant*.

Product and output mean the same thing. The total product curve is the number of units of output produced as one input is varied, holding the other constant.

STEP Click the TPL and TPK buttons to see the product curves for labor and capital.

In addition to the total product curves, there are average and marginal product curves. The *average product* is simply output per unit of input. Thus, the average product of labor is Y/L and the average product of capital is Y/K.

The *marginal product curves* tell us the *additional* output that is produced as input is increased, holding the other input constant. Marginal product can be computed based on finite-size changes in an input or via the derivative.

Via calculus, the marginal product is simply the derivative of the production function with respect to the input. For the Cobb-Douglas function in the *Technology* sheet, the marginal products are found by taking the partial derivatives with respect to L and K:

$$MP_L = \frac{\partial Y}{\partial L} = (1-\alpha)AK^{\alpha}L^{(1-\alpha)-1} = (1-\alpha)AK^{\alpha}L^{-\alpha}$$

$$MP_K = \frac{\partial Y}{\partial K} = \alpha AK^{\alpha-1}L^{1-\alpha}$$

STEP Scroll down and click on cell C52 to see that the marginal product is computed via the change in output from an increase of 2 hours of labor, with $K = 4$.

This computes the marginal product of labor as the rise over the run from $L = 0$ to $L = 2$ on the total product curve.

STEP Click the Show MPL button and then click on cell C58 to reveal the marginal product computed via the derivative.

Since the total product is a curve, the slope of the tangent line at $L = 2$ is not the same as the rise over the run from one point to another.

STEP Now look at the total, marginal, and average product curves.

Notice how the product curves are drawn based on a given amount of capital. If the amount of capital changes, then the product curves shift.

Marginal and average product can be graphed together because they share a common y axis scale, output per unit of input. The total product curve can never be graphed with the marginal and average product curves because the total product curve uses output as its y axis scale.

The graphs demonstrate that when total product increases at a decreasing rate, marginal product is decreasing. When total output increases at a decreasing rate as more input is applied, ceteris paribus, we are obeying the *Law of Diminishing Returns.* As long as alpha is between zero and one, our Cobb-Douglas production function exhibits diminishing returns.

The Law of Diminishing Returns does not deny that there can be ranges of input use where output increases at an increasing rate. It says that, eventually, continued application of more input along with a fixed factor of production must lead to diminishing returns in the sense that output will increase, but not as fast as before. Thus, the Law of Diminishing Returns is simply a statement that marginal productivity must, eventually, be falling.

As with utility, the Cobb-Douglas functional form is convenient, but there are many, many other functional forms available.

STEP Proceed to the *Polynomial* sheet to see a different functional form. The charts are strikingly different than before.

Unlike the Cobb-Douglas functional form, which always shows diminishing returns, the polynomial production function exhibits all three different phases of returns: increasing, diminishing, and negative returns.

At low levels of labor use, output is increasing at an increasing rate so the total product curve is curved upward and marginal product is increasing. In this range, as long as marginal product is rising and output is increasing at an increasing rate, output rockets upward, growing faster and faster.

When the marginal product curve reaches its peak, the total product curve is at an inflection point. From here, additional labor leads to increases in output, but at a decreasing rate, leveling off as L increases. We say that diminishing returns have set in.

The *Polynomial* sheet is color coded so it is easy to see where the total product curve changes character. Cells with yellow backgrounds signal the range of labor use where diminishing returns apply.

As more and more labor is used, total product reaches its maximum point (where marginal product is zero). Beyond this point, we are in a range of negative returns. This is a theoretical possibility, but not a practical one. No profit-maximizing firm would ever operate in this region because you can get the same amount of output with fewer workers.

It is worth remembering that the Law of Diminishing Returns does not say that we always have diminishing returns for every level of labor use. Instead, the law says that, eventually, diminishing returns will set in. It is also important to understand the difference between diminishing and negative returns. The former says output is rising, but slower and slower, while the latter says output is actually falling.

Notice the relationship between the marginal and average product curves. It is no coincidence that the marginal product curve intersects the average product curve at the maximum value of the average product. There is a guaranteed relationship between marginal and average curves: Whenever the marginal is greater than the average, the average must be rising and whenever the marginal is less than the average, the average must be falling. Thus, the only time the two curves meet is when the marginal and average are equal.

