LUÍS M B CABRAL

THE ECONOMICS OF ENTERTAINMENT

This book was typeset in LaTeX using the memoir style as well as multiple macros developed by the author. All images included in this book are in the public domain. Whenever appropriate, links to the license and author are included.

Cover design based on *Misty III* (2015), oil on canvas (18x18in), by Luís Cabral.

Luís Cabral is Paganelli-Bull Professor of Economics at NYU's Stern School of Business. For up-to-date information about this book, visit luiscabral.net/economics/books/entertainment/

Version 0.0.2 (September 13, 2023)

CONTENTS

CHAPTER 1
INTRODUCTION

1.1. WHAT IS SPECIAL ABOUT THE ECONOMICS OF ENTERTAINMENT

As one reads the title of this book, *The Economics of Entertainment*, the natural question to ask is: Why? Why should we care about the economics of a specific industry? Economists don't talk much about the economics of chocolate or the economics of titanium dioxide, to give the example of two multi-billion dollar industries. Why then entertainment, specifically the *economics* of the entertainment industry?

Each industry possesses certain characteristics which, while not necessarily unique, contribute to the distinctive nature of its business. What is special about the entertainment industries? Do the basic principles of economics apply to movies, popular music, professional sports? In some sense, the fields of economics and entertainment could not be farther apart: the latter is about cheering up people, whereas the former is often described as the dismal science. But as I will argue next, while it would be foolish to ignore the idiosyncratic features of the entertainment industries, it would be equally wrong to ignore the rich economics they are made of. (I should recognize that other authors have a different view on the matter.)

WikiCommons

George Martin, a leading talent in the art of talent management.

MANAGING CREATIVE TALENT

Most entertainment offerings result from the confluence of various creative talents: from the scriptwriter to the movie director to the actor; from the songwriter to the composer to the singer; and so forth. It is no exaggeration to say that frequently artists have a high opinion on themselves and are easily irritable. Managing such source of creative value is no small task.

The concept of prima donna originates in opera, and opera is one of the best examples of the difficulties of dealing with difficult people whose contribution is essential for the running of an operation. Those who can manage such a daunting task are nothing less than management hall of famers. Take Joseph Volpe, general manager of the New York Metropolitan opera from 1990 to 2006. Some criticized him for the "dictatorial" way in which he handled artists: "Superstar singers are pampered and granted wide-ranging artistic control, but anyone who doesn't guarantee a sold-out house is treated with a shocking disdain" (source). Be that as it may, most agree that, under Volpe's tenure, the Met solidified its position as one of the leading opera companies worldwide. He was a model of efficiency.

Take another example from the music world. Producer George Martin is frequently referred to as "the fifth Beatle," and not without reason: Martin crafted song structures, wrote beginnings and endings, harmonies and solos (*All You Need Is Ears* is the title of his autobiography). But equally important, he knew how to deal with each band member according to his unique personality. He would sit at the piano to work with McCartney; do his best to accommodate Lennon's vague requests ("I want you to score this for me"); attempt to encourage Harrison's songwriting (thought admittedly too late); and generally try to encourage the always underrated and somewhat under-appreciated Starr. "I taught them a few tricks and they were very quick to learn — like hothouse plants, they just sprung up" (source).

One could go on and on giving examples of why people are important and special — and difficult to deal with. One of my favorite stories is that of *The Third Man* (1949), the classic *film noir* set in post-war Vienna. The fact the movie was finished at all is a miracle of sorts, considering all the conflicts that arose among the various parties involved: screenplay writer Graham Greene and producer David Selznick clashed over the choice of an ending; director Carol Reed had difficulty in getting lead actor Orson Welles to set a time for shooting in Vienna; and then, when Welles finally arrived, the actor refused to continue filming the scenes set in the Vienna sewers, an important part of the thriller. Carol Reed, besides being a first-class director, should also be credited for his patience and tact in dealing with such a group of strong egos. He was also humble enough to welcome Welles' multiple suggestions and ad-libs, including the now-famous "cuckoo clock" speech. Finally, Reed also showed his ability to spot talent when he heard zither player Anton Karas at a Vienna party and eventually lead him to work on the movie's soundtrack, including the unforgettable "The Third Man Theme."

To sum up: It is a common mistake — and a frequent one among economists — to take labor as just any other generic production input. In the case of creative industries, it is a particularly big mistake.

BUYERS, SELLERS AND ENGAGEMENT

In a certain sense, a star is a seller who supplies a service (entertainment) to fans. But the relation between stars and fans goes well beyond that of buyer and seller. Suppose that ACME Widgets — a company you know nothing about — decides that, henceforth, each customer will pay whatever he or she feels like. It's an easy guess that the widget supplier will quickly file for bankruptcy.

But witness the experiment by British rock band Radiohead: their seventh studio album, *In Rainbows*, was released online on October 10, 2007. The price? — Whatever each downloading fan decided was fair. While exact figures are not available, some analysts estimate

Wikipedia

Dumbest Moment in Business? — Not likely.

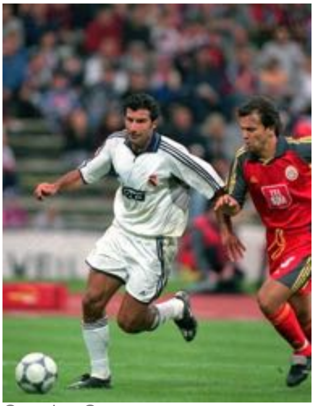
that 1.2 million copies were sold, well above the 200–500k range of previous releases. *CNN Money* found Radiohead's move worthy of their "101 Dumbest Moments in Business" list, but the math suggests that the band may have come out well ahead.

The comparison between widgets and songs is not entirely fair: the production cost of an *additional* music download is zero, whereas that of a widget is not. A better comparison might be between Radiohead and a software developer (since the seller's cost of an additional software download, like music, is zero). Would downloaders be willing to pay for free software if they were offered the same deal as Radiohead? Probably not.

In Chapter 2 we will look at additional examples of the general principle that the relation between star and fan is more complex than the relation between buyer and seller. While the examples are all from the music industry, the idea extends to other entertainment industries, including movies and professional sports.

IT'S NOT THE MONEY, IT'S THE PRINCIPLE

David Puttnam will not win any popularity contest in Hollywood, but he is unquestionably one of the industry's most influential figures. In his *Undeclared War*, Puttnam sets out his thesis that movies and TV shape attitudes and values, create conventions of style and behavior. For this reason, he argues, creative artists share a heavy moral responsibility to inspire and affirm, not just to entertain. For example, Puttnam stated that "films featuring violence and aggression devoid of human consequences lead to the growth of bullying


Creative Commons

In one of the most controversial sports transfers ever, Portuguese soccer legend Luís Figo left FC Barcelona for Real Madrid Madrid in the summer of 2000 for what was then a world record transfer fee. "For me [joining] Madrid was a new stage. I joined them to earn more in every way — prestige, titles and money — and I got them." Barcelona fans, however, were furious and let him know it: on Figo's first re-appearance in Barcelona, the crowd's jeers and missiles led the referee to suspended the match for 13 minutes, fearing for the players' safety.

in the playground, with children imitating what they saw on the big and small screens" (source). More generally, he blames the entertainment industry for creating a world of "moral ambiguity." In other words, the movie and TV businesses are not just like any other business. "Films must transcend profit and loss." They should make a profit "but profit isn't enough. We do have other responsibilities." (source) Nor is Puttnam just talking the talk, as shown by an impressive list of financial and artistic hits such as *Chariots of Fire*, *The Mission*, and *The Killing Fields*.

The idea that the bottom line is not the bottom line extends beyond beyond movies and television. In professional sports, a related issue is that of team loyalty. When F.C. Barcelona star Luís Figo left for archrival Real Madrid — thus becoming the first "Galatico" — he immediately became persona non grata at Nou Camp — or worst than that, some would argue. The argument that the deal made of lot of business sense seemed irrelevant — at least to his former Catalan fans.

The conflict between financial and other objectives may also imply a conflict between the management and the creative agents of a business. Consider for example the problem of monetizing newspaper content. In 2011, *The New York Times* re-introduced a "paywall" limiting access to its content. This is probably the best business solution for a newspaper struggling with declining readership of its paper edition; but is it the best outcome for content creators, such as the newspaper's columnists? Thomas Friedman, for example, has an interest in the *Times*' financial success, just as every employee is concerned with his or her employer's economic viability. But Friedman is equally or more interested in attracting a wide audience, and raising a paywall may considerably limit readership.

TECHNOLOGY AND CONTENT

History has shown that the relation between technology and content is not always what it seems at first. In 1982, as the consumer video-cassette recorder (VCR) became increasingly popular, Jack Valenti, then president of the Motion Picture Association of America, declared that "the VCR is to the American film producer and the American public as the Boston strangler is to the woman home alone."

Since the days of the video-cassette recorder, various entertainment-relevant technological developments have taken place, including in particular the Internet. Together with the trend towards greater globalization, these have changed (and continue changing) the business models of the music industry, the movie industry, professional sports, and so forth. Have the new combinations of technology and content increased or decreased value? How do the various players stand to gain or lose from technology change: hardware vs software, talent providers vs promoters, small stars vs superstars? These are some of the important open questions in many entertainment industries (and media industries, one might add).

For example, in Chapter 8 we consider the impact of digital technologies and the Internet on the evolution of the music industry. Some argue that the 2000s were the industry's lost decade: CD sales, for example, dropped by more than 50% from 2000 to 2010. But if we consider all of the music related revenues — including digital downloads, concerts, merchandizing and in particular music related hardware —, then we observe a healthy growth rate throughout the entire decade. We thus have a case of value shift rather than one of value destruction.

COMPETITORS AND COMPLEMENTORS

Co-opetition, the business best-seller by economists A. Brandenburger and B. Nalebuff, popularized the concepts of competitors and complementors. All too frequently, the authors argue, we think of business strategy as an attempt to capture the greatest slice of the pie. But more often than not what matters is how different players within the same industry come together to *increase the size* of the pie. When that happens, they act as complementors — even if, strictly speaking, they are market competitors.

The phenomenon of complementors is not unique to entertainment industries (a classical example is that of Microsoft and Intel). However, most entertainment industries feature one or more instances of complementarities of the sort described in *Co-opetition*. Take for instance the *Lion King* franchise. In addition to the movie, there is the Broadway play, the theme park attractions, the book, and a whole series of merchandizing products. To the extent that these are owned by different players, then we have a situation of complementors. When a promotional effort by the Broadway play leads to increased ticket sales, many other Lion King-related businesses improve as well. The trick is then to arrange things so that the various players have the right incentives to contribute to their common good.

Professional sports leagues are another important instance of complementors. There is a clear and obvious way in which the different teams in a given league are competitors: when F C Barcelona and Manchester United meet on the field, they play a zero-sum game ("my gain is your loss"). But there are many business and sports issues on which the two teams' interests are aligned: for example, the design and governance of the Champion's league or a possible European super-league.

WHAT IS *NOT* SPECIAL ABOUT THE ENTERTAINMENT INDUSTRY?

Having argued that the entertainment industries are in some respect unique and special (and closely related to each other), I should now add that there are also many ways in which the entertainment industries are very similar to other industries. For example, the distribution of movie revenues has been shown to be quite skewed, with quite a few blockbusters that fall outside Normal distribution patterns. Box-office "fat tail" revenue distributions turn out to be very similar to those of pharmaceutical drugs. In fact, there is a certain similarity between a medical drug's creative process and that of a motion picture, both in terms of the talent it entails and the outcome uncertainty it leads to.

A second example is given by music and beer (two industries that are closely related in the eyes of the consumer). Some say that the music industry is evolving in the direction of a greater number of very small artists (“the long tail”) and a small number of increasingly big artists (“superstars”). If we look at the evolution of the beer industry, we observe (in the U.S., perhaps also in other countries) an increase in size of the top players but also the emergence of a large number of very small producers — just as in the music industry. Can the parallel be extended?

A third example is given by pricing. In 2003, the New York Mets, a Major League Baseball (MLB) team, introduced variable pricing. Until then, all tickets for a certain stadium seat were priced the same regardless of the game being played. Beginning with the 2003 season, “better” games were classified as “gold” and priced accordingly, followed by “silver,” “bronze” and “value” games. Beginning with the 2009 season, the San Francisco Giants, another MLB team, introduced dynamic pricing, whereby a given game’s ticket prices are adjusted in “real time” according to the starting pitcher, the opponent’s current standing, etc.

Anyone flying on a commercial airline should be familiar with these pricing strategies, both in terms of quality tiers and in terms of dynamic adjustment. Can sports teams learn from airline pricing? The same might be said of rock concert organizers. And are movie theaters leaving money on the table by setting uniform prices (that is, prices that do not vary as a function of each movie’s popularity)?

In summary, the study of entertainment industries is greatly enriched by the knowledge of industries such as air travel or pharmaceuticals. Conversely, the knowledge gained by studying the business of entertainment may help understand the workings of other industries as well. There is an element of generality in economics that transcends the particular industries it is applied to.

1.2. ECONOMICS

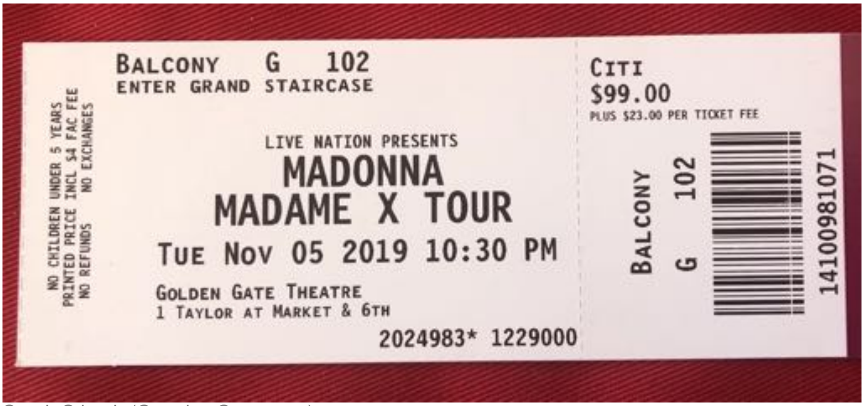
I expect this book to be of some interest to readers with an economics background who are interested in learning more about the entertainment industries, as well as entertainment industry readers who are interested in learning more about economics. This section is primarily written for the latter. The idea is to go over some of the central concepts in economics, the concepts that define the economics way of thinking. Abraham Maslow famously stated that "if all you have is a hammer, everything looks like a nail." Economists are no exception: we, too, have our own biased way of looking at reality. What are the concepts that make up the economics "hammer"?

MARGINAL THINKING

As often happens with other disciplines, economics is more than a series of ideas: it's a way of thinking about things. Non-economists may have noticed it — and be annoyed by it — when talking to economists. First, economists have a tendency to think about problems strictly in terms of costs and benefits (monetary or otherwise), i.e., to follow a **cost-benefit** approach.

Let us consider a specific example. On May 2019, Madonna's Madame X Tour was officially announced, with concerts starting in September 2019. One important decision for Madonna is how many tour performances to include. Any time there is a "how many" question, economists think about it in terms of **marginal** variations. Suppose the current plans call for 84 dates (it did as of January 2020). Is this an optimal number concerts? To answer the question, Madonna should evaluate the benefits (monetary and otherwise) from an additional concert, both in terms of ticket sales and other related revenue stream. This should then be compared against the cost (monetary and otherwise) of an additional concert. If cost is great than benefit, go for it; if not, don't.

But there is more: Suppose that Madonna really thinks like an economist. Then it must be the case that 84 concerts is the number such that the benefit of an additional concert is about the same as the cost of an additional concert. Why? Well, if the benefit of the 85th concert is greater than the cost, then Madonna should extend the tour to 85 dates (at least). By contrast, if the benefit is lower than the cost, then it's also likely the case that the benefit of the 84th concert itself was lower than the cost, in which case Madonna would be better off by shorting the tour to 83 dates (or less). We will return to this issue in Chapter 4, when discussing issues of optimal pricing.



Sarah Stierch (Creative Commons)

How many shows should Madonna include in her Madame X tour? She should think at the margin, an economist would say.

OPPORTUNITY COSTS AND SUNK COSTS

Suppose I own Maddison Square Garden (MSG), New York's famous multi-purpose indoor arena. I am considering hosting one of Madonna's Madame X concerts. As per the previous paragraphs, I compare benefits (expected ticket sales, etc) to costs. Since I own MSG, I don't need to pay anyone to use the space, that is, there is no monetary payment involved with using the space. However, for an economist there is a cost that should be taken into account, namely the **opportunity cost** of not using the space for an alternative purpose. Suppose I have the option of renting out MSG on the same night when Madonna is available to perform. A different tour operator is willing to pay me $300,000 for one night at MSG. Then I should include $300,000 as part of my cost of hosting the Madonna concert.

The concept of opportunity cost suggests that economists have a tendency to see costs where others don't. Well, the opposite is also true, that is, there are cases when others see costs where economists don't. Such is the case of **sunk costs**. In 2019, Benfica Lisbon, one of Europe's leading soccer teams (my mom's a big supporter), recruited Spanish player Raul de Tomas (a.k.a. RDT) for the hefty transfer fee of $22.8m. RDT did not adapt well to Benfica's playing style and did not get much playing time. At a press conference, one journalist challenged Benfica coach's decision to bench a player who cost the club so much money. An economist would respond by saying that the decision of whether to field RDT is a comparison of benefit and cost, and RDT's transfer price is an irrelevant cost. It's sunk: regardless of whether RDT plays or does not not play, you will have to pay the transfer fee (in fact, you already paid it). And if it's sunk, it's irrelevant. You may regret having paid all of that money for a player that's not being used, but you should not let the transfer fee affect your choice of the best starting team.

Baseball provides another example of the so-called sunk cost fallacy. Chris Davis, a 33-year-old slugger for the Baltimore Orioles, went through a long stretch of hitless at-bats in 2018, the worst batting average in major league history (.168). Still, the Orioles manager kept fielding him.

> In truth, the decision to keep playing Davis almost certainly has more to do with his $17 million salary this year and the $93 million the Orioles owe him beyond 2019 in salary and deferred payments, which will have the team sending him paychecks through the 2037 season. (source)

The point, again, is that the money the Orioles owe Davis is a sunk cost. As such, it should be irrelevant to the decision of whether to field him. The fact he was fielded suggests that not all behavior is consistent with basic economic thinking. This may not surprise you, but it continues to puzzle economists, especially in situations when the stakes are high.

DECREASING MARGINAL BENEFIT

To be completed

VALUE IN USE AND MARKET VALUE

To be completed

GAINS FROM TRADE

In March 2019, Disney acquired 21st Century Fox from the Fox Corporation in a deal valued at $71.3 billion. (In January 2020 Disney dropped the name "Fox" from the acquired assets so as to avoid confusion with the Fox family of TV channels.) Several analysts have declared the merger to be a game-changer for Hollywood. According to Bob Iger, Disney's CEO,

> The acquisition of this stellar collection of businesses from 21st Century Fox reflects the increasing consumer demand for a rich diversity of entertainment experiences that are more compelling, accessible and convenient than ever before. (source)

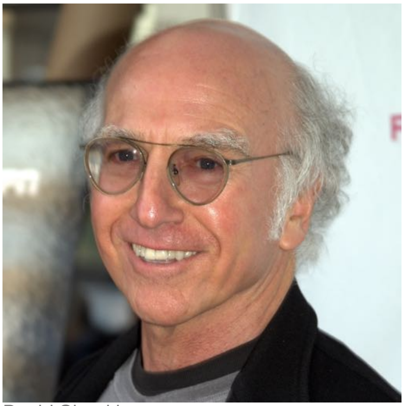
One way to rephrase Iger is say that the 21st Century Fox assets create more value in the hands of Disney than in the hands of Fox. This may result from a series of factors. Optimists will say that the acquisition will allow Disney to offer a better service to viewers. Pessimists will say that this is the beginning of a major consolidation of the major studios which will eventually hurt consumers. Whichever is the case, transactions like this illustrate the very important economics principle that exchange creates value, that is, there are **gains from trade**. All too often we think of transactions as a **zero-sum game** ("my gain is your loss"). This may be true in chess and other such games, but not in most business situations.

COMPARATIVE ADVANTAGE AND SPECIALIZATION

Adam Smith, arguably the founder of the discipline of economics, observed a fairly obvious fact: Scotland produces excellent wool but horrible wine. France, in contrast, produces horrible wool and excellent wine. This clearly calls for an exchange of Scottish wool for French wine, thus making both countries happier with what they wear and drink.

So far, this is an example of the previous point, namely that exchange creates value. Fast forward a few decades and David Ricardo refines Smith's idea with a more extreme example: Portugal produces better wine than England (easy), but it is also more efficient at producing textiles (i.e., does so at a lower cost).

Put this way, it does not seem there is much room for value-creating exchange. But there is. The genius of Ricardo is to develop the concept of **comparative advantage**: What matters is not whether Portugal is better than England at producing wine or textiles. What really matters is which activity Portugal is *relatively* better at. As it happens — not very surprising — Portugal was better than England at producing textiles but *much better* than England at producing wine.


David Shankbone

Some people argue that Larry David is a better actor than Jerry Seinfeld. Should he have taken a more active acting role in *Seinfeld*?

It follows that the two countries can jointly create value by exchanging English textiles for Portuguese wine.

To give a more recent and more entertainment-like example, consider artists Larry David and Jerry Seinfeld, who created the hugely successful show *Seinfeld*. David's contribution was primarily in writing. Some people argue that David is a better actor than Seinfeld. Witness, for example, his work in the also very successful show *Curb Your Enthusiasm* — or, more recently, his SNL appearances impersonating Senator Bernie Senders. Should David then have taken a lead role in *Seinfeld*? An economist might answer that the relevant question is not whether David is a better actor than Seinfeld, rather whether he is *relatively* better at writing or acting. In other words, if David is a better actor but also a much better writer than Seinfeld, then it makes sense for David to spend more time writing than acting (and Seinfeld the opposite).