STEP Change the parameter for the b coefficient from 30 to 40.

Notice that the S shape becomes much more linear. The range of increasing returns is larger and we do not hit negative returns over the observed range of L from 0 to 25.

STEP Set the parameter for the b coefficient to 80.

Over the observed range of L from 0 to 25, we see only increasing returns.

STEP Change the δL parameter from 1 to 2. This makes L go up by two and the range goes from 0 to 50.

Diminishing returns do kick in; it just takes more labor for the Law of Diminishing Returns to be observed when the b coefficient is set to 80.

Diminishing versus Decreasing Returns

One extremely confusing thing about the Law of Diminishing Returns has to do with another concept called *returns to scale*. Unlike the Law of Diminishing Returns—which is based on applying more and more of a particular input while holding other inputs constant—returns to scale focuses on the effect on output of changing *all* of the inputs by the same proportion.

There is no law for returns to scale. A production process may exhibit increasing, decreasing, or constant returns to scale, across all values of input use. For example, the Cobb-Douglas function in the *Technology* sheet has constant returns to scale because if you double L and K, you are guaranteed to double output.

You can see this is true by comparing the points 2,2 and 4,4 in the table in the *Technology* sheet. A more complete demonstration uses a little algebra. We begin with the production function:

$$AK^{\alpha}L^{1-\alpha}$$

Next, we double both L and K:

$$A(2K)^{\alpha}(2L)^{1-\alpha}$$

We expand the terms with exponents:

$$A(2^{\alpha})(K^{\alpha})(2^{1-\alpha})(L^{1-\alpha})$$

We collect the "2" terms:

$$A(2^{\alpha+(1-\alpha)}(K^{\alpha})(L^{1-\alpha})$$

The alphas add to zero ($\alpha - alpha = 0$) so we get:

$$A2K^{\alpha}L^{1-\alpha}$$

Thus, we have shown that doubling the inputs from any input levels leads to doubling the output, and this is called constant returns to scale. If the exponents in the Cobb-Douglas function do not sum to 1, then the function does not exhibit this property.

The Cobb-Douglas function in the *Technology* sheet obeys the Law of Diminishing Returns for each input (with $0 < \alpha < 1$), yet it has constant returns to scale. Do diminishing returns imply decreasing returns to scale? No, absolutely not. The two concepts are independent. They ask different questions. The Law of Diminishing Returns is about what happens to output when a single input is increased, ceteris paribus, and decreasing returns to scale says that output will less than double when all inputs are doubled.

Isoquants

In addition to product curves, another way to represent the production function uses the isoquant. The prefix *iso*, meaning equal or the same (as in isosceles triangle), is combined with *quant* (referring to the quantity of output) to convey the idea that the *isoquant* displays the combinations of L and K that yield the same output.

STEP Return to the top of the *Technology* sheet and click the Isoquant button (near cell H28) to see the isoquant map, as displayed in Figure 10.3.

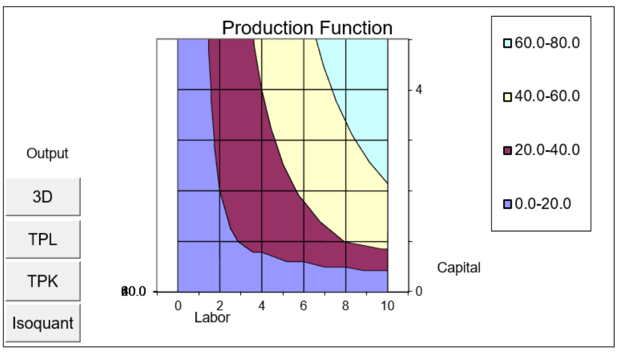


Figure 10.3: Isoquants for a Cobb-Douglas technology.
Source: ProductionFunction.xls!Technology.

An isoquant is simply a 2D, top down view of the 3D surface. Unlike the product curves, which give a view from the side, the isoquant shows L and K on the x and y axes, respectively, and suppresses output.

Notice that Excel cannot correctly draw the isoquant map, putting garbled characters in the bottom left-hand corner of the chart and producing a squiggly, jagged display at the bottom.