CAUSAL INFERENCE

One of the pitfalls of doing social science is that there is no lab where you can control for everything (or almost everything) and get a clear picture of the causal effect of X on Y. In economics we observe X and Y happening in sequence. Does that mean that X caused Y or is it simply a case of two variables that are correlated? For example, in January we lower the price of a videogame we sell, and in February we observe increased sales. Was it the drop in price? Another example: During Season 2, a given TV show's slot is moved from Tuesday to Wednesday. Audiences drop by 20%. Was it the change in the time slot or was it the quality of the content during the new season?

Economists are obsessed with the issue of correlation vs causality — an issue often referred to as **causal inference** — and get very upset when researchers in other fields don't do the same. Consider for example an article published in a 2012 issue of the prestigious *New England Journal Of Medicine*. The authors showed that there is a strong cross-country correlation between per-capita chocolate consumption and the per-capita number of Nobel laureates (Switzerland comes out on top of both rankings). This was not a spoof, this was a bona fide scholarly article. The journal's editor remarked that

> Chocolate consumption could hypothetically improve cognitive function not only in individuals but also in whole populations. The principal finding of this study is a surprisingly powerful correlation between chocolate intake per capita and the number of Nobel laureates in various countries. Of course, a correlation between X and Y does not prove causation but indicates that either X influences Y, Y influences X, or X and Y are influenced by a common underlying mechanism. (source)

To the editor's credit, he was careful to make the distinction between correlation and causality, but he omitted the fourth and most likely explanation for the correlation: the sample is too small. If you have time and would like to understand this while having some fun, then direct your browser to tylervigen.com. There you will find a plethora of correlated time series such as the one in Figure 1.1. It's hard to believe that there is *any* causal relation between the number of sociology doctorates awarded in US in a given year and the number of worldwide non-commercial space launches. It's also hard to come up with a theory by which "X and Y are influenced by a common underlying mechanism." How come we get such a high correlation?

There is a theorem by a famous statistician which essentially states that if you are given a fixed number of observations of a large enough number of random variables, and if you consider all of the possible combinations between these variables, and if you then select the pair with the highest correlation, then you will very quickly come up with a very high correlation indeed. That's what the site www.tylervigen.com does, obtaining in the process a series of funny **spurious correlations**.

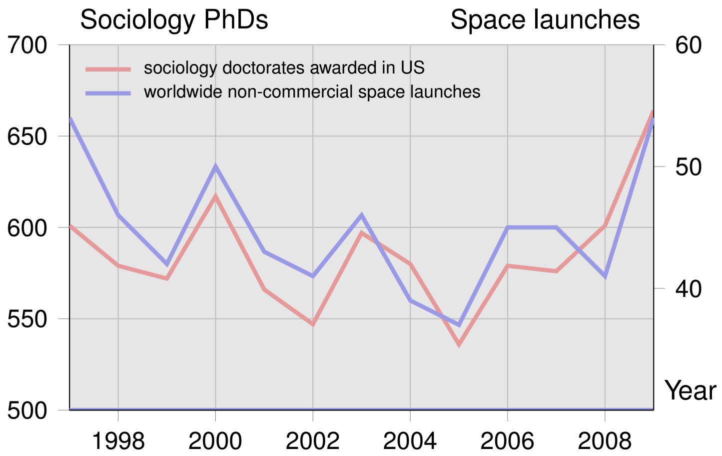


FIGURE 1.1
Spurious correlation

Economists deal with this sort of problem in a variety of ways. If you can actually run a controlled experiment, then this might be your best bet. For example, change the price set for a randomly selected set of buyers but not for a different set of buyers (the control group). This is what drug companies do to test the efficacy of a new drug, for example. Unfortunately, in many if not in most cases this is not possible, and all we have is a bunch of historical data to work with. In Chapters 2 and 3 we will consider specific examples of how to go about data analysis with the goal of causal inference.

KEY CONCEPTS

cost-benefit | marginal | opportunity cost | sunk costs

gains from trade | zero-sum game

comparative advantage | causal inference

spurious correlations

REVIEW AND PRACTICE PROBLEMS

■ **1.1. Entertainment.** What is special about the economics of entertainment industries?

■ **1.2. Co-opetition.** What do we mean by co-opetition? Provide an example from an entertainment industry.

■ **1.3. Marginal thinking.** What do economists mean by "marginal thinking"?

■ **1.4. Opportunity cost and sunk cost.** What is the difference between opportunity cost and sunk cost?

■ **1.5. Gains from trade.** What do economists mean by "gains from trade"? Provide an example from an entertainment industry.

CHAPTER 2

DEMAND

Creating Demand: Move the Masses to Buy Your Product, Service, or Idea (Updated for the 21st Century). You will find this and many other such titles at your local bookstore (if it still exists, which is an interesting demand-related question). If there is no demand, then there is no business. Therefore, before deciding what to do in business it's important to have some knowledge of what the demand for your product is. Our goal in this chapter is to introduce some key concepts regarding demand in general and demand for entertainment goods in particular.

2.1. DEMAND FUNCTION

Our first concept is that of a **demand function**. It is given by

$$q = D(p, X, A, Y) + \epsilon$$

In the above equation, $D(p, X, A, Y)$ is a function, a mathematical object that has (in the present case) p, X, A and Y as "inputs" and produces a value (in this case a value of q as an "output." (If you are not familiar with the concept of function, you may want to check the appendix on functions.)

As to the variables in the above equation, q is quantity demanded (e.g., # tickets sold); p is price; X corresponds to a set of product characteristics (e.g., talent quality); A measures marketing variables such as advertising, branding, etc; Y refers to income and other consumer demographics (average age, gender, etc).

Finally, ϵ measures uncertainty, in other words, stuff that we cannot easily measure. As an aside: Artificial intelligence is changing many aspects of our lives. One important way is improving our ability to predict the demand for entertainment goods, that is, AI is reducing the relative importance of ϵ with respect to X. Recommender systems, for example, are ways of predicting consumer demand, thus avoiding the unpredictability element embodied in the variance of ϵ.

The underlying idea is that, if we have enough information about the nature of a product and about its consumers, then we should be able to predict how much it will be in demand, that is, how much of it consumers will be interested in purchasing. For example, suppose I want to predict how many tickets will be sold for a New York Mets vs Washington Nationals game next Saturday evening. Based on ticket sales for past games, I have an idea of how ticket sales are determined by the opposing team, whether it's a weekday or a weekend, etc. By plugging in all of these variables, the function D gives me the desired number. This is not an exact science, but as we will see in Chapter 5 (and later in this chapter) there is actually a lot that can be known about the value of q.

The concept of the **demand curve** is related to that of the demand function. Fix everything but p, and assume that $\epsilon = 0$. The resulting relation between q and p is given by $q = D(p, \overline{X}, \overline{A}, \overline{Y})$, which we call the demand curve. (The bar over the variables means that they are fixed at a certain value.)

■ **The law of demand.** The Law of Demand is one of economics' most important results. It simply states that ***quantity demanded falls as price increases***. There are several reasons why the law of demand works, including the availability of substitutes and limited budget. For example, if the monthly subscription of Spotify were to increase dramatically I would consider alternative music services, or switching to the ad-supported free version, or simply going back to buying CDs. Even if there were no alternatives, if price were to increase dramatically I would at some point need to give up: there are other things in life one needs to spend money on.

Figure 2.1 illustrates the concept of a demand curve. For, suppose that Taylor Swift is giving a concert at Madison Square Garden, and

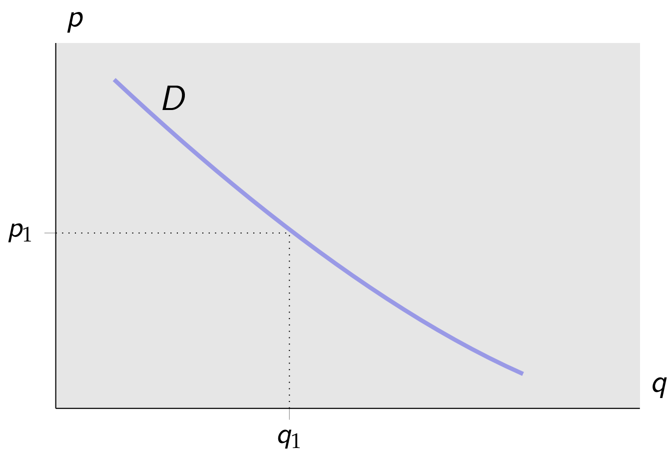


FIGURE 2.1
Demand curve

suppose D depicts the relation between ticket price p and the number of tickets sold, q. We then refer to D as the demand curve for Taylor Swift tickets. Note that normally we represent the demand curve with p on the vertical axis. Most people familiarized with functions would expect price (the explanatory variable) to be on the horizontal axis. If you find it confusing that p is on the vertical axis, don't blame me, blame 19th century economist Alfred Marshall, who first had the idea. Note also that, as per the law of demand, D is a downward sloping line: the greater p, the lower q, and vice-versa.

■ **Luxury goods.** Are luxury goods (e.g., fashion items) an exception to the law of demand? I get that question a lot. The answer is: not really. The source of confusion is that we associate the idea of luxury with the idea something really exclusive, which in turn we associate with high prices. For example, suppose there is this high-end club in New York City. It's super expensive to get in and everyone wants to be there (high demand), though few can afford. You might think that, were the price really low, then its exclusivity would vanish and so would demand. This suggests an upward sloping demand curve. However, the real test of the law of demand would be something like the following conceptual experiment:

> I'm offering you admission to the high-end, very exclusive club for a considerably lower price (say, 1% of the

regular price). Moreover, I guarantee that you are the only person being offered the deal. Moreover, no one will know that you were offered this deal.

Would you be more likely or less likely to join the club at the super low price? I would argue more likely. If that's the case, then the demand curve is downward sloping, as the law of demand predicts. (As an aside: There is one historical exception to the law of demand: potatoes in Ireland during the 19th century famine. The law of demand is not an absolute law, rather an empirical regularity so often observed that we call it a law.)

■ Inverse demand. There are two ways of reading the demand curve:

(a) for a given p, how many units will be sold. For example, if the price of the Taylor Swift concert is set at p_1, then q_1 tickets will be sold. This is the "normal" way of reading a demand curve.

(b) for a given q, what is the willingness to pay for the last unit. For example, the buyer of the q_1th ticket was willing to pay up to p_1 for it. When we read the demand curve in this way, we refer to it as the **inverse demand curve**.

■ Demand curve and demand function. It's important to distinguish the demand function from the demand curve. In the demand curve, q depends on p. In the demand function, q depends on p and on many other factors, such as the prices of substitutes or complements; population and income; advertising; and lots of other things.

Consider, for example, the demand for Netflix. The demand curve tells me how the number of subscribers depends on the monthly fee. The demand curve, by contrast, includes other factors in addition to price. Important factors might include what hit shows Netflix is currently carrying, the price of the HBO service, whether the economy is doing well, etc.

Similarly, it's important to distinguish between:

(a) movements along the demand curve (changes in price)

(b) shifts in the curve itself (changes in other factors).

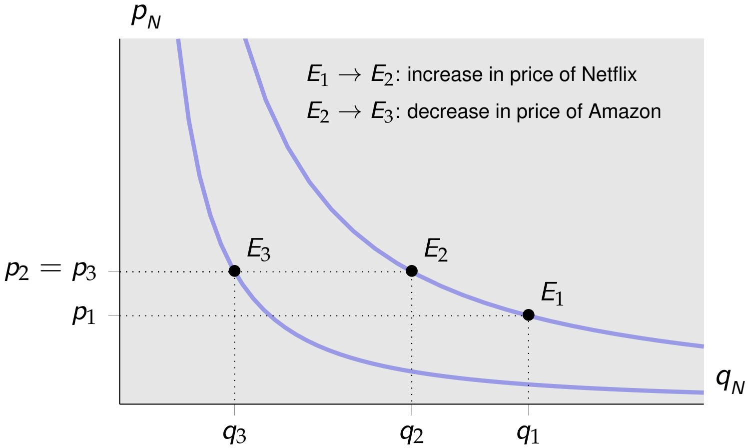


FIGURE 2.2
Example: Amazon and Netflix

Figure 2.2 illustrates this distinction. Suppose that initially (time 1) Netflix is priced at p_1 and there are q_1 subscribers. Now suppose that, at time 2, Netflix decides to increase is monthly fee. As a result, quantity demanded drops to q_2. This change corresponds to a *movement along the demand curve*. Now suppose that the price of Amazon Prime decreases. Some consumers will decide that Amazon Video is sufficient for their entertainment needs and thus cancel their Netflix subscription. This implies that, at time 3, even though the price of Netflix has not changed the quantity demanded of Netflix declines from q_2 to q_3. This change in q corresponds to a *shift in the demand curve*.

SPILLOVER EFFECTS

Entertainment goods are unique; and there are a great many of them. For example, the number of movies released as DVDs during the 2000s was of the order of 25,000. As such, **awareness** is a major issue: a typical consumer is not aware of the existence of all of the products.

Awareness is important when explaining an important phenomenon in music and movies: **demand spillovers**. A hit album by a given band typically implies increased sales by the same band's older albums. For example, research shows that when *Jagged Little Pill* by Alanis Morissette was released in 1995, not only was it a suc-


Discutivo

When *The Vow*, starring Rachel McAdams, was a hit at the box office, sales of older DVDs starring McAdams increased.

cess but it also led to an increase of *Alanis*, another album by Alanis Morissette, released in 1992.

Similar patterns are observed in movies. *The Vow* (2012), starring Rachel McAdams, was a big box-office success. Soon after it hit the theaters, sales of the *Wedding Crashers* (2005) DVD, also starring Rachel McAdams, increased considerably.

There are several possible narratives for these backward spillover effects. One is the awareness story. Once you find a great album you go and look for other albums by the same band. A second, related one, is that, although you knew about the band and its previously released albums, now that you heard a good album by them you update your beliefs regarding the value of the old ones, and accordingly go buy them. Third, it may simply be an instance of the **halo effect**, the tendency for an impression created in one area to influence opinion in another area.

SOCIAL GOODS

A final note regarding demand for entertainment goods relates to the fact that frequently these are social goods: I bought a copy of Harry Potter because everyone around me was reading Harry Potter, and if I didn't I would be left out of most conversations. When these social effects are strong, there is the possibility that demand will be multi-valued: for a given price, there are two possible demand values. This corresponds to the artist or work of art being "in" (high demand level) or "out" (low demand level).

Figure 2.3 illustrates this point. The panel on the left corresponds to a "normal" demand curve: the lower price is, the greater the quan-

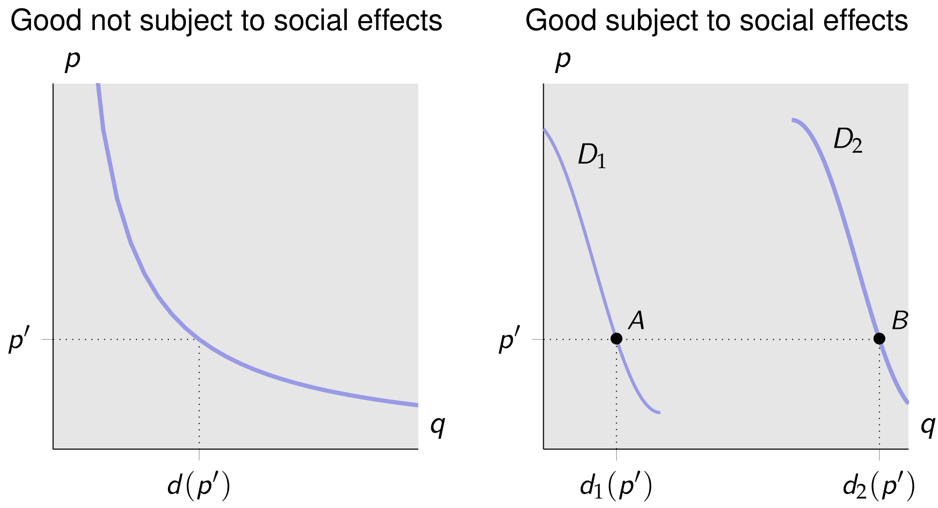


FIGURE 2.3
Multi-valued demand curve

tity demanded. The panel on the right corresponds to the possibility of multi-valued demand curves. If price is equal to p', then the quantity demand can be either $d_1(p')$ or $d_2(p')$, which is greater than $d_1(p')$. We will return to the issues of multi-valued demand curves in Chapter 4.

2.2. DEMAND ELASTICITY

How sensitive is demand to price changes? A possible way to rephrase the question is: how large is the slope of the demand curve? Is the demand curve very steep, in which case quantity demand is not very sensitive to price changes; or is the demand curve very flat, in which case quantity demanded is very sensitive to price changes? (Once again, this may be confusing on the account that we measure price on the vertical axis. See the comment on page 20.)

Even if we know the slope of the demand curve, it's not clear this will be of much use. Consider a first example: world oil demand decreases by 1.3 million barrels a day when price increases from \$50 to \$60 dollars per barrel. Would you consider the demand for oil very sensitive or not very sensitive to price changes? Consider a second example: demand for sugar in Europe decreases by 1 million tones per day when average retail price increases from €.80 to €.90 per

Product and market	Elasticity
Norwegian salmon in Spain	-0.8
Norwegian salmon in Italy	-0.9
Coffee in the Netherlands	-0.2
Natural gas in Europe (short-run)	-0.2
Natural gas in Europe (long-run)	-1.5
US luxury cars in US	-1.9
Foreign luxury cars in US	-2.8
Basic cable TV in US	-4.1
Satellite TV in US	-5.4
Ocean shipping services (worldwide)	-4.4

TABLE 2.1
Demand elasticity examples

kilo. Would you consider the demand for sugar very sensitive or not very sensitive to price changes?

To make matters even more complicated, can you compare the demand for sugar in Europe to the worldwide demand for oil in terms of sensitivity to price changes? The problem, as these examples illustrate, is that, by measuring the slope of the demand curve, we are stuck with units: barrels, dollars, kilos, euros, and so on.

The concept of demand elasticity addresses the limitations of using slope as a measure of demand sensitivity. If in lieu of changes we measure *percent* changes, then units cease to play a role and numbers become easier to interpret. Specifically, the definition of demand elasticity (normally represented by the letter ϵ) is the following:

$$\epsilon \approx \frac{\%\,\Delta\text{ quantity}}{\%\,\Delta\text{ price}} \tag{2.1}$$

where the sign $\approx$ stands for "approximately equal." For example, suppose Hulu gets 21 million subscribers when price is \$8 per month. When price increases to \$9, number of subscribers drops to 14 million. Then

$$\epsilon \approx \frac{\frac{14-21}{21}}{\frac{9-8}{9}} = \frac{-7/21}{1/8} \approx \frac{-33.3\%}{12.5\%} \approx -2.66$$

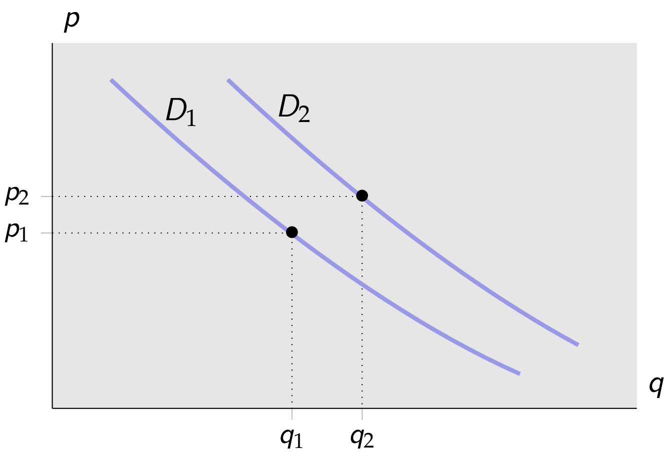


FIGURE 2.4
Pitfalls of demand estimation with mispecified model

Table 2.1 displays the elasticity estimates for several products, including some media and entertainment goods. Notice that the value of demand elasticity is always negative. Why?

A word of caution. When estimating demand elasticity from time series data, one must take into account factors that shift the demand curve. Otherwise, nonsensical estimates of demand elasticity may result. Figure 2.4 illustrates this point. If all the data we have corresponds to the two data points in the figure, we might be tempted to draw a curve along those points and call it the demand curve. This would not make sense: we would get a positively-sloped demand curve, a violation of the law of demand. The problem is that the two data points correspond to two different demand curves. We will talk more about this in Section 2.3 and elsewhere in the book.

ELASTICITY, PRICE, AND REVENUE

An immediate application of the concept of demand elasticity is the relation between price changes and changes in demand and revenues. Since elasticity is approximately given by $(\Delta q/q)$ divided by $(\Delta p/p)$, it follows that

$$\frac{\Delta q}{q} = \frac{\Delta p}{p}\,\epsilon \tag{2.2}$$

The relation between price change and change in revenues is a little more complicated. First notice that revenue is given by $R = p \times q$. It follows that

$$\frac{\Delta R}{R} \approx \frac{\Delta p}{p} + \frac{\Delta q}{q} \tag{2.3}$$

In words, the percent change (or proportional change) in revenue is approximately equal to the percent change in price plus the percent change in quantity sold. For example, if the San Francisco Giants are having a great season and their chances of making it to the World Series get a major boost after a series of good results, then I expect the Giants management to increase ticket prices (they've done it before) *and* for more tickets to be sold. (This is not a violation of the law of demand. Why not?) Specifically, suppose that prices increase by 10% and quantity sold increases by 5%. Then revenues increase by approximately 15% = 10% + 5%. Specifically, suppose initial p was \$100 and initial q was 20 (thousands of tickets). Then initial ticket sales revenue is 100 × 20 = \$2m. Suppose now that p increases to 110 (10% increase) whereas q increases to 21 (5% increase). The new value of ticket sales revenue is 110 × 21 = \$2.31m, an increase of approximately 15% with respect to \$2m.

Combining (2.2) and (2.3), we get

$$\frac{\Delta R}{R} \approx \frac{\Delta p}{p} + \epsilon\,\frac{\Delta p}{p} \;=\; \frac{\Delta p}{p}\,(1+\epsilon)$$

In words, the change in revenue resulting from a change in price is equal to the change in price (percentwise) times $(1+\epsilon)$, where ϵ is demand elasticity.

We can now apply this relation to estimate the impact of a price decrease (for a price increase, change the sign in the statements that follow). Specifically, revenue increases if $\epsilon < -1$ (that is, $|\epsilon| > 1$, where $|\epsilon|$ is the absolute value of ϵ); whereas revenue decreases if $\epsilon > -1$ (that is, $|\epsilon| < 1$). There is also the rather unlikely possibility that $\epsilon =$ is exactly equal to -1, in which case revenue remains unchanged following a price decrease or increase.

Let us consider a specific example. Suppose that the price of coffee decreases by 1%. As per Table 2.1, suppose the elasticity is equal to $-.2$. Then the demand for coffee increases by $.2\% = -1\% \times (-.2)$; whereas revenue falls by $0.8\% = -1\% \times (1 + (-.2))$. Regarding US luxury cars, a 1% price decrease leads to a demand increase of 1.9% and a revenue increase of 0.9%. (Check.)

As the above analysis shows, the threshold $\epsilon = -1$ (or $|\epsilon| = 1$) is very relevant, namely for the purpose of the relation between price and revenue. We say that, if the demand elasticity is greater than 1 (in absolute value), then the demand is **elastic**, whereas, if the demand elasticity is lower than 1 (in absolute value), then the demand is **inelastic**. We thus conclude that, ***if demand is elastic, then a price decrease leads to a higher revenue, whereas, if demand is inelastic, then a price decrease leads to a lower revenue.*** The opposite is true for a price increase.

MORE ELASTICITIES

So far we've considered how sensitive the demand for product i is to changes in the price of product i. What about changes in the price of another product, say, product j? The concept of **cross-price elasticity** provides an answer to this question. It is defined as follows:

$$\epsilon_{ij} \approx \frac{\%\,\Delta\,\text{quantity}_i}{\%\,\Delta\,\text{price}_j}$$

Based on the value of ϵ_{ij} we can define the economic nature of a pair of goods. In terms of economics jargon,

- If $\epsilon_{ij} > 0$, then we say i and j are *substitutes*
- If $\epsilon_{ij} < 0$, then we say i and j are *complements*
- If $\epsilon_{ij} = 0$, then we say i and j are *independent*

For example, NY Mets tickets and hotdogs in the park are complements, whereas theater tickets on Wednesday and theater tickets on Thursday are substitute products. Can you think of other examples? It's fun to think about pairs of independent products; for example *Ben and Jerry*'s ice cream and *Michelin* tires.

It is not always obvious whether two goods are complements or substitutes. Consider the following quote regarding the history of the music industry:

> Convinced that radio broadcasts were crowding out music sales, record companies in the 1920s waged a series

> of court battles demanding high royalties for songs, leading some networks to stop playing major-label music altogether. It soon became apparent, however, that radio airplay dramatically increased record sales, and by the 1950s record companies were paying large bribes to get their songs onto disk jockeys' playlists.

In terms of the above jargon, early on industry experts thought that radio play and record sales were substitutes when in fact they are complements.

Our elasticity gallery tour ends with the concept of **income elasticity**. The idea is to measure how sensitive demand is to changes in consumer income. Its definition is given by

$$\epsilon_y \approx \frac{\% \, \Delta \text{ quantity}}{\% \, \Delta \text{ income}}$$

In terms of economics jargon, we say that a good is a

- **inferior good** if $\epsilon_y < 0$
- **normal good** if $\epsilon_y > 0$
- **necessity** if $0 < \epsilon_y < 1$
- **luxury** if $\epsilon_y > 1$ (fraction of income increases as income increases)

Table 2.2 shows aggregate US consumer expenditure data. At this aggregate level, would you say that entertainment is a normal good: as household income increases, entertainment expenditures increase. In fact, the *fraction* of income spent on entertainment increases, from 5.18 to 5.83. In other words, as income increases the percent increase in entertainment expenditures is greater than the percent increase in income. This implies that the income elasticity of entertainment expenditures is greater than 1, i.e., in aggregate terms entertainment is a luxury (in the economics sense of the word.

Two notes are in order. First, 5.18 is very close to 5.83, which suggests the fraction of one's income spent on entertainment is approximately constant. Second, Table 2.2 divides expenditures at a very aggregate level. If we were to separate entertainment expenditures by type of entertainment I suspect we would find that the income elasticity of movie theater tickets, for example, is lower than that of NFL games.

	Household income	
	<$50k	>$50k
Percentage of households	52.5	47.4
Income before taxes ($)	24,643	106,754
Percentages of total expenditures		
Food	14.54	11.98
Housing	37.25	32.41
Transportation	17.32	16.95
Healthcare	7.40	5.22
Entertainment	5.18	5.83
Personal insurance and pensions	5.50	13.66
Other	17.99	19.78

TABLE 2.2
US consumer expenditures (Source: US Bureau of Labor Statistics, 2008)

2.3. CASE STUDY: WEDNESDAYS AT CINEMEX

Cinemex, the Mexico-based movie theater chain, is a success story in movie exhibition. It also provides an excellent example of how to estimate demand and demand elasticity — and how not to. Before getting into data issues, some information about the case in question.

MARKET ENTRY

Cinemex started with a student business plan. Matthew Heyman and two of his business school classmates, Adolfo Fastlich and Miguel Angel Davila, speculated that Mexico was ready for world-class movie theaters. Decades of regulation, including fixed (low) ticket prices, had produced an installed base of old and dilapidated theaters. When the regulations were lifted, Heyman and his colleagues decided that Mexico City offered an attractive market for a high-end chain of theaters. Rejecting job offers from Blockbuster, Goldman Sachs, McKinsey, Pepsico, and others, they took their plan on the road in search of investors.

In 1994, they secured $21.5m in equity financing from JPMorgan Partners and a partnership of the Bluhm family of Chicago. The deal

Wikimedia Commons

A Cinemex movie theater complex

is generally acknowledged to be the largest venture capital start-up in Mexican history. In December the economy collapsed, with real output falling by 15% and the value of the peso falling in half between early December 1994 and mid-1995. Although this made Mexico a less-attractive market in the short run, it also made land cheaper and scared off potential competitors (AMC and Lowes) who had seen the same opportunity. Cinemex opened its first complex, Cinemex Altavista, in August 1995.

From the start, Cinemex followed a strategy of differentiation through branding. Since all theaters have access to the same films, and in some cases the same or similar locations, Heyman felt that the greatest leverage was in the quality of the theater itself. The low quality of existing theaters presented an opportunity to develop a brand associated with quality, including bigger and better screens, complete carpeting in all rooms, well-illuminated interiors, emergency lights on the floors, modern light cards for promotional placards, and attractive marquees.

These amenities, considered standard for decades in most American theaters, were seen as almost revolutionary when first introduced into Mexico. The candy shops were the same in all the complexes, with large displays and well-maintained cash registers that allowed for quick service. Management trained its employees to be courteous and helpful. It was also the first movie chain in the world to introduce its own system for customers to purchase and reserve tickets by telephone and the Internet, and was the only chain in the world with 100% digital sound.

This commitment to quality was rewarded by the market. By

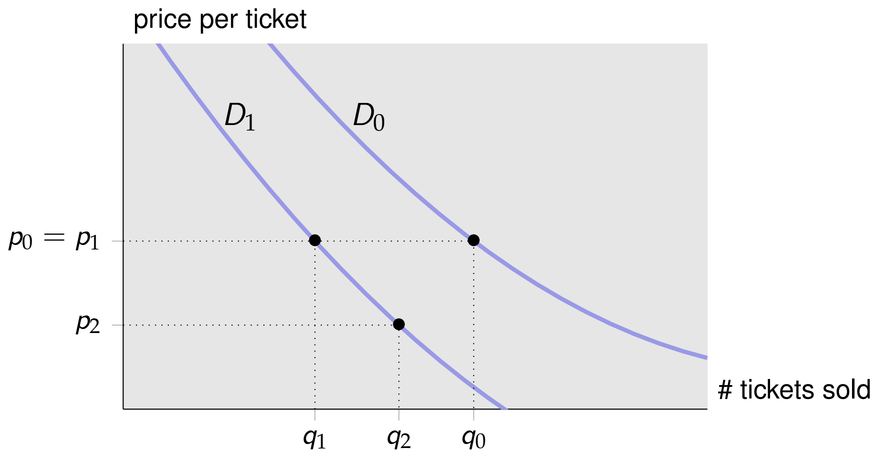


FIGURE 2.5
Shifts in and along Cinemex's demand curve

2000, Cinemex had captured 52% of Mexico City's movie market and 23% of Mexico's national market. The National Organization of Theater Owners named Cinemex "International Exhibitor of the Year" in 2001. By mid 2002, the company had 349 screens in 31 locations and had generated a reported compound annual rate of return well in excess of 20% for its initial investors.

2-FOR-1 WEDNESDAYS

In the spring of 2001, Cinemex's competitors began offering a special deal: any customer who purchased a ticket to see a film on a Wednesday (traditionally a slow day at the box office) would receive a second ticket at no additional charge. This ploy cut into Cinemex's attendance figures. On five of the first six Wednesdays after the deal's introduction, Cinemex's attendance was less than in the same week during the previous year.

Heyman faced a difficult decision. Should he offer his own two-for-one deal on Wednesdays? This might raise attendance, but since many tickets would be given away for free, it might also reduce ticket revenues. Or should he do nothing, hoping that the appeal of Cinemex's customer service package would eventually bring customers back?

Eventually, Heyman did match the rival's 2-for-1 deal. In terms of the demand curve, the 2001 events are illustrated in Figure 2.5. Initially, Cinemex's price was at p_0 and the number of tickets sold at q_0. Then the rival lowered their price. This implied a shift in Cinemex's demand curve, from D_0 to D_1. So, even though Cinemex did not change its ticket price, its ticket sales dropped from q_0 to q_1. When Cinemex drops its own price from p_1 to p_2, we observe a movement along Cinemex's demand curve, so that the number of ticket sales increases from q_1 to q_2.

The value in Figure 2.5 only show the qualitative effects of the various events. In particular, we do not know how big the changes were — not even whether q_2 is greater or smaller than q_0. In order to get a better idea of whether Heyman did the right thing or not, we need to know how Cinemex's demand is sensitive to changes in the rival's price as well as its own price (one of the central themes in this chapter). And in order to estimate demand and demand elasticity, we must look at the data Heyman collected over the years.

DATA ANALYSIS

Yogi Berra, the famous baseball player, once said that "you can see a lot of things just by looking." This is a good rule for data analysis: start by plotting the data and see what you can learn from it.

Figure 2.6 plots the time series of Cinemex's Wednesdays data. The left panel show the number of tickets sold, per week, for the years 2000 and 2002. We exclude 2001 on account of the various price changes that occurred during that year. Several things stand out from the figure. First, the 2002 values are uniformly greater than the 2000 values. This reflects the secular growth in Cinemex's size. In particular, during this period Cinemex opened a number of new theaters in Mexico City.

A second interesting feature from Figure 2.6 is that there are some noticeable seasonable patters. In particular, we notice an increase during summer months (in both years) and during the Christmas season. There are other peaks, but these are not repeated in both years, and so we should not call a seasonal pattern.

Last but not least, the left panel of Figure 2.6 suggests that there is a lot of noise in the data, that is, significant variations from week to week. This is likely due to the quality of the movies being shown, as

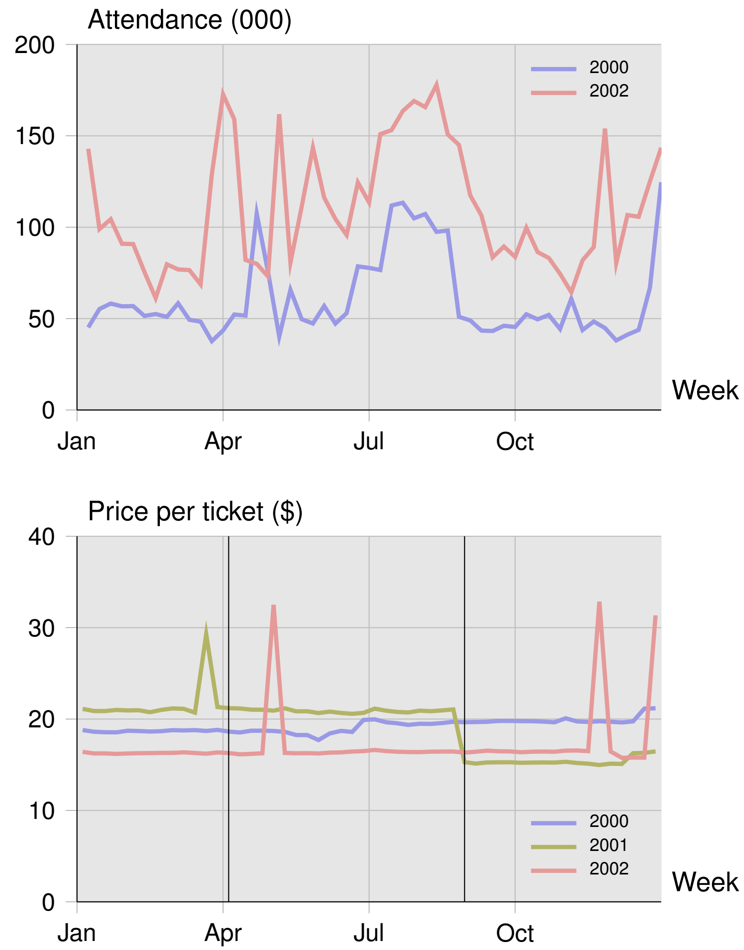


FIGURE 2.6
Cinemex Wednesday attendance and average price per ticket

well as other factors such as national holidays, etc. Noise in an issue that data analysts have to deal with. We'll come back to it later.

The right panel plots the time series of average price, which we compute by dividing sales in $ by number of tickets sold. As mentioned before, prices changed during 2001, not during 2000 or 2002. However, we find some spikes in average price during 2002. I have not had a chance to confirm this with Heyman, but I suspect this is simply due to observation error. This is one learning point from data analysis: in addition to all unobserved factors that lead to changes in demand (leading to the volatility observed in the left panel), there is also noise due to measurement error (for example, someone pressed the wrong key when entering the data). There's not a lot we can do

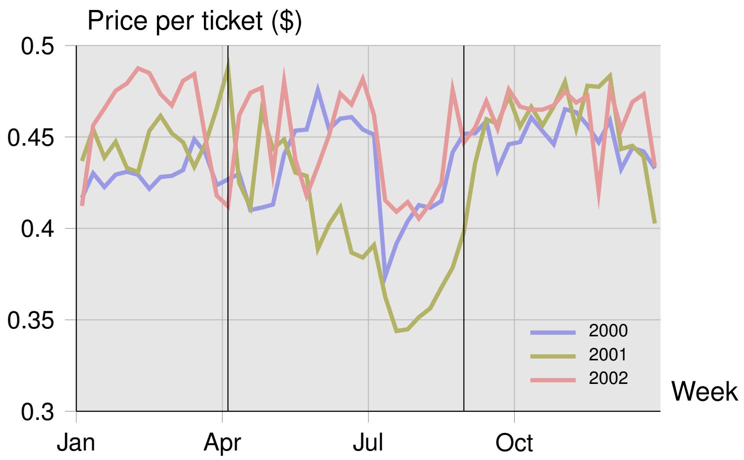


FIGURE 2.7
Cinemex Wednesday market share

about it (besides being careful when entering data).

Consider now the data for 2001 (solid line). Again, we find a spike in March which is likely due to measurement error. We also notice a drop in September. This is expected, for we know that Cinemex eventually decided to follow the rival's 2-1 deal on Wedensday. So far, so good, but there is an issue: 2-for-1 suggests that the average price per ticket sold should be something like one half of what it was before. However, we observe a much smaller drop in price. There are two possible explanations for this. First, Cinemex increased ticket price at the same price as it offered the 2-for-1 deal. Second, there is a large number of solo moviegoers (or an odd number of them), so that they do not take advantage of the 2-for-1 deal.

Our next step is to estimate the impact that the rival's price change, as well as Cinemex's price change, had on Cinemex sales. This is tricky. Suppose that I compare ticket sales the week before the rival offered a 2-for-1 with ticket sales the week after the rival offered a 2-for-1. The problem is that a number of other things may have changed in addition to the rival's price change. For example, it could be that the week the rival lowered price the movie shown (by Cinemex and by the rival) was considerably less appealing than the week before. We would then observe a drop in ticket sales but such drop would result from a shock in demand independent from the rival's price change. It's the usual correlation vs causality problem.

Time period	Wednesday deals	Cinemex's attendance
A: weeks 1–11, 13; year 2000	No deals	626,724
B: weeks 14–34; year 2000	No deals	1,564,004
C: weeks 35–52; year 2000	No deals	937,960
D: weeks 1–11, 13; year 2001	No deals	758,842
E: weeks 14–34; year 2001	Rival offers 2-for-1 deal	1,589,713
F: weeks 35–52; year 2001	Both offer 2-for-1 deal	1,306,623

TABLE 2.3
Counterfactual analysis

One way to address this is to compute Cinemex's market share, the idea being that confounding factors such as movie appeal affect both Cinemex's and the rival's number of tickets sold proportionately, so that market shares are invariant to such shocks. We have no proof that this is the case, but it seems reasonable to assume that.

With that in mind, in Figure 2.7 we plot Cinemex's market share (in Mexico City) for the years 2000, 2001 and 2002. The data suggests that Cinemex's market share is around 45%. We notice a drop in market share during the summer (July and August) in all three years. However, the drop in market share is particularly pronounced in 2001. This suggests that the rival's price drop did cut into Cinemex's market share, whereas Cinemex's reply re-established its initial position.

Market share analysis is helpful as it helps control for a lot of extraneous factors, but it does not help estimating the impact of price changes on the number of tickets sold. The reason is that the rival's price decrease may have cut into Cinemex's sales or simply increased total sales. In principle, its possible that the rival's price decrease did not decrease Cinemex's sales at all!

COUNTERFACTUAL ANALYSIS

When faced with inference problems of the sort presented here, economists — and historians and many others — are wont to follow a process of **counterfactual analysis**. What impact did the rival's price change have on Cinemex's ticket sales? Suppose we first answer the question, what would Cinemex's ticket sales have been had the rival not changed its price. If we can answer the second question, then answer the first one is easy: the impact of the rival's price change is simply the difference between the counterfactual and the actual number.

To illustrate this, consider the values in Table 2.3. We divide year 2001 into three periods:

- the first 13 weeks, a period during which neither Cinemex nor the rival offered 2-for-1 on Wednesdays
- weeks 14 through 24, the period when the rival but not Cinemex offered 2-for-1
- weeks 35 through 52, the period when both Cinemex and the rival offered 2-for-1.

Furthermore, we exclude week 12 since, as per the discussion in the previous subsection, we suspect there is some anomaly in the value of price.

What would Cinemex have sold in period E (weeks 14 to 34 in 2001) had the rival not offered the 2-for-1 deal? One natural answer is to look at the 2000 sales pattern and estimate what we would have gotten had 2001 been just like 2000 in terms of annual sales patterns. Specifically, period B was greater than period A by a factor of

$$B/A = 1,564,004/626,724 = 2.496$$

Had the same pattern repeated in 2001, we would have expected E to be given by

$$E' = D \times 2.496 = 758,842 \times 2.496 = 1,893,708$$

(Note: In this and the following computations, I display the rounded values of the ratios but carry the entire value when computing the subsequent values.) In other words, the E' is our counterfactual of E under the assumption of no rival's 2-for-1 deal and the repetition in 2001 of the pattern observed in 2000. Notice that E' is greater than E. This makes sense: we estimate that, had the rival not introduced 2-for-1, then Cinemex's sales would have been greater. Specifically, we compute the impact of the rival's 2-for-1 deal as given by

$$E - E' = 1,589,713 - 1,893,708 = -303,995$$

In sum, we estimate that the rival's introduction of 2-for-1 cost Cinemex 303,995 fewer tickets sold during weeks 14 to 34 in 2001.

This is not the only way to get to this number. We could also have measured year-on-year changes. Specifically, focusing on the first 13 weeks of the year, for which no 2-for-1 deal was offered in either 2000 or 2001, we observe that

$$D/A = 758,842/626,724 = 1.21$$

that is, a 21% increase. Had the rival not introduced 2-for-1, we would have expected a value of E given by

$$E' = B \times 1.21 = 1,564,004 \times 1.21 = 1,893,708$$

Bingo! As expected, we get the same number as before.

We can now use the same methodology to estimate the impact of Cinemex's own price decrease. In 2000, the ratio between the third and second period is given by

$$C/B = 937,960/1,564,004 = 0.60$$

We would therefore expect that, if the rival continued its 2-for-1 deal and Cinemex had *not* introduced its own, then ticket sales during period F would have been

$$F' = E \times 0.60 = 953,378$$

We thus estimate that, by offering 2-for-1, Cinemex increased ticket sales by

$$F - F' = 1,306,623 - 953,378 = 353,245$$

Finally, we can use our estimates to estimate the price elasticity of Cinemex's demand. Had Cinemex not changed its price, p would have been $p = 20.9$ (the observed average price before the 2-for-1 deal. The data shows that, after then 2-for-1 deal was introduced, average price dropped to 15.4. Our counterfactual analysis suggests that, had the price not changed, ticket sales would have been $F' = 953,378$. Ticket sales were actually $F = 1,306,623$. In other words, a $20.9 - 15.4 = 5.5$ price decrease is associated with a $353,245$ increase in ticket sales. In terms of elasticity, we get the estimate

$$\epsilon = \frac{353,245/953,378}{-5.5/20.9} \approx -1.41$$

DID HEYMAN DO THE RIGHT THING?

Was Heyman right to match the rival's 2-for-1 pricing strategy? The demand elasticity estimates (around -1.4) tell us that the demand increase from dropping the price is greater (percent wise) than the price drop, which in turn implies a revenue increase. Assuming that Cinemex's costs do not vary with respect to the number of movie goers — which seems a reasonable approximation — what is said about revenues is also true for profits: Cinemex's price increase let to an increase in profits. So, in terms of Wednesday ticket sales (and profits) it seems Heyman changed price in the right direction.

There are at least two important qualifications to the above statement. First, it may be true that Cinemex's costs do not depend on the number of movie goers. However, Cinemex's total revenues are likely to increase when it attracts more movie goers: remember that the theatrical exhibition business is more than just selling admission tickets, it's also about selling popcorn. In this sense, the strategy of lowering price looks even better than before: not only did Cinemex increase tickets sales, it also increased popcorn sales.