You might be thinking that it looks a lot like an indifference map. There are definitely strong parallels between isoquants and indifference curves. Both are top-down views of a 3D object and, therefore, both are level curves or contour plots. Both are used to find and display the solution to an optimization problem.

However, there is one critical difference: unlike an indifference curve, each isoquant is, in principle, directly observable and the isoquants can be compared on a cardinal scale. With indifference curves, the utility function is a convenient fiction and the numerical values merely reflect rankings. No one cares that a particular indifference curve yields 28 utils of satisfaction. This is not the case for isoquants because the suppressed axis, output, is measurable. You can certainly say that one isoquant gives twice the output as another or that one isoquant gives 17 more units of output than another.

One way in which indifference curves and isoquants are the same is that we can compute the slope between two points or the instantaneous rate of change at a point on an isoquant. To avoid confusion with MRS, we call this slope the *technical rate of substitution*, TRS. With labor on the x axis and capital on the y axis, the TRS tells us how much capital we can save if one more unit of labor is used to produce the same level of output.

From one point to another, the TRS can be computed as the rise over the run, $\frac{\delta K}{\delta L}$. At a point, we compute the TRS as the ratio of the derivatives with respect to L and K from the production function:

$$TRS = -\frac{MP_L}{MP_K} = -\frac{\frac{\partial f(L,K)}{\partial L}}{\frac{\partial f(L,K)}{\partial K}}$$

Whereas MRS is universally used for the slope of an indifference curve, MRTS (marginal rate of technical substitution) is sometimes used for the slope of the isoquant. MRTS and TRS are perfect synonyms. We will use TRS.

The TRS (like the MRS) is a number that expresses the substitutability of labor for capital at a point on an isoquant. So, the TRS of two different L and K combinations on the same isoquant might be -100 and -2. The TRS $= -100$ value says that the firm can replace 100 units of capital with 1 unit of labor and still produce the same output. The isoquant would be steep at this point. If a point has a TRS $= -2$, 1 unit of labor can replace 2 units of capital to get the same output. The isoquant at this point would be much flatter than the point with the TRS $= -100$.

Just like the MRS, the TRS tells us how steep the isoquant is at a point. The steeper the isoquant, the more capital can be replaced by labor and still produce the same output.

Technological Progress

Over time, technology—our ability to transform inputs into output—improves. Electric power and computers are examples of technological progress that enables more output to be produced from the same input.

There are two kinds of technological change. The Cobb-Douglas functional form can be used to illustrate each type.

Suppose increased education improves the productivity of labor. This would be modeled as an increase in the exponent for labor in the Cobb-Douglas production function. Small changes, say from 0.75 to 0.751, lead to large responses (e.g., in output or labor use) because we are working with an exponent. This is known as labor-augmenting technological change.

We could also have a situation where the coefficient A in the function $AK^\alpha L^\beta$ increased over time. As A rises, the same number of inputs can make more output. This technological progress is said to be neutral (in terms of the utilization of L and K) because TRS does not depend on A.

We can show this by walking through the steps needed to find the TRS. First, we compute the marginal products of L and K from the function, $Y = AL^\alpha K^\beta$:

$$MP_L = \frac{\partial Y}{\partial L} = \alpha A L^{\alpha-1} K^\beta$$

$$MP_K = \frac{\partial Y}{\partial K} = \beta A L^\alpha K^{\beta-1}$$

The TRS is minus the ratio of the marginal products:

$$TRS = -\frac{MP_L}{MP_K} = -\frac{\alpha A L^{\alpha-1} K^{\beta}}{\beta A L^{\alpha} K^{\beta-1}} = -\frac{\alpha K}{\beta L}$$

The A terms cancel out, which means that the ratio of the marginal productivities of each input depends only on each input's exponent and the amount of the input used.

The Firm as a Production Function

The production function is the starting point for the Theory of the Firm. As with utility, many, many functional forms can be used to represent real-world production processes.

Economists represent the production function not as a 3D object, but in two dimensions. We get product curves (total, marginal, and average product curves) by focusing on output as a function of a single input, holding all other inputs constant. An isoquant suppresses the output and shows the different combinations of L and K that produce a given level of output.