A second qualification, which goes against the price decrease strategy, is that we are only looking at Wednesday sales. One tantalizing possibility is that a lower Wednesday price increases sales *at the expense of sales during other days of the week*. This is the old cannibalization problem faced by multi-product firms. We will return to it in Chapter 4.

CINEMEX: BY WAY OF CONCLUSION

This has been a long mini-case study, but I hope with interesting learning points. Here's a summary of the main ones. First, as the case illustrates data analysis is not something mechanical: it requires some judgment, and people can differ about what judgments make the most sense. (You may have had different ideas about tackling the issues of demand estimation, and I would be delighted to hear abou them.)

In the present case, key issues of judgment regarding data include secular trends related to population, income, number of movie theaters, etc; as well as seasonal variation, that is, the fact that some times of year are better for movie attendance. Failure to take into account secular trends or seasonal patterns may lead to biased estimates.

Another important point about the data is that there is a lot of noise, both due to unobserved variables (how good is are the movies being shown) and possibly measurement error (someone may have pressed the wrong key when entering the data). In this case, the trick is to make use of the **law of large numbers**: as we aggregate more and more observations, these noisy shocks partly cancel out and become less significant as a fraction of the total value. Specifically, our strategy in the present case was to aggregate several weeks in each of the relevant periods, hoping that the average quality of a movie in each period is similar from year to year.

POSTSCRIPT

In July 2002, Heyman announced the sale of Cinemex to Canadian buyout firm Onex and Los Angeles-based Oaktree Capital Management. He planned to focus on Digital Projection Partners, a startup he formed in 2001 that was in discussions with MPAA companies (movie studios) regarding the funding and implementation of the studios' digital projection initiative.

APPENDIX: FUNCTIONS

In economics (and other fields, too), we often use relations between two variables: demand depends on price, cost depends on quantity produced, and so on. We call these relations "functions." More formally, a function f assigns a (single) value y to each possible value of x. We write it this way: $y = f(x)$. In a spreadsheet program, you might imagine setting up a table with a grid of values for x. The function would then be a formula that computes y for each value of x.

Perhaps the easiest way to think about a function is to draw it: put x on the horizontal axis and plot the values of y associated with each x on the vertical axis. Some examples are given in Figure 2.8. We will generally be interested in functions that are "continuous" (they don't have "jumps," as in (b)) and "smooth" (they don't have "kinks," as in (c)).

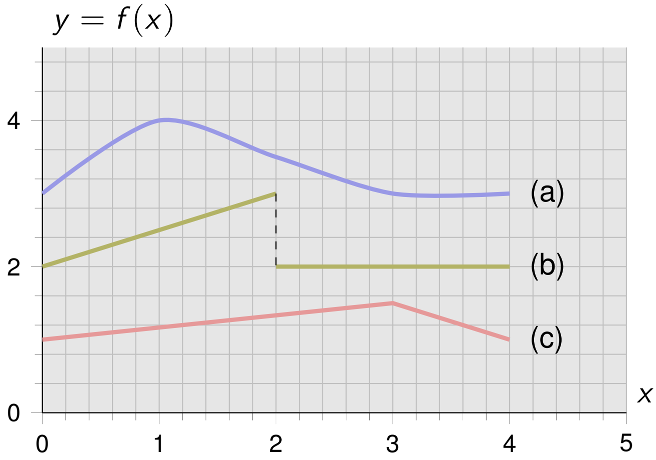


FIGURE 2.8
Examples of functions.

KEY CONCEPTS

demand function | demand curve | inverse demand curve | awareness | demand spillovers | halo effect | elastic | inelastic | cross-price elasticity | income elasticity | inferior good | normal good | necessity | luxury | counterfactual analysis | law of large numbers

REVIEW AND PRACTICE PROBLEMS

■ **2.1. Law of demand.** What is the law of demand? Can there be exceptions to the law?

■ **2.2. Entertainment demand.** Identify one aspect of entertainment demand that differs from the demand for "common" goods.

■ **2.3. Demand curve.** What is the difference between a demand curve and a demand function?

■ **2.4. Elasticity.** What is the price elasticity of demand? What is it good for? In addition to the price elasticity, what other demand elasticities are there?

■ **2.5. Elasticity.** What is the sign of the price elasticity of demand? Why?

■ **2.6. Complements and substitutes.** Consider the following pairs of goods. Would you consider them substitutes or complements? Why?

(a) Computers and paper.

(b) Music file sharing and recorded music.

(c) Music file sharing and concerts.

(d) Online and offline retailing.

■ **2.7. Demand.** Consider the good "tickets to movie X at a Cinemex theater on date t" and the following prices: (a) price of a ticket to movie X at a Cinemex theater on date t, and (b) price of a ticket to movie X at a Cinemex's rival theater on date t. Based on these three variables, indicate the difference between a movement along the demand curve and a shift of the demand curve.

■ **2.8. Demand elasticity.** Compared to 2017, the price of the music service CabralTunes increased by 10%, while the number of subscribers decreased by 20%. What is your estimate of the price elasticity of demand for CabralTunes? What assumptions do you need to make in order to come up with such estimate? What additional information would you need to obtain a more accurate estimate?

■ **2.9. Elasticity** At the current price, the demand elasticity at a given movie theater is given by -0.8. Would revenues increase or decrease if the theater were to increase price? What would you advise the theater owner to do? Suppose that concessions represent an important portion of total revenues and that concessions revenues are proportional to the number of tickets sold. How would you change your advice to the theater owner?

■ **2.10. Social effects.** The demand for many entertainment goods is frequently subject to significant social effects. Explain the nature of these social effects and what implications they have for pricing and other dimensions of a seller's marketing strategy.

CHAPTER 3
SUPPLY

It takes two to tango. In order for economic transactions to take place, we need both demand and supply (or, as someone aptly put it, "it's not supply *or* demand, it's supply *and* demand"). In the previous chapter, we looked at the main determinants of demand, in general and in particular for entertainment goods. In this chapter we take the perspective of supply. Specifically, we examine the main determinants of the supply of goods in general and of entertainment goods in particular.

3.1. PRODUCTION FUNCTION

At the risk of oversimplifying, we can think of a firm as a process of transforming inputs into outputs. This is easier to see for a firm that makes actual things. For example, a bagel bakery uses water, flour and other ingredients, together with machinery (an oven) and labor (someone has to put it all together), to produce tasty bagels. Firms that offer services also go through a similar process. For example, a consulting firm uses hours of labor — many, many hours, I'm told — together with some capital (mainly laptop computers) and materials (paper and paper clips), to offer solid advice to corporations that need it.

The firm's **production function** is the mapping that tells us, for a given set of inputs, how much output a firm is able to produce. Normally, this depends on the particular firm, as some firms are more efficient than others at transforming inputs into outputs. It also depends on the quality of inputs, for example skilled versus unskilled labor. A standard production function looks like the following:

$$Y = f(K, L, M)\,\xi + \epsilon \tag{3.1}$$

where Y stands for output level (e.g., tons of steel), f represents the production function itself, K is the level of capital (e.g., the number and size of furnaces), L the quantity of labor employed (e.g., the number of workers), and M the quantity of materials used in the production process (e.g., coal, iron, etc).

Two additional factors deserve particular attention. First, ξ is a productivity factor. The idea is that, for given input levels (number of workers, quantity of capital, etc), some firms are able to produce more than others. We then model this by assuming such firms have a higher ξ parameter (frequently referred to as **total factor productivity**). Second, to the value of the function f we add the value of ϵ, which stands for uncertainty in the production process. The idea is that we can only account for a certain number of variables, but in the real world there are multiple other factors which affect the actual output level. (Can you think of any?)

What about entertainment goods — does the above framework also apply? From our discussion in Chapter **??**, we know that entertainment goods share some important characteristics not always found in other goods. First, the output is unique. You cannot count the number of *Bohemian Rhapsody* songs recorded by the band Queen: there is only one. You might talk about the number of times it was sold or downloaded or listened to, but that's an indicator of demand, not supply.

Second, and related to the first one, the output of an entertainment production function is typically a non-rival good. To better understand this, it's probably better to begin with an example of a rival good, for example, an apple (the fruit, not the computer). If I eat an apple, you cannot eat that same apple. Similarly, a car, a house, etc, are examples of **rival goods**: consumption by consumer a precludes consumption by consumer b. By contrast, a **non-rival good** is one such that consumption by a does not preclude consumption of the same product by b. The fact you listen to *Bohemian Rhapsody*, for example, does not prevent me from doing so.

Third, unlike commodity production functions — to give an extreme example — it can be very difficult to measure inputs in an entertainment production function. For example, it takes five players to put together a basketball team. However, measuring the number of workers (as I suggested earlier) is clearly not a good way of measuring the inputs into the basketball entertainment production function: It makes a big difference who those players are, and quantifying the quality of players can be difficult. (In the chapter on sports, we will talk about *moneyball*, the process of valuing sports players. Pioneered by Billy Beane in baseball, moneyball can be thought of as an attempt to measure the sports production function with greater precision.)

With all of these caveats in mind, we might say that an entertainment production function looks like the following:

$$Y = f(X, S) + \epsilon \tag{3.2}$$

where Y stands for output level (e.g., the number of potentially interested viewers), f represents the production function itself, X corresponds to observable (and measurable) inputs (e.g., number of hours spent writing an opera score), S is some measure of the value of unique inputs (e.g., star power), and ϵ, as before, measures uncertainty in the production process. In what follows, we focus on these three important determinants of entertainment supply: X, S and ϵ.

MEASURABLE INPUTS

In more traditional industries, there is a relatively regular relation between quantity and quality of inputs and quantity and quality of outputs. There is also the issue of total factor productivity, measured in (3.1) by the factor ξ. One of the puzzling issues in microeconomics is that two different firms produce very different output levels despite using the same amounts of inputs. Something similar happens in entertainment industries: you get you pay for.

Take soccer, for example. Figure 3.1 shows the relation between team budget and team performance in the English Premier League (EPL) during the 2018–2019 season. Irrespective of whether we consider payroll or the sum of payroll and transfer fees, we see that teams that spend more on getting strong players perform better: you get what you pay for.

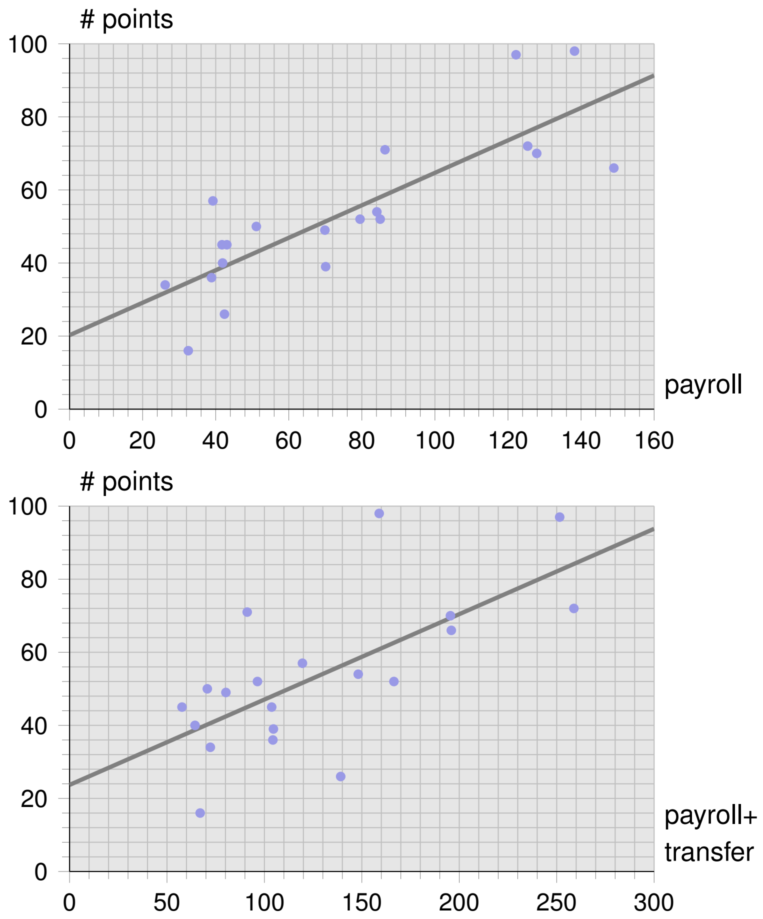


FIGURE 3.1
Soccer production function: English Premier League 2018–2019

In movies, too, there is evidence that hiring an expensive star increases revenues (at the box office) but does not increase profits. One interpretation of this result is, again, that you get what you pay for: the compensation paid to superstars is commensurate to their contribution to the movie's performance at the box office.

Continuing with movies, Figure 3.2 plots the values of budget and box-office revenue for about 5,500 movies for which the data is available. The values are deflated and expressed in 2000 dollars. In this way we properly measure the impact of movies such as *Gone With The Wind*, which cost less than 4 million dollars in 1939, when a dollar was worth a lot more than it is today.

Since the values of budget and revenue vary enormously across

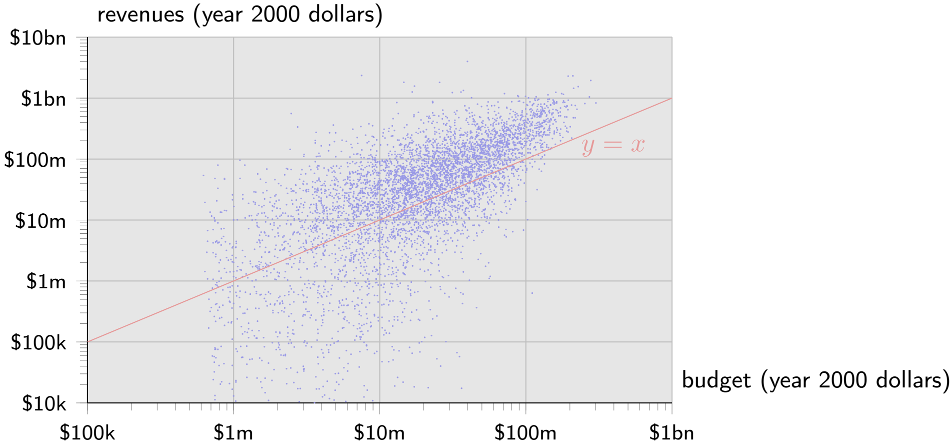


FIGURE 3.2
Movie budget and movie revenue (source). Each dot represents a movie.

movies, the values are plotted on a logarithmic scale. Unlike a linear scale (the "common" scale), where a fixed increment along the axis corresponds to a fixed increase in value, in a logarithmic scale a fixed increment along the axis corresponds to multiplying the variable by a factor of 10. We thus get $1k, $10k, $100k, $1m, $10m, and so one, tick after tick.

The values in Figure 3.2 suggest two things. First, there is clearly a positive correlation between budget and performance at the box-office. Like soccer, this is an instance of the rule: you get what you pay for. However, there is also considerable variation. In particular, if we look at movies with a budget of less than $1 million — mostly indies — then the data look more like a cloud than like a clear positive correlation. We will later return to the issue of uncertainty in the entertainment production function.

UNIQUE INPUTS

One of the more idiosyncratic features of the entertainment industries is the uniqueness of some of the inputs used in the entertainment production function. As much as the New York Knicks want, they cannot use LeBron James as an input unless they hire LeBron James himself, and there is only one LeBron James.

This uniqueness of talent, in turn, leads to star power: superstars hold considerable star power because they know that, absent their

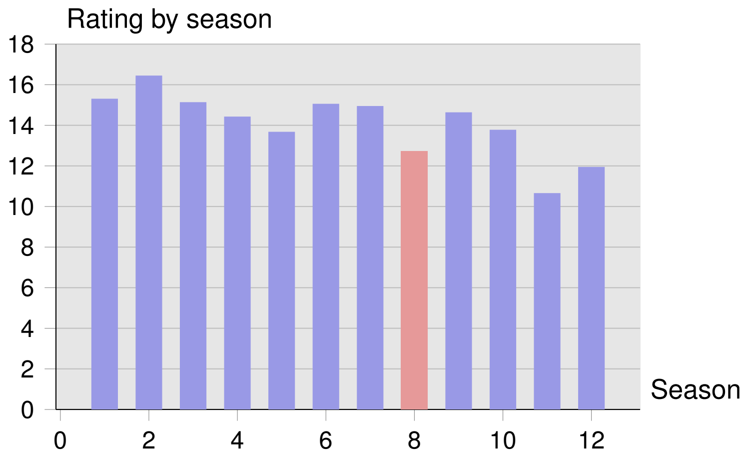


FIGURE 3.3
Two and a Half Men: the Charlie Sheen effect (or lack thereof)

contribution, the value of a given production function is considerably lower. Referring to Tom Hanks' role in the movie *Cast Away*, Bill Mechanic (former chairman of Twentieth Century Fox) commented that

> A guy stranded on an island without Tom Hanks is not a movie. With another actor, [the movie *Cast Away*] would gross $40 million. With Tom Hanks it grossed $200 million. There's no way to replace that kind of power.

That said, the case can be made that the uniqueness of inputs, in particular star power, can at times be over-estimated. Consider the TV show *Two and And a Half Men*. In 2011, during the show's eighth season, Charlie Sheen — the show's leading star — was effectively fired (it's a long story). Soon after, CBS announced that Ashton Kutcher would join the cast. "I can't replace Charlie Sheen but I'm going to work my ass off to entertain the hell out of people!," he promised. Could the show survive without Sheen?

Figure 3.3 suggests that, notwithstanding Sheen's star power and how unique his contribution may have been, the show did reasonably well without him. We will return to this in Chapter 7, when discussing power games.

In addition to TV shows, the uniqueness of inputs is particularly salient in movie sequels or prequels. Take the case of *Hunger Games*.

Jonathan E.S.A. on flicker.com (cropped), and Wikicommons

Can Lionsgate successfully add to the *Hunger Games* franchise without Jennifer Lawrence?

Based on a novel trilogy written by American novelist Suzanne Collins, Lionsgate produced a series of four blockbusters from 2012–2015. The series' main character, Katniss Everdeen, was played by Jennifer Lawrence. Lawrence was already a known actress by 2012, having received her first Best Actress nomination for 2010's *Winter's Bone*. However, it was *Hunger Games* that cemented her star value. By 2016 she was the world's highest paid actress.

Given the trilogy's success, it was no surprise that Collins and Lionsgate worked on adding a new title to the franchise. The problem is that, as early as 2015, Jennifer Lawrence stated she would not be interested in working on the project. "I wouldn't be involved. I think it's too soon." Lawrence is clearly an essential "input" for the Everdeen character, so no Lawrence means no movie featuring Everdeen. After years of wait, a new volume, *The Ballad of Songbirds and Snakes*, was released on May 2020. Instead of Everdeen, the plot follows a young Coriolanus Snow and the events that eventually lead him on the path to becoming the tyrannical leader of Panem. The corresponding movie is scheduled to be released in the US in November 2023. It does not feature Lawrence.

UNCERTAINTY: NOBODY KNOWS ANYTHING

As much as we measure the quality and quantity of inputs, the output of an entertainment production function is typically subject to significant uncertainty. Referring to the movie industry, screenwriter William Goldman famously stated that "nobody knows anything." Former Disney CEO Michael Eisner expanded on it by stating that "the movie business has always been like the wild-catting oil business. Everyone wants a gusher."

There is actually something one can say about uncertainty beyond "nobody knows anything." Suppose that you toss a coin 100 times and count the number of heads. Most of the time you get about 50% heads. Sometimes — less often —- you get 70% heads (or 30% heads). Sometimes — even less often — you get 90% heads (or 10% heads). It turns out that the **frequency** with which you get each value between 0 and 100 is measured by a bell curve like the one on the left panel of Figure 3.4 (the dashed line).

One of the most important theorems of probability theory, the Central Limit Theorem, states that most measurable events are distributed according to a frequency pattern that looks like the bell curve on the top panel of Figure 3.4 (the red line): middle values are most likely, whereas values away from the average are less and less likely to happen (e.g., 95 heads out of 100 coin tosses). This pattern of frequency distribution is so common that this bell distribution is known as the **normal distribution**. (The normal distribution is also known as Gaussian distribution. For aficionados: if you measure the quantity in question in logarithms, as is the case on the left panel of Figure 3.4, then we say the distribution is log-normal.)

It turns out that the market success of entertainment ventures tends to depart from this normal pattern (in the technical sense of "normal"). Specifically, very "large" outcomes are disproportionally likely to happen. In terms of the movie industry, this means that mega-blockbusters occur more often than would be predicted by the normal distribution. (Movies, by the way, are not the only exception to the rule that events tend to follow a "normal" pattern, that is, being distributed according to a normal distribution. Another exceptions is given by the market value of medical drugs. More generally, the interested reader is referred the book *The Black Swan*.)

In addition to the red normal-distribution curve, the top panel of Figure 3.4 plots the estimated frequency of box office success (blue line). As can be seen, the distribution is bi-modal, that is, there is a high probability that a movie "bombs" (revenues around 10 thousand dollars, the left "hump" of the distribution); but there is also a high probability that the movie does very well (revenues close to 100 million dollars, the right "hump" of the distribution). Part of this variation in movie revenues is due to different budget levels. How-

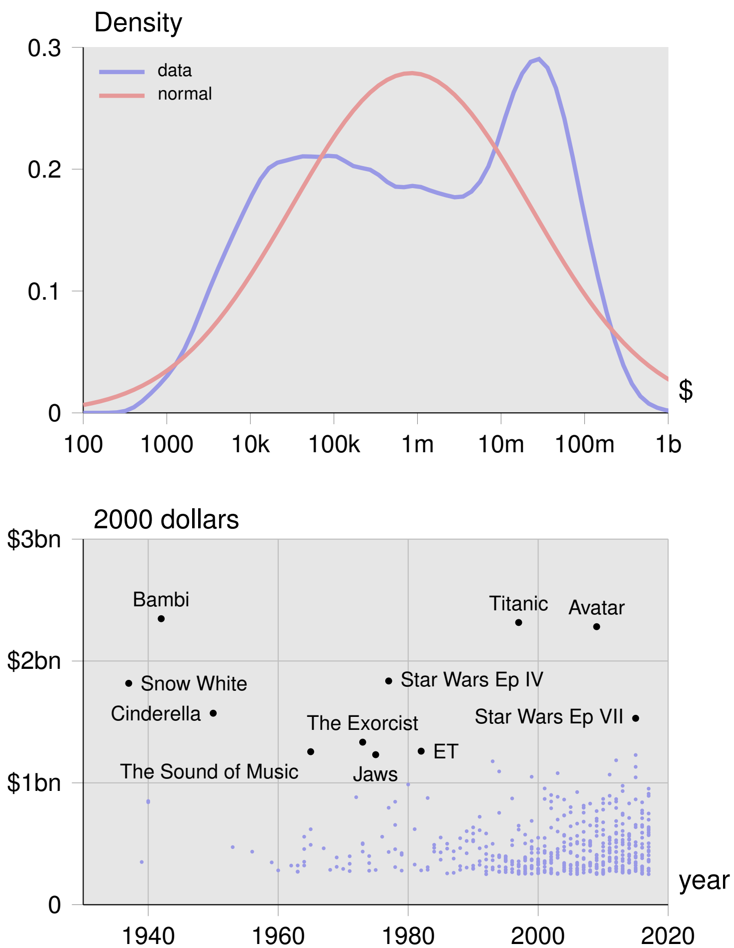


FIGURE 3.4
Distribution of US box-office revenue (top panel) and blockbuster revenues (bottom panel) (source: boxofficemojo.com, number.com, and author's calculations)

ever, it can be shown it also results from the very high variance of the ϵ factor in equation (3.2).

To continue with movies, the right panel of Figure 3.4 plots the top 500 blockbusters of all time for which we have data. As before (cf Figure 3.2) the values are expressed in 2000 dollars. As can be seen, most blockbusters gross less that one billion dollars, which is not surprising, considering that one billion dollars is a lot of money. What is perhaps surprising is the frequency with which we observe blockbusters that hit it "out of the park," for example James Cameron's *Titanic* or *Avatar*. (Trivia question: Figure 3.4 actually only plots 499

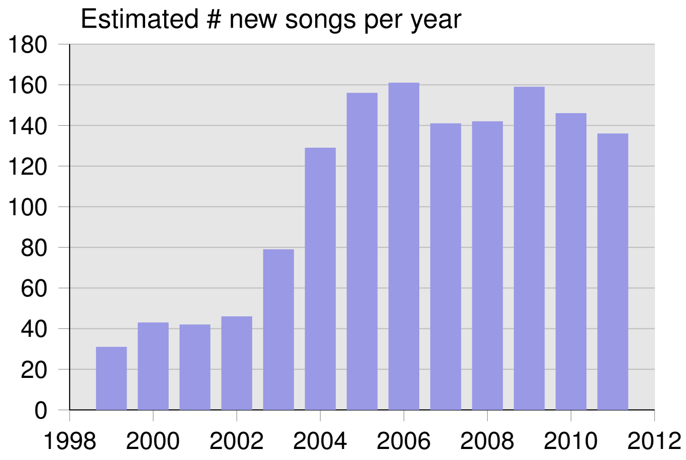


FIGURE 3.5
The “digital renaissance” effect in music

of the top 500 blockbusters of all time. Number 1 is missing as it falls outside of the graph’s limits. Can you guess which one it is? Hint: it’s quite old.)

Finally, while there is significant uncertainty regarding the output of entertainment production functions, there is also significant uncertainty regarding input levels. This is not specific of entertainment. For example, it is common wisdom that construction projects are subject to considerable variance in terms of total cost (or time for completion). In the case of entertainment, one interesting example is that of the TV game show *Power of 10*, where variability of input costs may have been one of the factors leading to the show’s cancellation. Specifically, the prize paid to contestatns may vary all the way from zero to 10 million dollars. Box 3.1 has the details (this box is largely based on a *Planet Money* episode).

3.2. QUANTITY AND QUALITY

It’s no secret that the cost of creating and distributing entertainment content have decreased in the past two decades. As a result, and as mentioned in Chapter **??**, the 21st century has witness a significant “explosion” in content production. Economist Joel Waldfogel refers to the this process as the “digital renaissance” in media. Much of this section is based on his ideas.

Box 3.1: Power of 10

The TV game show *Power of 10*, created by Michael Davies and produced by Vincent Rubino for Sony Pictures Television, first aired on CBS on August 7, 2007. In many ways, the show was similar to countless other TV game shows. It did, however, include some specific features.

The rules are relatively simple: At a given level, a contestant is presented with a question regarding a national poll. The questions are chosen purposely not to require trivia-type knowledge on the contestant's part — and to be funny. For example, one question asked, "What percentage of Americans said they are afraid of circus clowns?"

The organizers know the result of the pre-game national poll; the contestant does not, and the challenge is to guess a range of values where the actual number lies. The stakes increase from question level to question level — by a factor of 10. The range for a correct answer, in turn, decreases as the stakes increase, from a 40% range for a $1,000 question all the way to the exact percentage value for a $10 million question.

High stakes is one of the key ingredients in entertainment, and Power of 10 made a strong bet. But the supply of suspense in the form of high stakes has its costs: enormous uncertainty regarding the show's cost. Moreover, contestants themselves may or may not be willing to take on large bets. The very first contestant, 19-year-old Jamie Sadler, guessed ranges correctly all the way to $1,000,000, at which point he walked away from guessing the exact percentage of American women who consider themselves feminist. Several subsequent contestants did not make it past the first round.

The show only ran in the US for one season. It was subsequently adopted in dozens of other countries.

Consider the case of music. Figure 3.5 shows the estimated number of new songs from 1998 to 2012 (source). As can be seen, compared to the 1990s the number of new songs increased by a factor of 3 to 5 in one decade. This is an unprecedented growth rate. Similar "explosions" in content supply have been observed in book publishing (Amazon now lists dozens of millions of different titles) and movies (especially indie productions). What are the main implications of the rapid increase in the supply of entertainment goods that was brought about during the first decades of the 21st century?

UNCERTAINTY, ORDER STATISTICS, AND QUALITY

Has the "digital renaissance" brought about any value to the music, movie, publishing, videogame, etc, industries? One possible take is that it's just more of the same. For example, several music critics complain that there is a lot of repetition in the supply of new music. An even more negative view is that the average quality of new music is necessarily lower as we dig deeper into the talent pool. However, it is possible that both of the above statements be true and still the case can be made that the explosion in content supply has brought about significant benefits to consumers. The key to understanding this possibility is to understand the concept of order statistics.

An order statistic is simply the value of the highest or second lowest or top 5 or whatever of a given set of values. For example, suppose that the national basketball team of each country is composed by its tallest players. (I know this is not the case: you don't need to be tall in order to be a basketball star, and many tall people are rather lousy at basketball; but for the sake of illustration let us stay with this assumption.) Then the average height of each national basketball team is the average height of the five tallest people. The average of the top five values of a sample is an example of an order statistic. In this case the sample is given by all men or all women in a given country, as the case may be, and the value in question is height.

Why are we interested in order statistics? Let us continue with the height and basketball example. Let us compare two different countries: China and the Netherlands. The average height of Chinese men is 5′ 6.5″ (169.5cm), whereas the average height of Dutch men is 5′ 11″ (180.8cm). However, the top five players on China's 2019 men's national basketball team was 6′ 10.6″ (2.098m), whereas the corresponding value for the Dutch team was a mere 6′ 9.9″ (2.08m).

How do we get a reversal in order as we move from the population average to the average of the top 5? The key is that, in 2019, there were about 700 million Chinese males but only 9 million Dutch males. As we pick the top 5 from a larger pool we get a larger average. That's what order statistics do for you.

Wikimedia Commons

Dutch men are taller on average than Chinese men, but the Chinese men's national basketball team is taller than the Dutch one.

What does this all have to do with music? Suppose we can measure the quality of a new song as $s = 0$ (really bad), $s = 1$ (OK) and $s = 2$ (hit). Suppose that each new song has a $1/3$ change of falling into each of these categories. If there is only one new song each year, then average quality of new music is given by

$$\mathbb{E}(s^*) = 0 \times \tfrac{1}{3} + 1 \times \tfrac{1}{3} + 2 \times \tfrac{1}{3} = 1$$

Now suppose that two new songs are written each year and suppose moreover that only the best is selected by distributors and consumers. We now have nine different possible outcomes to consider: Song 1 may be bad, OK or hit quality (0,1,2); and likewise Song 2 may be bad, OK or hit quality (0,1,2). The following matrix represents all nine possible combinations, with each cell representing the quality of the best song (that is the top order statistic of the sample of all new songs):

		Song 2		
		0	1	2
	0	0	1	2
Song 1	1	1	1	2
	2	2	2	2

For example, if Song 1 scores $s = 2$ (hit) and Song 2 scores $s = 0$ (bad) then the best new song scores $s = 2$ (hit), and so forth. The result of all of these combinations is that there is now a $1/9$ chance that the best of the two new songs is a bad song: such outcome would require both new songs to be bad. By contrast, there is a 5/9 chance that the best new song is a hit. Overall, the average quality of the best new song is now given by

$$\mathbb{E}(s^*) = 0 \times \tfrac{1}{9} + 1 \times \tfrac{3}{9} + 2 \times \tfrac{5}{9} = \tfrac{13}{9}$$

which is greater than 1, the previous expected value if there was only one new song. In other words, even though the average quality of each new song has not changed, the average quality of the *best* song has increased, on account of the number of new songs having increased. In other words, order statistics is good news for entertainment consumers: The context explosion effect on quality is likely to be positive. In fact, if the number of new content units increases by a lot — as it did in music, film, publishing, etc — then it is quite possible that average quality of the best increases even if average quality of each element decreases.

In the previous section I noted that uncertainty is an important element in the process of entertainment supply. In other words, the ϵ in equation (3.2) can be of considerable importance. This is very important for the order-statistics story presented above. To see why, suppose there is no uncertainty in the entertainment supply process. Even if only one song is produced each year, then we can choose only to produce one that will be a hit. If this is the case, increasing the number of new songs produced does not increase the average quality of the top song. In fact, if we *really* increase the number of new songs then at some point we will dig deeper and deeper into the talent pool, thus lowering the quality of the *additional* songs produced each year. This is the argument of the pessimists with respect to the music and other entertainment industries, and I'm sure there is something to it.

> Imagine that traditional gatekeepers were really good at predicting what was going to succeed. In that case, if the cost of distribution fell relative to revenue, they could greenlight more of the products they used to say no to. But all of those newly greenlit products would be worse than the former lowest threshold, so it wouldn't be that helpful to consumers to get a lot of not-very-appealing stuff (source)

However, it is very hard to predict ex-ante whether a new song will be a hit, and this is the reason why numbers matter. In the above example I made the extreme assumption that a new song can have value 0, 1 or 2 with equal probability, that is, we know very little about its quality. In reality, producers are good at predicting whether a new song will be a hit. However, they are not perfect. This (partial) unpredictability of success is what makes the content explosion an explosion in quality as well. Not average quality, but quality of the best.

Is there any evidence that the content explosion has actually improved quality? A similar question is: did the digital renaissance bring us content that would not have been produced before? In book publishing, we do observe a significant increase in self-published books, which suggests the answer is yes. The *USA Today* best-seller list (150 titles per week) now typically includes 15% self-published titles, a percentage that used to be zero. That said, the *New York Times* notables list, which focuses on more erudite material, continues to include no self-published books. This suggests that self-publishing — the core of the digital renaissance in book publishing — is affecting the mass-market segment but now the higher-brow segment.

3.3. DISCOVERY

Polymath Herbert Simon once wrote that

> A wealth of information creates a poverty of attention and a need to allocate that attention efficiently.

Take the case of music. As Figure 3.5 suggests, the number of new songs has increased considerably since the start of the century. So far in this section we have argued that an increase in quantity of content may lead to an increase in quality, the reason being that the best pieces of content are selected from a wider pool.

One key step in the above process is selection. Specifically, we have (implicitly) assumed that the best content is selected from the pool of new content created each year. However, if the supply of content has increased by a lot — as Figure 3.5 suggests is the case for music — then the question may be asked whether the content that is selected is indeed the best. In fact, one may inquire how "hits" get selected in the first place.

We know that it's hard to predict what will become a hit. But even if experts or lead consumers identify a particular content as a hit, what is the mechanism whereby such content is adopted by the market as a whole? In other words, how do consumers discover new content? To paraphrase Simon, *wealth of entertainment content creates a poverty of awareness and the need to discover new content efficiently*.

The mechanisms for **content discovery** vary from industry to industry. In movies, for example, the rise in the supply of independent movies (a.k.a. indies) has led to the emergence of various festivals which effectively function as gatekeepers. According to the now-defunct website withoutabox.com, there were at least 5,000 movie festivals in 2017. Being accepted at one of these festivals is an important step — a necessary step, one might say — in order to gain access to the wide public. Nowadays, Amazon prime video direct helps "studios, distributors, and independent filmmakers reach audiences worldwide," and through Amazon's recommender system helps consumers discover new content as well.

In book publishing, recommender systems such as Amazon's also play an important role. Sites such as goodreads.com promise that "we'll give you surprisingly insightful recommendations," and by many readers' testimony they do deliver on their promise. (Goodreads.com has since been acquired by Amazon.)

In music, radio has traditionally played an important role (and continues to do so, though less so than in the radio golden era). Music critics — individuals or websites such as Pitchfork.com — are also influential. Finally, similar to movies and books, recommender systems — e.g., Spotify's playlists — also help selecting from an increasingly large pool of available content.

The above channels may be summarized into two different categories: expert advice and crowd advice. In what follows, we provide some evidence regarding the effectiveness of these channels in the increasingly important discovery process.

BEST SELLERS

Many newspapers publish lists of best-selling books. Do these lists affect consumers' awareness and purchase choices? In other words,

pxfuel.com royalty-free photo

Inclusion in the *The New York Times* best-seller list increase awareness of debut authors.

are past consumer choices a factor in the discovery process in new books?

Statistically, one problem with addressing this question is to circumvent the so-called endogeneity problem. The fact that we observe a correlation between ranking in best-selling lists and subsequent sales does not imply that the former causes the latter. Correlation does not imply causality. It could simply be that some books are better, and book quality implies both a high rank in best-selling lists and high sales to subsequent consumers, even if the latter do not pay any attention to a particular best-selling list.

To make this point clear, suppose that I create a new newspaper, *Cabral News*, with a total readership of one (my mom). One of my newspaper features is a list of best-sellers, which I compile from data given to me by a local bookstore. Although this is a small bookstore, and certainly one that does not represent the world of book publishing and reading, it is likely that the titles included on my list today will be highly correlated with next month's worldwide sales of each title. I would love to claim that my January best-selling list *caused* worldwide sales in February. I could make that claim and I could offer the high correlation between February sales and January ranking as evidence for my claim. But I know that even if my mom — the sole reader of *Cabral News* — were to talk to a lot of people, it's very unlikely that this was the cause for high sales in February.

Faced with this conundrum, economists try to find ways of identifying causality other than simply measuring correlation. In Chapter 2 we looked at the idea of computing historical counterfactuals to identify the effect of price changes on demand. In the present context, one alternative path is to make use of a so-called **natural experiment**.

Consider the *New York Times* best-seller list. It's been published as a regular feature since 1942. Although many other prominent lists now exist, the *Times'* is generally considered the most influential in the industry. Does being listed as a *New York Times* bestseller cause an increase in sales? How can we avoid the correlation/causality problem mentioned before? It turns out that the process used to generate the *New York Times* list (surveying a limited number of booksellers) is inexact, that is, there may be books which should have been included (actual sales were in the top) but were not included (the *Times*-surveyed sales were not in the top). These 'mistakes' provide a so-called natural experiment of the theory that the *Times* list has a causal effect on sales.

Specifically, during the years 2001–2002, there were 182 instances in which a hardcover fiction book was not listed as a *New York Times* bestseller when in fact it should have been. The majority of these (roughly 70%) were narrow misses, that is, books that would just have made it to the list if the *Times* had computed sales correctly. We can now estimate the causal effect of inclusion in the *Times* list by restricting to these observations.

The results suggest that the effect is positive and significant: The week a book appears on the *New York Times* best-seller list, its sales get a boost of about 8%. The increase is particular significant for debut authors, which suggests that the effect of best-seller lists is to increase consumer awareness (the "Herbert Simon effect").

OPENING WEEKEND AT THE BOX OFFICE

Consider now the case of movies. A particularly important ranking — the equivalent of a book best-seller list — is the ranking during a movie's opening weekend. Particularly important in this ranking is the #1 position. Every Monday morning, *Variety* — the industry's leading periodical — announces the identity of the weekend's top grossing film. We observe a high correlation between high sales in the opening weekend and subsequent sales. However, as usual this may simply be a case of correlation, not causality: The fact that the #1 movie does very well in subsequent weekends may simply result from the fact that, being a blockbuster, it does well during the opening weekend *and* during subsequent weekends. Correlation, not causality.

In this context, one strategy for causal inference is to implement a **regression discontinuity** design. Here's the idea: If being #1 does not make a difference beyond the movie's underlying appeal, then being #1 during the opening weekend or being #2 by a difference of a dollar (i.e., a very small difference) should not make a difference in terms of subsequent weekends' performance. In other words, two movies with the same opening weekend sales should perform equally well in subsequent weekends. However, being #1 matters even in this case.

Further evidence shows that the #1 effect is lower when the leading movie was heavily advertised in the week prior to opening day. This suggests that the effect of being #1 is to create greater awareness of the movie's existence. In other words, pre-opening advertising or being #1 are substitute channels for creating greater awareness of the movie's existence (once again, the "Herbert Simon effect").

PITCHFORK BEST NEW MUSIC

Consider now the case of music. In the pre-digital era, radio play and label promotion fulfilled a central role in the discovery of new content. The digital era brought us dramatically lower production costs, as mentioned earlier; but it also created multiple new platforms and web sites with abundant and free information, including in particular music reviews. To be sure, critics and reviews have always existed and have always played a role. However, considering the ease of access to webpages full or ratings, rankings and reviews, one may conjecture that reviews are particularly important.

> Along with many other effects of digitization, the Internet has led to an explosion of outlets providing critical assessment of new music. Since 1995 the number of outlets reviewing new music — and the number of reviews produced per year — has doubled. (source)

Consider specifically the case of Pitchfork, one of the leading music websites. (An *LA Times* reviewer characterizes the site as "an essential part of the iPod generation's lexicon.") Since 2003, Pitchfork lists three new titles with the classification "Best New Music" (BNM).

Does the BNM seal have any effect on discovery of new music? As usual, determining a causal effect of this sort is not an easy task. A simple correlation analysis shows quite clearly that BNM titles are played more frequently. However, this largely follows from the simple fact that Pitchfork's critics and music listeners have similar tastes.

Pitchfork attributes a score to all of the music it reviews (and assigns the BNM label to a subset of these). Statistical analysis suggests that the BNM seal has an effect on the number of listeners and number of plays over and above the score given by Pitchfork critics and other critics. Moreover, this effect is (a) higher for new music acts that; and (b) lower during the years when Spotify was already in play. Fact (a) is consistent with the "Herbert Simon effect" mentioned earlier. Fact (b) is consistent with the idea that crowd-based recommendations (Spotify playlists based on plays) are a substitute to expert advice.

3.4. INCENTIVES AND CREATIVE TALENT

What makes creators tick? Ultimately, that is the really important question when it comes to the supply of entertainment content. One view is that the creative muse is something from the soul, something imaterial, something as far from economics as possible. But creators must eat like everyone else. And creators understand extrinsic motivation like everyone else.

To put a more specific question: does the intellectual property (IP) right over an artist's creation — and the corresponding income flow — lead the artist to create more and better content, or are IP rights an afterthought in the creative process?

As often is the case with social science, this is a hard question to answer because we have very few instances of before and after, with and without, to allow us to estimate the effect of IP rights on creative effort. One of such few instances is precisely the time when copyrights were first introduced in a given country, for example, in the Italian states during the early 19th century. We next turn to this case study (source).

NINETEENTH CENTURY ITALIAN OPERA

Back in the 1800s there was no soccer or basketball, no Netflix or Spotify, no YouTube or Internet. By contrast, in Italy — in the states that today comprise Italy — opera was big. Very big. Opera was not an erudite and somewhat elitist form of entertainment as it is nowadays, it was a popular form of entertainment. Many new operas were constantly written, and there was great excitement and demand for each performance.

Before our story begins, around the end of the 18th century in what is now Italy, a composer's work was not legally protected from copy and imitation. As such, a composer's income would essentially be derived from the initial performance. After that, piracy was rampant and authors were left with little or nothing. Impresarios would

> either steal an authentic score (as a rule by bribing a copyist) or pirate it by getting a minor composer to work up a new orchestral setting from the printed vocal score. ... An impresario who wanted to give a recent opera would commonly try to knock down the cost of hiring the authentic score by pointing out that he could get one elsewhere at half the asking price (source).

In this context, there was little a composer could do besides "recycling" some of their compositions and take them to a different town.

The life of a composer was not an easy one. For most, selling opera scores was the main if not the only source of income. Take the great Gioachino Rossini, for example.

> His mother ... was a seconda donna of very passable talents. They went from town to town, and from company to company; the husband playing in the orchestra, and his wife singing on the stage. Poverty was of course the companion of their wanderings (source).

Meanwhile, in neighboring France, things were changing rapidly: The age on Enlightenment, the 1989 Revolution, Napoleon's rise to power. Two events in particular came to be of great importance for Italian opera writers. First, Napoleon's military expansion reached Italy with the invasion of the Island of Sardinia in 1796. Second, some Italian states adopted France's 1793 copyright law.

Napoleon Bonaparte (portrait by Andrea Appiani). Gioachino Rossini (portrait by Étienne Carjat).

Wikimedia Commons

A key word in the last sentence — *the* key word, for our purposes — is *some*. For a variety of reasons, the Italian states of Lombardy and Venetia adopted French laws early on, whereas other states took a lot longer to do so. The details are complicated and involve many military and political considerations. From the point of view of our social science research, the important point is that copyright law — in particular the incremental way in which it was introduced in Italy — was a side effect of the gradual conquest of the Italian states. This is nirvana for a researcher: we have what's known as a **natural experiment**, the closest you can get to a laboratory experiment when using historical data. Specifically, we can use the gradual introduction of copyright law to study its differential effect on opera composers.