The TRS is similar to the MRS, and it will play an important role in the understanding the firm's cost minimizing input choice.

Remember to keep straight the difference between the Law of Diminishing Returns and idea of returns to scale. The former applies more and more of a single input, holding all other inputs constant; the latter reports what happens to output when all inputs are changed by the same proportion. Those are two different things.

Exercises

1. Starting from a blank workbook, with $K = 100$, draw total, marginal, and average product curves for $L = 1$ to 100 by 1 for the Cobb-Douglas production function, $Q = L^{\alpha}K^{\beta}$, where $\alpha = 3/4$ and $\beta = 1/2$. Use the derivative to compute the marginal product of labor.

   Hint: Label cells in a row in columns A, B, C, and D as *L, Q, MPL,* and *APL*. For *L*, create a list of numbers from 1 to 100. For the other three columns, enter the appropriate formula and fill down. For *MPL*, do not use the change in Q divided by the change in L; instead enter a formula for the derivative for the *MPL* at a point.

2. For what range of L does the Cobb-Douglas function in question 1 exhibit the Law of Diminishing Returns? Put your answer in a text box in your workbook.

3. Determine whether this function has increasing, decreasing, or constant returns to scale. Use the workbook for computations and include your answer in a text box.

4. From your work in question 3 and the comment in the text that you cannot have constant returns to scale "if the exponents in the Cobb-Douglas function do not sum to 1," provide a rule to determine the returns to scale for a Cobb-Douglas functional form.

5. Is it possible for a production function to exhibit the Law of Diminishing Returns and increasing returns to scale at the same time? If so, give an example. Put your answer in a text box in your workbook.

6. Draw an isoquant for 50 units of output for the Cobb-Douglas function in question 1.

   Hint: Use algebra to find an equation that tells you the K needed to produce 50 units given L. Create a column for K that uses this equation based on L ranging from 20 to 40 by 1 and then create a chart of the L and K data.

7. Compute the TRS of the Cobb-Douglas function at $L = 23$, $K = 312.5$. Show your work on the spreadsheet.

References

The epigraph comes from page 152 of "A Theory of Production" by Charles W. Cobb and Paul H. Douglas, *The American Economic Review*, Vol. 18, No. 1, Supplement, Papers and Proceedings of the Fortieth Annual Meeting of the American Economic Association (March, 1928), pp. 139–165, www.jstor.org/stable/1811556.

Douglas, an accomplished professor and US Senator from Illinois, explained how he and Cobb used the functional form that would be named after them:

> I was then temporarily lecturing at Amherst College, and consulted with my friend and colleague, Charles W. Cobb, a mathematician. At the latter's suggestion, the formula $P = bL_kC_{k-1}$ was adopted, a form that had also been used by Wicksteed and Wicksell.

See p. 904 in Paul H. Douglas, "The Cobb-Douglas Production Function Once Again: Its History, Its Testing, and Some New Empirical Values," The *Journal of Political Economy*, Vol. 84, No. 5 (October, 1976), pp. 903–916, www.jstor.org/stable/1830435

Chapter 11

Input Cost Minimization

> The term "isoquant" was introduced by R. Frisch but originally for a different concept, for which it should have been reserved.
>
> Joseph Schumpeter

11.1 Initial Solution

Input cost minimization is one of the three optimization problems faced by the firm. It revolves around the question of choosing the best combination of inputs, L and K, to produce a given level of output, q.

The best combination is defined as the cheapest one. The idea is that many combinations of L and K can produce a given q. We want to know the amounts of labor and capital that should be used to produce a given amount of output as cheaply as possible.

Of course, we answer this question by setting up and solving an optimization problem; then we do comparative statics. Because there is a constraint (we must produce the given q), we will use the Lagrangean method.

Setting up the Problem

The economic approach organizes optimization problems by answering three questions:

1. What is the goal?
2. What are the choice variables?
3. What are the given variables?

The goal is to minimize *total cost*, TC, which is simply the sum of the amount paid to the workers, wL, and the amount spent on renting machines, rK.

The endogenous variables are L and K. Labor is measured in hours and capital is the number of machines. The firm can decide to produce the given output by being labor intensive (using lots of labor and little capital), or roughly equal amounts of both, or by renting a lot of machinery and using little labor.