Figure 3.6 illustrates the results from this historical investigation. Prior to the introduction of copyright law in Lombardy and Venetia, the number of new operas introduced each year was not very different from that in other states. By contrast, after 1801 — the date when copyright is introduced — we observe the emergence of a significant gap between Lombardy and Venetia and the other Italian states.

Notice that there is also a slight positive trend in the number of operas introduced in the states with no copyright law. This illustrates the benefit of having a control group. If we were only to examine the evolution of operas in Lombardy and Venetia we would observe a post-1801 increase. However, this increase could have been caused by factors other than the introduction of copyright protection, specifically other factors which occur simultaneously with the introduction of copyright law. The ability to create a control group (states where copyright protection was not introduced) allows us to perform a

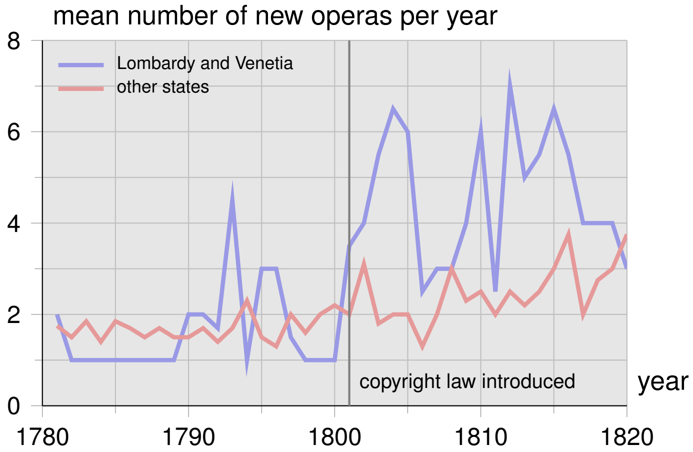


FIGURE 3.6
New operas per state

difference-in-differences analysis (similar to what we did when estimating Cinemex's demand elasticities in Chapter 2). Specifically, we compute (a) the difference between Lombardy-Venetia and the other states before 1801; (b) the difference between Lombardy-Venetia and the other states after 1801; and finally (c) the difference between (b) and (a), a difference in differences.

A closer look at historical accounts shows that, with the advent of copyright protection, composers were paid for repeat performances, not just the first performance. For example, a 1802 (i.e., post-copyright) contract between impresario Francesco Benedetto Ricci and composer Giuseppe Mosca specifies that

> Francesco Benedetto Ricci is obliged to pay Giuseppe Mosca the sum of 3,500 francs for the score and 250 francs for each repeat performance in the current season (source).

Better financial terms, in turn, led composers to compose more operas. Specifically, Lombardy and Venetia — the early copyright adopters — created 2.2 more new operas per year after 1801 compared with other Italian states without copyrights. This increase results from composers in Lombardy and Venetia working harder. It also results from Italian-born émigré composers returning to Lombardy and Venetia after 1801.

Equally important, the increase in quantity of operas did not come at the cost of quality, quite the opposite. There are many ways of measuring quality. One that I find particularly interesting is the test of time: was the piece still being performed at New York's Metropolitan Opera between 1900 and 2014? The data suggests that, by this measure, Lombardy and Venetia experienced a 7-fold increase with respect to no-copyright states.

COPYRIGHT IN THE 21ST CENTURY

Under the international Berne Convention for the Protection of Literary and Artistic Works of 1886, the signatory countries are required to provide copyright protection for a minimum term of the life of the author plus fifty years. Additionally, they are permitted to provide for a longer term of protection.

In the US, the Copyright Term Extension Act (CTEA) of 1998 extended copyright terms in the United States to life of the author plus 70 years and for works of corporate authorship to 120 years after creation or 95 years after publication, whichever end is earlier. This implies that a work created and registered in 1922 would fall under the old law and be turned into the public domain by 1998, whereas a work created and registered in 1923 would fall under the new law and be turned into the public domain by 2019. To give a prominent example, Disney's Mickey Mouse, having first appeared in 1928, will be in the public domain by 2024, 20 years after it was previously scheduled. (Some people derisively refer to CTEA as the Mickey Mouse Protection Act.)

The Walt Disney Company, the estate of composer George Gershwin, Time Warner, Universal, Viacom, and the major professional sports leagues (NFL, NBA, NHL, MLB), all supported the extension of the copyright term. Part of their argument was that some works would be created under a longer copyright that would never be created under the existing copyright.

At first, the evidence from Italian operas in the 19th century seems to provide credence to this argument: With no copyright, there were few operas and many variations on the same theme. With copyright, the number and the quality of operas increased substantially.

However, there is more to the Italian 19th century experience. Having first introduced copyright with a term of life plus 10 years, Lombardy and Venetia extended the term to life plus 30 years in 1840, and then to life plus 40 years in 1865. Contrary to the initial introduction of copyrights, the data shows no evidence of a positive effect of copyright extension on supply.

This is, after all, a classical example of the economics law of **decreasing marginal benefits**. If you don't have any copyright protection, then a bit of protection matters a lot. By contrast, if you already have some copyright protection, then an extension of the term may not benefit you that much, that is, you may not care that much about it — or at least it is unlikely to make you work harder.

KEY CONCEPTS

production function · total factor productivity · rival goods · non-rival good · frequency · normal distribution · content discovery · natural experiment · regression discontinuity · natural experiment · difference-in-differences analysis · decreasing marginal benefits

REVIEW AND PRACTICE PROBLEMS

■ **3.1. Entertainment supply.** What is special about the supply of entertainment goods?

■ **3.2. Inputs and outputs.** "In the entertainment business, you get what you pay for." Comment, if possible with examples.

■ **3.3. Unique inputs.** True or false (justify your answer): The entertainment production function differs from other production functions in that some inputs into the entertainment function are unique.

■ **3.4. Black swans.** What is the statistical meaning of the concept of "black swan"? How does it apply to entertainment industries?

■ **3.5. Power of 10.** Consider the TV show Power of 10, described in Box 3.1.

(a) What were the main differences between *The Power of 10* and other game shows — and why did these make the show so promising?

(b) What does economic theory have to say about the contestants' decision to go for $10 million?

(c) Is there anything in *The Power of 10* you would change were you to run the show for a second season?

■ **3.6. Audio books.** An argument can be made (see Section 3.2) that the "digital renaissance" (or "content explosion") has benefited consumers in the form of an increase in quality. Does this argument apply to audio books as well? Why or why not?

■ **3.7. Quantity and quality.** In the digital world, quantity implies quality. Comment.

■ **3.8. Attention.** Polymath Herbert Simon once wrote that

> A wealth of information creates a poverty of attention and a need to allocate that attention efficiently.

How does this relate to the so-called digital renaissance?

■ **3.9. Copyright.** Is copyright an important factor in providing proper incentives to content creators? Contrast the 19th and 21st centuries.

CHAPTER 4
PRICING

In 2012, Melissa Leong — writing under the pseudonym Wynne Channing — published her first eBook on Amazon. As she later recalled,

> On June 6, almost a month after I made the decision to self-publish a book (or two, I hope), I uploaded my manuscript and cover to Kindle Direct Publishing. I set the price at $2.99.

Why $2.99? What took Leong to set that particular price? What if she were also to sell a paperback version and a hardcover one: what prices should she set?

This chapter deals with the issue of price setting in media and entertainment. We begin in Section 4.1 with the very basic economics of optimal pricing, the case of a seller who sells one single product. Then in Section 4.2 we consider the more realistic case of a seller with multiple products or revenue streams, especially when these revenue streams are related to each other. For example, as we saw in Section 2.3, movie theaters sell movie tickets and popcorn, and the sales of one are related to the sales of the other. Section 4.3 deals with another extension that is quite important in the world of media and entertainment: the pricing of goods subject to dynamic of social effects (as in "I buy it because you also buy it"). Finally, in Section 4.4 we focus on one of the most distinctive features of entertainment goods: the fact that the relation between seller and buyer is more than the simple relation between a seller and a buyer.

4.1. OPTIMAL PRICING

If you worked as a consultant to Melissa Leong, you might tell her that the more sensitive demand is to price changes, the lower a price you should set. This is fairly intuitive and consistent with what we've already seen in Section 2.3. Another point you might make is that, everything else constant, the higher the production cost, the higher the optimal price. For example, it costs more to print a hardcover book than it does to print a paperback. For this reason, one would expect the optimal price of a hardcover book to be greater than that of a paperback. (In Chapter 5 we will consider additional reasons for a price difference between the hardcover and paperback versions of a book.)

Let the the sensitivity of demand to price changes be measured by the value of the price elasticity of demand, ϵ, a concept we introduced in Section 2.2, and let the cost of each unit be given by c. Then it can be shown that the optimal price level, p^*, is given by

$$p^* = \frac{c}{1 + \frac{1}{\epsilon}} \tag{4.1}$$

This expression may be rewritten (check it) as

$$p^* = \left(\frac{|\epsilon|}{|\epsilon| - 1} \right) \; c \tag{4.2}$$

Consider for example the pricing of the hardbound version of *What Kills Me*. Suppose that the estimated elasticity is given by -2 and suppose that its value is constant (i.e., is the same at all points of the demand curve). Suppose moreover that each hardcover copy costs \$10 to produce. Then $|\epsilon| = 2$, $c = 10$, and the optimal price is given by $p^* = (2/(2-1)) \times 10 = 20$, that is, \$20 per hardcover copy.

Economists and businesspeople are wont to measure price levels by computing the price margin or the price markup. Unfortunately, different people mean different things by these terms: keep that in mind. For our purposes, we will define margin m and markup k as follows:

$$m \equiv \frac{p-c}{p}$$
$$k \equiv \frac{p-c}{c}$$

It can be shown that the "elasticity rule" presented above, that is, equation (4.1), corresponds to

$$m = \frac{1}{-\epsilon} = \frac{1}{|\epsilon|} \tag{4.3}$$
$$k = \frac{1}{-\epsilon - 1} = \frac{1}{|\epsilon| - 1}$$

Note that all of the above equations can be derived one from the other: they are all the same result, in other words, they are different ways of presenting the same result.

In the above example, margin is equal to $(20-10)/20 = 50\%$, whereas markup is equal to $(20-10)/10 = 100\%$. Since $|\epsilon| = 2$, we have $1/|\epsilon| = \frac{1}{2} = 50\%$ and $1/(|\epsilon|-1) = 1/(2-1) = 100\%$. So, we get $50\% = 50\%$ and $100\% = 100\%$. Bingo! (Exercise 4.4 might help solidify these concepts.)

To go beyond the math and a bit into economic intuition, consider the "margin" version of the elasticity rule, that is, equation (4.3). It states that the optimal margin m should be the inverse of the absolute value of demand elasticity, $|\epsilon|$. This encapsulates the first idea that we started the section with: the less sensitive demand is to price changes — that is, the lower the value of $|\epsilon|$ —, the higher the optimal margin set by a seller. As to the effect of cost, (4.2) states that, for a given value of ϵ, optimal price should vary proportionately with respect to cost: increase cost by 10%, say, and you should increase price by 10% as well.

PRICING OF MEDIA AND ENTERTAINMENT

Speaking of cost and optimal pricing, one of the most salient features of many entertainment goods is precisely that they have zero variable cost: Once you've recorded a song, the cost of making it available to an additional listener is essentially zero. How does this affect optimal pricing?

If we were to apply (4.1) directly we would get $p = 0$, which clearly does not make sense: giving your content away for free is not the way to maximize revenues if you only have one revenue source. Later we will consider the case when $p = 0$ is indeed an optimal policy. However, this only happens when you plan to collect revenue from other sources (now or in the future).

If there are no variable costs, then maximizing seller profit is the same as maximizing seller revenue. As we saw in Section 2.2, an increase in price leads to an increase in revenue if and only if $|\epsilon| < 1$. The idea, if you recall, is that $|\epsilon| < 1$ implies that the percent variation in q is lower than the percent variation in p. As a result, when price increases the negative effect (lower q) is lower than the positive effect (higher p) on revenues.

Similarly, a decrease in price leads to an increase in revenue if and only if $|\epsilon| > 1$. Together, this implies that, when setting optimal price p^*, ***it must be that demand elasticity is equal to one in absolute value***, that is, $|\epsilon| = 1$.

We can formally prove this by contradiction (a common method in mathematics): Suppose $|\epsilon| \neq 1$. Then either $|\epsilon| > 1$ or $|\epsilon| < 1$. If $|\epsilon| > 1$, then you're better off by decreasing price. If $|\epsilon| < 1$, then you're better off by increasing price. Therefore, if $|\epsilon| \neq 1$ then the current price cannot be optimal. End of proof.

This line of proof reflects the "think at the margin" approach so common in economics. In Section 1.2, we considered the problem of finding the optimal number of dates for Madonna's *Madame X* tour. The idea is that she should evaluate the benefits (monetary and otherwise) from one concert more or one concert less. Similarly, when finding the optimal price level we should compare the costs and benefits of a slightly lower price (or a slightly higher price).

A DIFFERENT PATH TO THE SAME RESULT

In this subsection I propose a slightly different — slightly more formal — approach to the optimal rule $|\epsilon| = 1$ for entertainment goods. If you understood the intuition from the previous subsection and math is not your forte, feel free to skip this subsection. If formality helps you understand things better, then this subsection is for you.

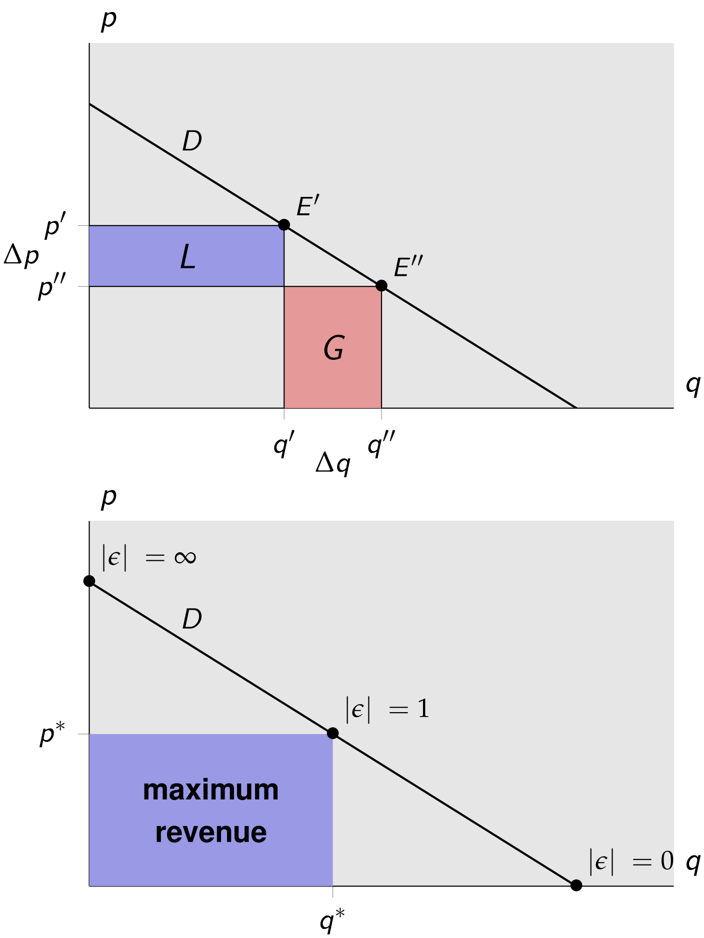


FIGURE 4.1
Optimal pricing: graphical intuition

Consider the top panel of Figure 4.1. The curve D represents the demand curve (which for simplicity we assume is linear). Suppose that initially price is given by p', so that the quantity demanded is given by q'. Consider the alternative price p'', lower than the initial price p'. By the law of demand, we know that such a price decrease will lead to an increase in demand, specifically, from q' to q''. What are the costs and benefits of lowering the price from p' to p''?

The cost of a price decrease, measured by the area L in Figure 4.1, corresponds to the loss in margin. Initially you were getting p' for each unit sold, now you only get p''. It follows you are losing $p' - p''$ per unit sold. Define $\Delta p \equiv p'' - p'$ as the price change. So, if you are selling q units then the cost of lowering price is given by

$L = -q \times \Delta p$. (We need a minus sign because Δp is negative, so if we want to think of L as a loss we much change the sign.)

The benefit of a price increase, measured by the area G in Figure 4.1, corresponds to the increase in sales. Initially you were selling q' units, now you are selling q'' units. It follows you are benefiting from an increase in sales given by $\Delta q \equiv q'' - q'$. If you sell for a price p, then the gain from bigger sales is given by $G = p \times \Delta q$.

Is it worth it to lower price from p' to p''? The answer is yes if and only if the benefit is greater than cost ("think at the margin"), that is, if $G > L$. Plugging in the expressions derived earlier, we get

$$p \times \Delta q > -q \times \Delta p$$

Now divide both sides by $-q \times \Delta p$. Since $\Delta p < 0$ (price is decreasing), we are effectively dividing both sides of the inequality by a positive number, which is totally legit. We get

$$\frac{p \times \Delta q}{-q \times \Delta p} > 1$$

or simply

$$-\frac{\Delta q / q}{\Delta p / p} > 1 \tag{4.4}$$

Now, in Section 2.2 we saw that demand elasticity is defined by

$$\epsilon = \frac{\%\,\Delta\text{ quantity}}{\%\,\Delta\text{ price}} = \frac{\Delta q / q}{\Delta p / p} \tag{4.5}$$

Since moreover $|\epsilon| = -\epsilon$ (since the demand curve is downward sloping, we have $\epsilon < 0$) we conclude that a decrease in price increases revenues if and only if $|\epsilon| > 1$. A similar argument implies that, if $|\epsilon| < 1$, then an increase in price leads to an increase in revenues. We thus conclude that the optimal price corresponds to the case when $|\epsilon| = 1$.

The bottom panel in Figure 4.1 shows the optimal price p^*. If the demand curve is linear — a particular case, but one that economists consider frequently — then the demand elasticity becomes very large (in absolute value) as q approaches zero and very small (in absolute value) as p approaches zero. The point where $|\epsilon| = 1$ turns out to be the mid-point between the two axis intercepts. (I note again that this is a particular feature of linear demand curves. Generally speaking, the point where $|\epsilon| = 1$ will be somewhere between $p = 0$ and $p = \infty$.)

4.2. MULTIPLE REVENUE STREAMS

The previous section dealt with optimal pricing in the simplest of all cases: the seller has one single revenue stream and sets only one price. Most real-world examples one can think of depart from this extreme case, that is, most sellers in the real world have multiple revenue streams. How does this affect the above results regarding optimal pricing?

SUBSTITUTES AND COMPLEMENTS

Let us start with the easy case: If a seller's multiple revenue streams are totally independent from each other, then you simply apply the optimal pricing rules from Section 4.1 to each revenue stream. For example, if Kabral Ventures sells both ice cream and truck tires then Kabral Ventures should apply the optimal pricing rules considering the specific values of cost and demand elasticity of ice cream and the specific values of cost and demand elasticity of truck tires.

The more interesting — and more challenging — case is when a seller's multiple revenue streams are related to each other. Consider first a seller who sells two substitute products. (See Section 2.2 for a definition of substitute products.) We already saw one such example, namely Cinemex (cf Section 2.3): In addition to selling Wednesday tickets, Cinemex also sells Thursday tickets. Suppose that we apply the optimal rule to Wednesday ticket prices. In particular, in Section 2.3 we estimated that $\epsilon = -1.4$, so $|\epsilon| > 1$. This suggests Cinemex is better off by decreasing price: the increase in quantity of tickets sold more than outweighs the decrease in price. Suppose however that all of Cinemex's increased Wednesday sales come at the expense of fewer sales on Thursday. This is not entirely unrealistic: moviegoers observe that the Wednesday price is lower and decide to change their plans from Thursday to Wednesday (after all, it's the same movie). In this alternative interpretation of the data, lowering the price on Wednesdays would result in lower revenue. More generally, ***if a firm sells two substitute products then the optimal price tends to be higher than what the formula (4.1) would suggest***.

Consider now the opposite case: a firm that sells two complement goods (see Section 2.2 for a definition). The New York Mets, for example, sell tickets to baseball games as well as hot dogs. (We will

Bradley Gordon

When baseball ticket prices are lower, more fans come to the park, which in turn increases de demand for concession goods.

talk about this example at length in Chapter 5.) Suppose that the price elasticity of demand for tickets is approximately equal to 1 at the current price (and for simplicity suppose that there is one price only). According to our rule, this implies that the current price is optimal: the Mets cannot increase revenues by increasing or by decreasing price.

However, this analysis is incomplete. Suppose the Mets do lower ticket price. Even though the revenues from ticket sales will not change — the increase in number of tickets sold cancels out the decrease in price — the fact that more people come to the park means that more hotdogs will be sold; and this in turn implies an increase in revenues. (And considering the astronomical price that hot dogs are sold in baseball parks this must be a significant revenue increase indeed!) It follows that even when elasticity is approximately equal to 1 (in absolute value) it pays to lower price. More generally, ***if a firm sells two complement products then the optimal price tends to be lower than what the formula (4.1) would suggest***.

THE WATERBED EFFECT

A particularly important application of the above principles is what economists call the waterbed effect. Let us start with a historical example: wireless telecommunications. During the 1990s and early 2000s there was an important difference between phone plans in Europe and phone plans in the US. In Europe you did not have to pay when you received a call, whereas in the US you paid as a function of minutes used, regardless of whether those minutes were spent making calls or receiving calls.

Why this difference in pricing strategy? The answer is: termination charges. Suppose your network was Orange and you wanted to call a user on the Vodafone network. Then Orange would need to pay Vodafone a fee for the service of taking the call from Orange to the Vodafone user receiving the call. Moreover, this was typically a large fee. In this context, it made sense for networks not to charge users for receiving calls, in particular calls from a different network: Each time a network A user received a call from a network B user, network A would earn a large fee. Therefore, network A had an incentive to encourage its users to receive such calls, and not charging them was an obvious way to do so.

Over the years, European regulators drastically reduced termination charges, that is, imposed lower and lower ceilings on how much a network could charge a different network for the call termination service. Currently, the levels of termination charges in Europe are similar to those in the US, that is, very low. As a result, monthly plans have also become more similar to those in the US, in particular with little or no distinction between calls initiated or calls received.

More generally, we may define the **waterbed effect** as follows: ***when you stifle one revenue stream* (e.g., termination charges) *an alternative revenue stream typically emerges* (e.g., receiver charges)**. The waterbed metaphor is a reference to the effect that jumping on one end of a waterbed — or an air mattress, for that matter — has on anything light you place on the other end of the waterbed. (At this point I believe I'm supposed to add the "don't try this at home" disclaimer, so here it is.)

A FOCUS ON MUSIC

The music industry provides an interesting application of many of the concepts presented above. At a very broad level, we may distinguish three eras in the history of the music business:

- The performance service era
- The product service era
- The service distribution era

As the revenues from music recordings decreased, prices (and revenues) from concert tours increased

Before music reproduction technology (recordings) were possible, the only way to enjoy an artist's music was to listen to them perform live. For this reason, listening to music by professional musicians was a luxury that only the well-to-do could afford.

The advent of recorded music "democratized" the access to music. During this second era (the product service era), different technologies were used — vinyl, CDs, tapes, etc — but the common element was that a given piece of content was played by means of a physical device (turn table, tape player, etc) and the consumer needed to have a physical copy of the recording.

The digital revolution of the 21st century disrupted this system rather drastically. Once recordings could be expressed digitally *and* transferred through the Internet, consumers were able to share digital files and thus avoid the cost of purchasing CDs (then the standard way of playing music). File sharing sites such as Napster emerged and increased rapidly as broadband Internet access became more common. At the time, estimates of music piracy rate ranged from 23% in the US to 55% in Brazil.

For music publishers (music labels) this resulted in a terrible negative shock. Some referred to the first decade on the 21st century as "the day the music died" (a reference to the famous Don McLean song "American Pie"). The top panel of Figure 4.2 illustrates this trend. As can be seen, total revenues earned by the music labels dropped from about twenty billion dollars at the start of the decade to less than ten at the end of the decade. Reflecting on this period, David Goldberg, former head of Yahoo music, lamented that

> The digital music business has been a war of attrition that

> nobody seems to be winning. The CD is still disappearing, and nothing is replacing it in entirety as a revenue generator.

The initial reaction by the major labels was to fight file sharing tooth and nail. Multiple law suits were filed against websites such as Napster and PirateBay, as well as individual users (including teenagers). While these suits were successful in shutting down some of the sharing sites, they did not stop the downward spiral in music sales revenues.

Goldberg was right in stating that "the CD is still disappearing." He was wrong, however, in stating that "nobody seems to be winning." Apart from consumers, who were the main beneficiaries of the digital revolution in music, many revenues streams increased as CD sales weaned. In fact, we have here an interesting application of the waterbed effect. For much of the 20th century, musicians used touring as a means to increase sales of recorded music. It was common, for example, for a band to go on tour to promote a new album. In this world, musicians had an incentive to set prices at a lower level, conscious of the fact that cheaper tickets mean more fans at concerts, which in turn implies hire record sales.

Once record sales are drastically down, the incentive to "subsidize" concert goers is no longer present. This was reflected in a significant increase in concert tickets (well ahead of inflation) and increased revenues from touring. As in the cell-phone waterbed, we see a downturn in one revenue stream being compensated by an upturn in a different revenue stream.

But this is not all. Another important beneficiary of cheap digital music was the music hardware industry. Exhibit A is provided by mp3 players such as the iPod. The bottom panel of Figure 4.2 shows total sales (in units) of the various versions of Apple's iPod (after 2015 we have no data on iPod sales, but we know they were already very small). iPod prices varied a bit, but for the sake of illustration let us assume a price of \$100 (likely a lower bond of average price). In peak years, Apple was selling more than 50 million units a year. This is about 5 billion dollars, which corresponds to about one half of the drop in music revenue sales. If we consider that mp3 players are only one of the revenue streams benefiting from cheap digital music; and if we consider that the iPod was not the only mp3 in the market;

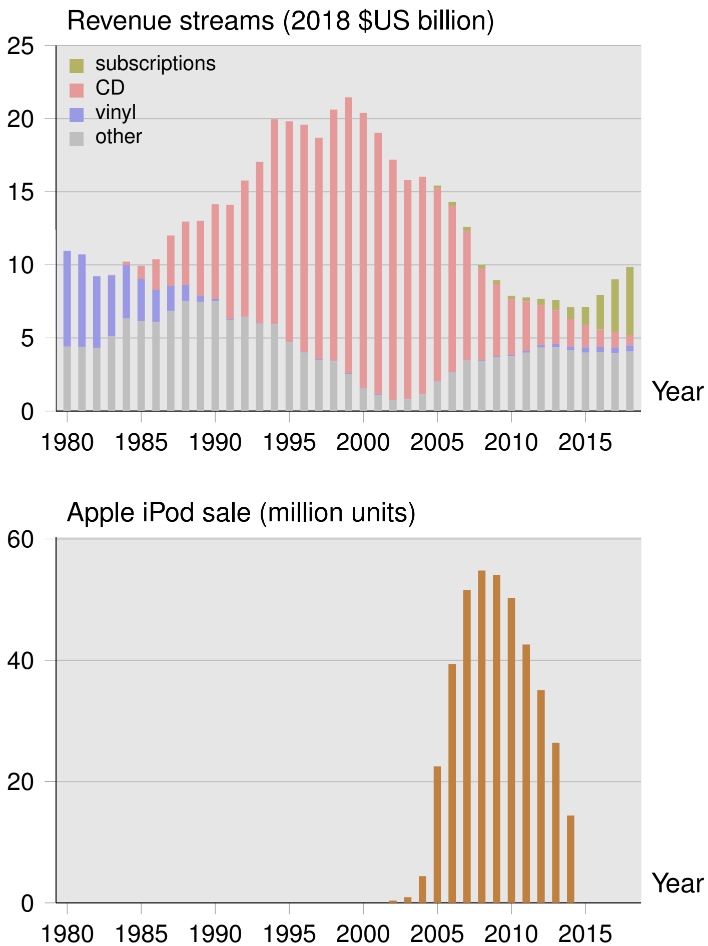


FIGURE 4.2
Music sales (top panel) and iPod sales (bottom panel)

then we conclude that, overall, music-related revenues did not do so badly after all. The first decade of the century may have been bad for music labels, but it was not so for all relevant players.

Hindsight 20-20, as they say. If one looks at the top panel in Figure 4.2, one realizes that there are cycles in the music business, largely determined by technology cycles. There was a time where CD sales were very small, namely the late 1980s. The CD technology was then just getting started. For music publishers, this new technology was great news: many music lovers replaced their vinyl collections with corresponding CD collections. Importantly, both the vinyl and the CD revenue streams were absorbed by the music labels. The digital revolution implied a different shift in revenue streams, one that benefited musicians and hardware manufacturers rather than music labels. It was a crisis for the latter, not for the industry as a whole.

Eventually, the key players realized that the digital revolution called for a change in the music business model. A first step was to create digital music sales services. Here, Apple's iTunes was the leader, largely because Apple was the first (or one of the first) companies to understand that there was a demand for quality certification: not content quality but rather quality of the music files. (File sharing sites were rife with corrupted flies, some perhaps placed there with the intent of sabotaging the sites.)

A second step in the evolution of the business model was the transition from individual file sales to the so-called service model. Here, Spotify was the leader and remains the main player. Meanwhile, music labels have been playing catch up and reorganizing their companies to adapt to the new reality of music revenue streams.

TWO-SIDED MARKETS

The concept of **two-sided markets** is similar (in fact, it generalizes) the previously mentioned waterbed effect. There is a lot of debate among scholars regarding the precise definition of two-sided markets. I will not get dragged into it. Instead, I discuss one of the most relevant examples of two-sided markets in the world of media and entertainment: TV content and TV advertising.

In the pre-cable years (yes, there was over-the-air television before cable), the major networks' main revenue source was advertising. In particular, viewers were not charged at all: all they needed was a TV set and an antenna. With the advent of online advertising, in particular targeted online advertising, the relative effectiveness of TV advertising decreased, and so did the demand for TV advertising. Moreover, viewers currently have more options to choose from besides TV as a source of news and entertainment, and so TV content reaches fewer eyeballs. Now that the major networks do not earn as much from advertising, their incentive to charge viewers is greater, and in fact they started to do so, either directly or indirectly (by indirectly I mean through charges to cable companies who then charge consumers).

Put this way, this example of two-sided markets looks just like another instance of the waterbed effect, and so it is. To a great ex-


BiblioArchives

There was a time when people did not pay to watch TV, that is, $p = 0$. The networks' revenues were almost entirely derived from advertising revenues.

tent, the waterbed effect is nothing but a particular case of pricing in a two-sided market: A given firm — say, CBS — has two revenue streams, viewer fees and advertising fees. These two revenue streams are clearly related: the more eyeballs I have, the more advertisers are willing to pay for a 30-second spot on my shows. There is also an effect going the other way: the more commercials I put on during my shows, the less viewers are interested in tuning in to my channel.

When setting advertising fees and viewer fees a seller should take these two-sided effects into account. Specifically, when setting viewer fees one should take into account that for each additional viewer the seller attracts, the seller is able to charge advertisers a higher fee. One way to think about this is that serving an additional viewer not only has no cost — as often happens with media and entertainment goods — but it actually implies a benefit! Or, if you prefer, the cost of serving an additional viewer is *negative*.

In terms of optimal pricing, this implies that the seller should set a lower viewer fee than it would otherwise (that is, ignoring the advertising benefits of an additional pair of eyeballs). In fact, if advertising revenues are sufficiently important, then the seller might even want to set a *negative* viewer fee, that is, the seller might want to pay consumers to watch their channel!

The idea of a negative price may seem a bit outlandish, but there are cases when it actually takes place. For example, many credit card issuers effectively compensate users for making purchases with their cards. Specifically, credit card users are paid in the form of cash-back offers, frequent-flyer miles, or some type of point system that is redeemable in the form of cash, products or services. Credit card companies don't do this out of charity. Rather, for each dollar that you spent on a purchase, the merchant pays a fee to the credit card issuer. If this fee is sufficiently high, then it pays for the credit card issuer to pay the credit card user: a negative price.

ZERO PRICES

In two-sided markets it may be optimal to set a negative price. However, many times — most times, one might argue — negative prices are difficult to implement. In this context, zero pricing may be the optimal strategy. There are many situations when goods and services are offered for free. For example, radio broadcasts are typically offered for free, satellite radio being an exception. For many decades, television broadcasts were also offered for free. If you ride the New York subway you will likely find people distributing free newspapers (actually, this seems to have stopped now). A lot of online content, including newspaper content, is **ungated**, that is, can be accessed at no cost.

A related reason why zero pricing may be optimal is that the optimal price would be close to zero and the transactions costs of charging a price are greater than the revenue collected. For example, suppose that the optimal price of the *AM New York Metro* newspaper is 5 cents per copy. It's nothing short of a miracle how newspaper hawkers manage to dole out so many copies to ultra-fast-walking subway riders. To do that and to physically collect a nickel from each of them would require more than a miracle. In any event, the costs from such an operation would likely not compensate the benefit.

(There is also a psychological effect of free. Laboratory experiments show that consumers are disproportionately more likely to "purchase" a product when price drops from 1 cent to free than when it drops from 16 to 15 cents.)

Two-sided markets are by no means the only reason why we observe zero prices. In the next two sections, we will look at situations where zero pricing might be part of a dynamic strategy to get consumers "on board" a certain entertainment product, in the hope that, on a future date, those consumers may become a source of revenue. At some level, this has similar features to the waterbed effect or to two-sided markets. However, there are also specific features that justify their treatment in a separate section.

4.3. DYNAMIC AND SOCIAL EFFECTS

Section 4.1 considered the simplest case of optimal pricing, namely a seller selling one good at one point in time. In the previous section we considered the extension to a seller who offers multiple *related* goods: tickets on Wednesday and tickets on Thursday, or tickets and hot dogs; and, more generally, a seller with multiple related revenue streams (recording and touring revenues, or advertising and viewer subscriptions).

In this section, we extend the analysis in a different direction. Specifically, we consider the case when sales today are related to sales tomorrow, or the case when sales to person or group A are related to sales to person or group B. In some ways, the ideas are similar to those in the previous section, but in other ways they are different.

Consider first the case of a seller pricing a good with network effects. We say there are **network effects** when the benefit that a consumer gets from good X depends on how many other people also buy X. For example, it's no fun for me to play an online multiplayer game if I'm the only one. Imagine playing *Fortnite* or *League of Legends* by yourself. The good news is, you win. The bad news is, it's not fun.

Videogame sellers are well aware of these network effects and will do everything they can to make sure the chicken-and-egg problem is resolved in their favor. By chicken-and-egg I mean that there are games that are popular because they are popular: I want to play them because there are a lot of people who play it as well (and thus it's easy to find people to play with).

In this context — and to continue with the the same example — one possible strategy is to launch a new videogame as a freebie, heavily advertise it, and perhaps even seek the endorsement of some important lead player, so that there is a large initial set of players. Later on the seller may be able to monetize its **installed base** (i.e., the set of faithful game players), either by selling a premium version of the game or by selling game extras. This strategy is normally known as

The argument can be made that many people bought and read Harry Potter books because many other people were reading Harry Potter books.

Pexels lincense

introductory pricing: a low price — possibly zero — intended to build up an installed base of users.

Introductory pricing may make sense even for single-player videogames, and more generally for goods with no network effects. One feature of goods such as videogames is what economists refer to as **habituation effects**: Once you learn how to play a certain video game, your willingness to pay for it increases — including your willingness to pay for complementary goods or services that go with playing the game.

More generally, the digital revolution has brought with it a considerable expansion in the so-called **freemium** strategy: to give the good away for free as a means to attract users and up-selling some of them to a premium level. This practice has become very popular for a variety of digital products and services, including in particular software.

A situation similar to network effects is that of goods subject to **social effects**. I have a confession to make: the main reason why I bought and read a Harry Potter book is that, at the time, everyone else was doing the same. If I didn't know the rules of Quidditch, or who Snape and Dumbledore were, then I would not be able to have a conversation, I would have become a social outcast. OK, perhaps I exaggerate, but you get the idea: FOMO (fear of missing out) is a real phenomenon, particularly in the context of social networks, and it may have an effect on the demand for various products.

There is a famous quote by Yogi Berra regarding Ruggeri's, a St Louis restaurant: Berra allegedly stated that "nobody goes there anymore; it's too crowded." Seriously, it should be either "nobody wants to go there because it's always empty;" or "everybody wants to go there because it's always full." It doesn't sound as funny as the Yogi Berrism but it makes more sense. Specifically, **self-reinforcing dynamics** — to use the economics jargon — may lead to different outcomes from the same starting point. Adding to the list of economics jargon, people sometimes refer to information cascades, herding, viral dynamics, chicken-and-egg, etc. Basically, they all refer to the same phenomenon. (Incidentally, this type of phenomenon can also be found in the technology space. For example, the fax machine technology was known for many decades until adoption "exploded" during the late 1980s: all of sudden, everyone had to have a fax machine because everyone else had a fax machine.)

Figure 4.3 depicts this situation in terms of demand curves. The left panel corresponds to a standard, downward sloping demand curve. The right panel, in turn, corresponds to a multi-valued demand curve. For some price values — say, p' — two things may happen: Either the good is not "in", in which case demand is given by $d_1(p')$; or the good is "in", in which case demand is given by $d_2(p')$. In economics jargon, we say this is a case of **fulfilled-expectations equilibria**. Lynne Truss, author of the best-seller *Eats Shoots and Leaves*, pithily stated that her book "sold well because lots of people bought it." Yogi Berra could not have put it better! Basically, Truss modestly admits that her book got to the D_2 branch of the demand curve but it could as well have stayed in the D_1 branch.

Again, the trick for a seller is to make sure its demand falls on the D_2 branch, not on the D_1 branch. How do you do it? Easier said than done. The optimal strategy depends on the industry and product in question, as well as on the nature of the social effect. If you published a book, you might go on a book tour or, better still, be featured on Oprah. In music, being played on the radio may be the difference between point A and point B. Or, if you are lucky enough to be included in someone important's playlist, this may be the beginning of a favorable viral effect. Generally speaking, if you are mentioned by an industry expert or a famous YouTuber or some other influencer, then chances are your book or your song or whatever you created is now "in". Going back to book publishing, Box 4.1 focuses on a particular strategy: to buy scores of copies of your own book so that the book makes it to a best-seller list so that people than buy lots more copies. It's not as unusual as you might think.

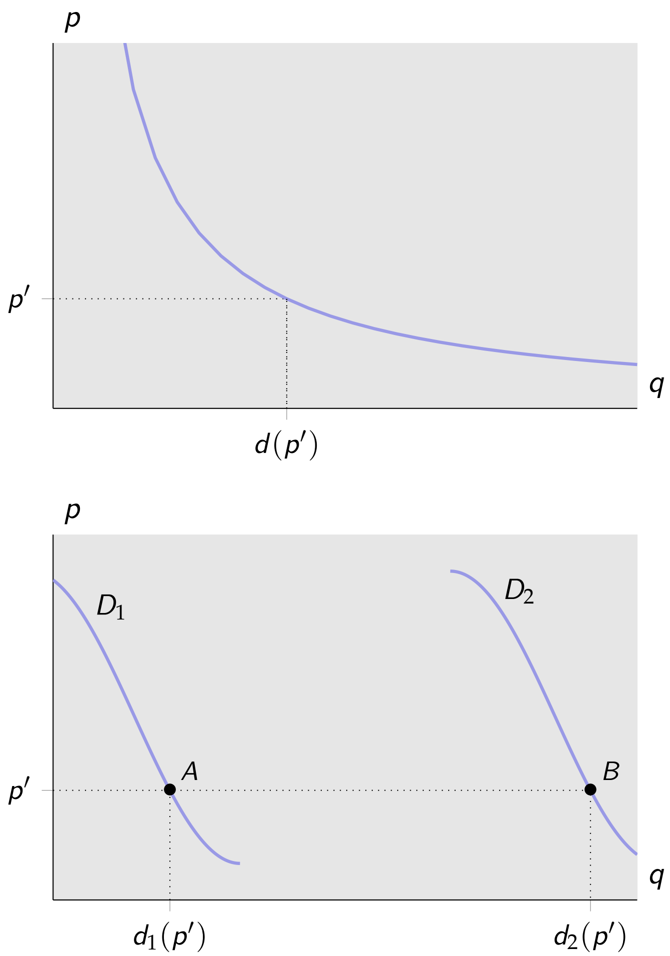


FIGURE 4.3
Good not subject to social effects (top panel) and good subject to social effects (bottom panel)

PRICING WITH CAPACITY CONSTRAINTS

To conclude this section, let us look at the pricing implications of goods subject to social effects, in particular when the seller is subject to capacity constraints. One of the more puzzling aspects of pricing entertainment goods is the practice of demand rationing: Boston Red Sox fans will wait in line for hours if not days in order to buy a season ticket; *Star Wars* fans were known to camp out on the sidewalk in anticipation of opening night; tickets for a U2 concert will run out in minutes if not seconds after they go on sale; and so fourth.

An economist's first reaction to situations of excess demand is

Box 4.1: Buying a spot on the best-sellers' list (source)

Authors, agents and publishers work hard to promote their books, hoping to get a spot on the *New York Times* best-seller list. One possible strategy is simply to purchase copies in bulk. Considering the benefits from making it to the list (cf Chapter 3), it may be a worthwhile investment.

And it's not a rare event. In 1995, Michael Treacy and Fred Wiersema arranged large purchases of their own book, *The Discipline of Market Leaders.* A consulting company hired by the authors paid $250,000 to buy 10,000 copies of their book and funneled 30,000 to 40,000 additional purchases through corporate clients, all made at stores.

In 2000, Alan Nevins, a prominent literary agent, sent orders worth more than $75,000 to four Los Angeles bookstores for thousands of copies of the new book, *Hope From My Heart: 10 Lessons for Life,* written by his client Rich DeVos, the 74-year-old co-founder of the marketing giant Amway.

More recently, in 2013, Soren Kaplan, a business consultant and speaker, hired ResultSource to promote his book, *Leapfrogging.* He was told that, in order to hit the *Wall Street Journal* best-seller list, he would need to order 3,000 books (about $70k). If order to hit the *New York Times* list, he would need to order 9,000 books.

The exact process by which the *Times* creates its best seller list is a secret. However, it is known that the newspaper polls a variety of bookstores precisely to avoid "strategic" moves such as ResultScourse's. Moreover, if the newspaper finds out that there were bulk orders then the book is listed with a dagger icon to denote that fact.

However, in this cat-and-mouse game, consultants like ResultScourse try to break their bulk sales into smaller sized or even individual purchases, effectively challenging the *Times*' mechanism to prevent "purchased" best-seller spots.

that price should be increased. In fact, this is one of the main tenets of the competitive market mechanism: when there is excess demand, then price tends to increase, whereas when there is excess supply price tends to decrease. But the above examples are clearly not examples of competitive markets. (A competitive market brings together

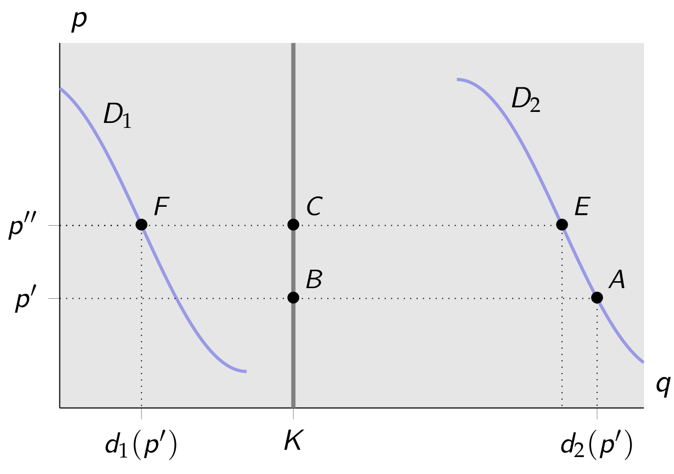


FIGURE 4.4
Limited capacity and pricing

many buyers and many sellers trading a homogeneous product, that is, a commodity.)

Why don't the Boston Red Sox simply increase their season ticket price? Why don't movie theaters opening the next episode of *Star Wars* do the same? Figure 4.4 illustrates the situation and suggests a possible answer. As usual, we represent price on the vertical axis and quantity (e.g., number of tickets sold) on the horizontal axis. In the present case we also measure capacity on the horizontal axis. Specifically, let K be the seller's capacity, that is, the seller is unable to sell more than K, that is, $q \leq K$.

Suppose that initially $p = p'$ and that, due to an effective "buzz" campaign, product demand is given by $d_2(p')$ (point A in the figure). Since the seller is capacity constrained, the most that it can sell is K (point B).

Since there is considerable excess demand, the seller may be tempted to increase price from p' to p''. Assuming that we continue with demand in the D_2 branch, this implies a movement along the demand curve from A to E. If this is so, then there is still considerable excess demand at p'', which implies that the quantity effectively sold remains at K (point C).

Since sales remain the same ($q'' = q' = K$), and price is higher ($p'' > p'$), the price increase strategy looks very good indeed. However, the risk of the price increase is that it may "rock the boat" in several ways. In particular, some fans may run to social networks and write angry posts about the price increase: "It's not fair" or something like that. Once that gets under way, it is not unlikely that a "fan revolt" takes place and that all of the sudden the artist in question is no longer "in", that is, demand reverts to D_1. This implies that, by increasing price from p' to p'' we have not just a movement along the demand curve but also a shift in the demand curve itself, from D_2 to D_1. In the end, the seller gets a demand of $d_1(p'')$ at a price p'' (point F), which is worse than the initial point, B.

As the popular saying goes, if it ain't broke don't fix it!

There may be other reasons why, in the entertainment world, demand rationing is a good idea. For example there is a certain buzz when a band sells out a concert minutes after tickets go on sale. They must be good, people will think. Second, by imposing a cost on fans who are desperate to get a ticket, sports teams effectively select diehard fans when they ration season tickets. (However, a millionaire might pay someone else to stay in line for hours, in which case rationing might have the opposite effect.) Finally, demand rationing might be a question of fairness, an idea we consider in the next section.

4.4. FAIR PRICING

One reason why capacity constrained entertainment sellers prefer not to increase price is the fear of being perceived as "unfair" or too "greedy". In this section we explore this dimension of pricing strategies. We do so by examining two specific examples: singer Neil Diamond and soccer club Liverpool F.C.

NEIL DIAMOND AT MSG

It was 2009 and Neil Diamond was scheduled to give a concert at New York's Madison Square Garden. Less than a minute after tickets went on sale, all tickets sold out. Soon after, more than 100 seats became available at TicketExchange.com for hundreds of dollars more than their face value.

In each case, the seller presented him or herself as a fan who had purchased the ticket but was unable to attend the concert. The ticket

Soon after tickets for a Neil Diamond concert went on sale they were sold out, and soon after that dozens of seats became available in secondary markets for considerably more than their face value. Allegedly sold by "fans", the tickets were actually sold by Neil Diamond himself.

listings were set at unusual prices, such as $1,164.01, which made them look more authentic. But you have to ask: what are the odds that you purchased a ticket a few minutes ago and now are unable to attend the concert? And what are the odds this happened to dozens of fans?

As it happened, the "fan" who was selling tickets in the secondary market was Neil Diamond himself.

If it's any consolation to Neil Diamond fans, the practice of selling premium-priced tickets on TicketExchange — priced and presented as resales by fans — is common among many other top performers. According to industry experts, the list includes artists such as Bon Jovi, Celine Dion, Van Halen, Billy Joel, Elton John and Britney Spears.

In the entertainment world, it is dangerous to increase prices when there is excess demand. Artists risk being seen by fans as "greedy", which in turn may result in a loss of fan support. In this context, secondary ticket sales are viewed by Ticketmaster, concert promoters and artists as a significant sources of revenue.

Many people complain that artists sell their own tickets for incredibly high prices but rarely admit doing so, thus avoiding the appearance of gouging fans. Paul McCann, a broker near Baltimore, says that "it's not fair for artists to hide behind Ticketmaster-TicketExchange." If fans do not find out, it's the best of both worlds for artists: you keep the image of "fairness", of someone who cares for fans and sets low nominal prices, and at the same time you manage to sell tickets for a higher price.

At minute 77 during the February 6, 2016 game against Sunderland, thousands of Liverpool FC fans walked out in protest against the proposed ticket price hike (from 59 to 77 pounds).

THE LIVERPOOL FC 2016 WALKOUT

By 2010, Liverpool FC were in dire straits. The owners were under pressure to repay loans owed to the Royal Bank of Scotland. Failure to do so could result in a nine-point deduction in the English Premier League (EPL), which in turn could lead to the club's relegation from the EPL.

Then American investors stepped in and offered to take over the club. The owners resisted — there were even a few lawsuits — but eventually the Fenway Sports Group's £300 million offer was accepted. (At the time, the bidder was presented as New England Sports Ventures.) Liverpool Football Club is now owned by the same parent company as the Boston Red Sox baseball team. Ian Ayre, one of the board members supporting the Fenway takeover, was appointed the new CEO. In the meantime, the Liverpool Supporters' Committee (LSC) was created to represent the major constituencies among the club's supporter base.

In February 2016, as construction on a £114 million new Main Stand continued, Liverpool announced a sharp change in ticketing for the next season. The announcement came after a period of thirteen months during which a specially appointed committee considered a variety of issues, including accessibility and affordability. The new plan made sure local fans would have priority access to over 20,000 tickets across the Premier League season. Moreover, prices as low as £9 provided easier access to local children.

However, what stood out in the eyes of many fans was the announced price of some seats in the main stands, a price hike from £59 to £77, as well as season tickets of more than £1,000. The LSC quickly came out protesting the changes, calling them "morally unjustifiable":

> We believe it is right and fair to lower ticket prices in order to sustain our support and subsequently the atmosphere inside Anfield [Liverpool's stadium]. Unfortunately, the decisions of the ownership are based purely on economics with no compromise.

Ian Ayre, the club's CEO stood his ground:

> We always carefully consider ticket pricing to ensure the long-term sustainability and competitiveness of the club while listening to the views of our matchgoing fans to understand the priorities around accessibility and affordability.

Then came the famous "walkout". On February 6, 2016, Liverpool were playing Sunderland at home. Roberto Firmino headed the home team in front and later provided an assist for Adam Lallana to put Liverpool up by 2–0.

Everything was going well on the field, but toward the end of the second half the shouts from the crowd switched from supporting the home team to attacking the home team management:

> Enough is enough, you greedy bastards, enough is enough!

The rest of Anfield, including the Sunderland supporters, echoed the same feeling with a roaring applause. Then, at the 77th minute, a large contingent of fans holding black flags walked out of the stadium. The choice of the 77th minute was a clear reference to the ticket price hike (from £59 to £77). Another symbolic move was that fans walked out while chanting "You'll Never Walk Alone," a song usually reserved for the last few moments of each game.

With about one quarter of the supporters left, the walkout produced a strong effect. It may or may not be related, but Liverpool's lead, which seemed sure to result in the 3 points for a win, was cancelled by two Sunderland goals (Adam Johnson's curling free-kick and Jermain Defoe 89th minute finish from 10 yards). Liverpool manager Jürgen Klopp, like so many other managers, often remarks that fans are very much part of the power of Liverpool, and the opposite of that (lack of fan support) seemed to play out during that end of game. As a fan later put it

> The dissenting mass totally sucked the life out of their team, Sunderland seized the moment, and the win was stolen.

The pressure on management continued. A fan' complained in an open letter:

> The point we were making was this: we're worth more. ... I've been giving Liverpool my money as a season-ticket holder for 30 years. They've got me. Like all of us, a slave to football's rhythm. We'd just like to feel they weren't laughing at us. ... 77 and out. So many of us gone — hopefully not for ever. But enough is enough.

As of the 2019-2020 season, main stand regular prices are £59. Season ticket prices vary from £685–869, well below the £1,029 announced in February 2016. The fans prevailed.

KEY CONCEPTS

- waterbed effect
- two-sided markets
- ungated
- network effects
- installed base
- introductory pricing
- habituation effects
- freemium
- social effects
- self-reinforcing dynamics
- fulfilled-expectations equilibria

REVIEW AND PRACTICE PROBLEMS

■ **4.1. Pricing.** You are the product manager for a media and entertainment company, in charge of a specific product. What factors should you take into account when setting the price of the product you manage?

■ **4.2. Margin and markup.** Define product margin and product markup.

■ **4.3. Elasticity rules.** If a seller sets its price according to the optimal pricing formulas derived in Section 4.1, what is the relation between demand elasticity and margin?

■ **4.4. Cable service.** Suppose the elasticity for cable service is constant and equal to -3.5. Suppose also that the cost of serving one additional customer is \$20 per month.

(a) Determine the profit maximizing price.

(b) Determine the values of margin and markup at this optimal price.

(c) Show that the equations relating margin, markup and elasticity hold when price is equal to the optimal price.

■ **4.5. Zero variable cost.** What is the optimal price of a good with zero variable cost.

■ **4.6. Ruth's concert pricing.** Ruth must decide what price to set for a summer music festival she's organizing. The festival takes place outdoors and there are no capacity restrictions. Moreover, variable costs are essentially zero (that is, the cost of organizing the festival is the same regardless of how many people attend). Based on her experience from previous summers, Ruth estimates that, at a price of \$20, the price elasticity of demand is equal to -1.8. If Ruth wants to increase revenues from ticket sales, should she increase or decrease price (or keep it the same)? Suppose that Ruth also gets revenues from concessions and that concession sales are proportional to the number of people attending the festival. If Ruth wants to maximize total revenues (from ticket sales as well as from concessions) should she increase or decrease price (or keep it the same)? Justify your answers.

■ **4.7. Multiple revenue streams.** Suppose that a firm's revenues come from two different but related streams. What implications does this have in terms of optimal pricing.

■ **4.8. Waterbed effect.** What do we mean by the waterbed effect? Exemplify.

■ **4.9. Two-sided market.** What is a two-sided market? What are some of the characteristics of pricing in a two-sided market? Exemplify.

■ **4.10. Zero pricing.** What are so many goods priced at zero? (Hint: there may be more than on reason.) Exemplify.

■ **4.11. Network effects.** What do we mean by network effects? What implications do these have for optimal pricing? Exemplify.

■ **4.12. Fulfilled-expectations equilibrium.** What do we mean by a fulfilled-expectations equilibrium? Exemplify.

■ **4.13. Excess demand.** If a seller observes excess demand at a given price, then the seller should increase price. Explain why this is or is not true. Provide specific examples.

■ **4.14. Pricing entertainment goods.** In what ways does the pricing of entertainment goods differ from other goods?

CHAPTER 5

MARKET SEGMENTATION

In the previous chapter we considered optimal pricing strategies. We started in Section 4.1 with basic optimal pricing rules and then considered several extensions. For all the variety of cases considered, they all share the feature that, at a given moment in time, all consumers pay the same price for the same good.

In many instances, however, different consumers pay different prices for the same good, or at least for approximately the same good. In this chapter, we consider the strategy of **market segmentation**, whereby a seller divides its customers into several groups, offering different selling deals to each segment.

Figure 5.1 provides the motivation for this chapter: if you set a single price, say p^*, so all of your consumers, then you are likely leaving money on the table. Specifically, your revenue is given by $p^*\, q^*$, the area labeled “revenue”. In this situation, you are missing out on two possible additional revenue sources: First, area A, which represents revenue lost to buyers who are willing to pay more than p^* but who only pay p^*. Second, area B, which represents revenue lost due to consumers who do not make a purchase when price is p^* even though they are willing to pay a positive price for a good that has zero variable cost.

The purpose of this chapter is to analyze seller strategies with the goal of capturing some of the revenue lost (to the seller) corresponding to areas A and B. The idea is that p^* is only optimal if the seller is constrained to setting the same price for all buyers. If the seller can

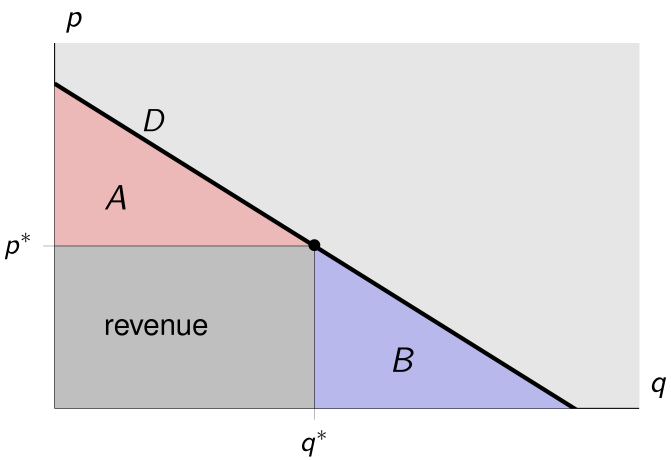


FIGURE 5.1
Motivation for market segmentation

change different prices from different buyers, then you can see how the seller might be able to do better than p^* for everyone.

We begin the chapter with the analysis of perfect price discrimination, the extreme case when the seller is able to set a different price for each buyer. This is a somewhat extreme and unrealistic case, but it helps understand the effects of market segmentation, as well as the limitations it might face.

The chapter then includes two important sections, corresponding to the two main approaches to market segmentation: segmentation by indicators and segmentation by self-selection. We conclude with a series of segmentation examples from media and entertainment industries.

PERFECT PRICE DISCRIMINATION

Consider the case of Jane, a personal trainer in LA. Jane works with clients from a variety of social and economics backgrounds, from middle class professionals to some very wealthy individuals, including some Hollywood celebrities. Jane does not have a fixed rate. In her mind, she has a lowest rate she is willing to accept. Beyond that, she charges an amount she thinks "is fair" and her clients can afford.

> Tomorrow, for example, I have Susan W. in the morning. She works as a middle manager in a mid-sized company. I charge her $80, my lowest rate. I don't think she can af-

cc0-icon CC0 Public Domain

In many markets, such as personal trainers, the seller has the ability to set a different price for each customer.

> ford more than that. Then in the afternoon I will be working at the home of a very well-known TV actor. I just read recently she is paid more than $200k per episode on this TV show. When she first asked me how much I charge I said $400 per session — that's my "celebrity" rate. She didn't blink, so that was that.

Situations like that of Jane the personal trainer are not uncommon, especially in service industries. Additional examples include include plumbers, lawyers, piano teachers, and many more.

Perfect price discrimination refers to the extreme case when each customer is charged a different price — exactly their willingness to pay. This is an extreme situation, one that is rarely if ever observed: in practice, sellers do not know exactly how much a specific customer is willing to pay. However, perfect price discrimination serves as a reference point to study the effects of market segmentation: it's as if each customer is him or herself a market segment.

Customer markets are a particularly important instance of (nearly) perfect price discrimination, especially in a business-to-business (B2B) context. In these markets, the number of customers is relatively small and the seller has considerable information about buyers. Examples include ready-mixed concrete, large commercial aircraft, enterprise software, tug boat push services. In these markets, although there is a list price (rack rate), each customer receives a discount (often negotiated); and the final price depends on the customer's ability to pay, their bargaining power, and possibly other factors.

Consider Haddad's, the leading film and television equipment

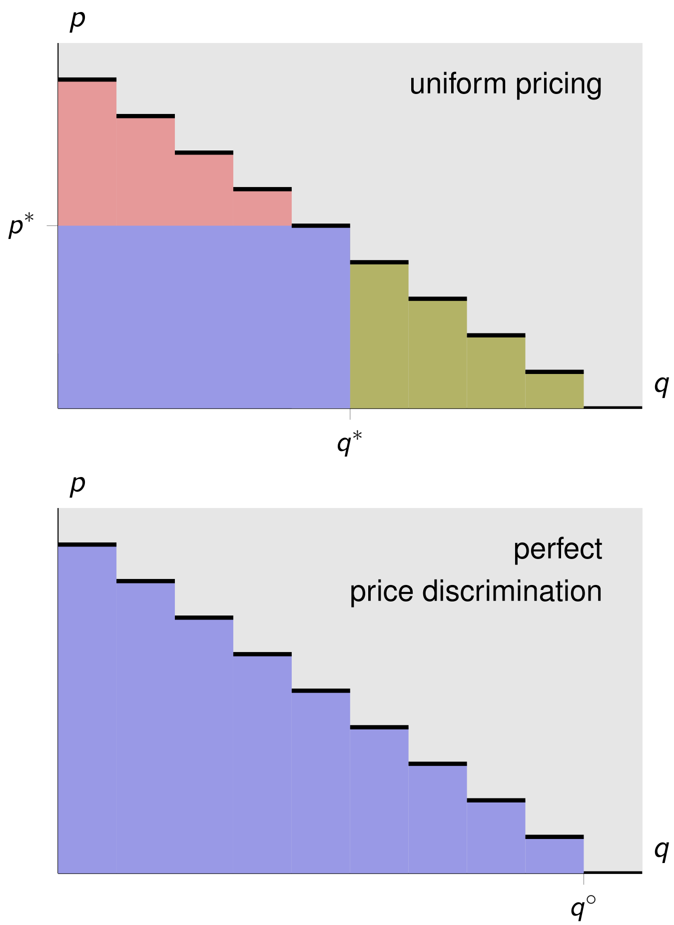


FIGURE 5.2
Perfect price discrimination

rental company in the US. Their rental rates vary according to a multiple factors. For example, the daily rental of a star trailer may vary from $550 to $700 a day. I suspect — though I have no hard information in this regard — that different film production companies will be offered different deals depending on their budget size.

Special mention should be made to **personalized pricing**. In the Internet era, sellers are able to obtain very detailed information about customers. By means of **cookies**, firms are able to learn very detailed customer histories (who clicked where and when, who bought what from whom). The large data sets thus created allow for **big data** analytics which provide a fairly accurate estimate of an individual consumer's willingness to pay. This, in turn, allows for nearly perfect price discrimination.

EFFECTS OF PRICE DISCRIMINATION

Figure 5.2 illustrates the difference between uniform pricing — the case we considered in Section 4.1 — and perfect price discrimination. In order to better understand the effects, we consider the case when demand is given by a finite number of buyers, nine in this case. Buyers are placed in decreasing order of willingness to pay, leading to a downward "sloping" ladder of willingness to pay — the demand curve.

Under uniform pricing — the top panel in Figure 5.2 — the seller sets price p^* for all buyers. At this price, a total q^* buyers make a purchase. These buyers would be willing to pay more than p^* (with the exception of the very last buyer, who pays exactly her willingness to pay). As far as these buyers are concerned, the seller is leaving money on the table, specifically the area in red in the top panel of Figure 5.1, specifically, what the seller is missing by not charging high-valuation buyers as much as they would be willing to pay.

The four lowest valuation buyers, that is, the buyers whose valuation is lower than p^*, refrain from making a purchase. However, their valuation is positive, and offering them the good in question would imply no additional cost to the seller. This corresponds to the green area in Figure 5.1, that is, the total value the seller is leaving on the table by not selling low-valuation buyers for as much as they would be willing to pay.

Under perfect price discrimination — the right panel in Figure 5.2 — the seller sets a different price p for each buyer, specifically the value of their willingness to pay. This means that all buyers with a positive willingness to pay make a purchase (nine of them).

So, with respect to uniform pricing, we first conclude that high-valuation buyers are worse off: before they were paying p^*, now they are each paying their willingness to pay, which is higher. Second, we see that low-valuation buyers are better off. Strictly speaking, they are just as well off as under uniform pricing: in one case, they do not make a purchase, in the other case they do but they pay a price equal to their willingness to pay. However, we must remember that perfect price discrimination is a limit case. In most real-world situations, sellers do not have perfect knowledge of the buyer's willingness to pay, and so some buyers who were not making a purchase under uniform pricing now do make a purchase under price discrimination, and pay a price that is lower than their willingness to pay.

Finally, we observe that perfect price discrimination implies a greater number of buyers make a purchase and that seller revenues increase. In the specific case depicted in Figure 5.2 seller revenues approximately double as we move from uniform pricing to perfect price discrimination, which suggests the potential gains from market segmentation are significant.

LIMITATIONS OF MARKET SEGMENTATION

If sellers seem to gain from price discrimination, why don't we observe market segmentation (a.k.a. price discrimination) more often? Partly because, as we will see next, there are several practical limitations to market segmentation.

The first impediment to implementing a market segmentation strategy is the difficulty of **market research**. If you do not have good information about buyers (how many there are, what their willingness to pay is, how sensitive demand is to price changes, and so on), then your market segmentation strategy is destined to failure. As the popular saying goes, garbage in, garbage out.

Market research requires first of all obtaining historical data. This may be purchase data from stores, click data from online sites, viewership data for various media products, etc. This data can be generated by the seller or purchased from third parties such as Nielsen. In addition to historical data, a seller may create their own data by setting up focus groups, perhaps experimenting in selected markets, and a host of other practices which allow the seller to get a better understanding of the values of willingness to pay and how these relate to various demographics (e.g., age, gender, income, and so on).

In sum, if sellers do not implement segmentation strategies, or if they do so unsuccessfully, the explanation often lies with insufficient information regarding demand. I will not be dwelling too much on the issue of market research. By no means should this be understood as implying that market research is not important. On the contrary, it's an absolutely crucial step in any pricing and market segmentation strategy. However, it does not correspond to the core of economic analysis, which is what this book is about.

A second important impediment to a successful segmentation strategy is **arbitrage** (related terms include **resale**, **gray markets**,

harvesting). If you set different prices to different market segments, then there is an incentive to buy in the cheap market and sell in the one where price is high. And if this happens, then the seller's strategy may completely backfire.

One example where arbitrage plays an important role is textbooks. Many publishers sell international student editions targeted at developing countries, where students have lower ability to pay. In at least one instance, a group of entrepreneurial students decided to reimport these cheaper editions into the US and sell them on campus at a price substantially lower than the hardcover US editions.

A second example of arbitrage is given by Levis, which was setting very different prices in Bulgaria and the UK. An entrepreneur decided to purchase jeans in Bulgaria at retail price and then sell them wholesale to Tesco, a UK chain. The price difference set by Levis was so big that the Bulgaria retail price was lower than the UK wholesale price.

A third example of arbitrage is given by the practice known as harvesting. Savvy buyers looking for digital camera storage would get a MuVo2 MP3 player, take out the drive from the player, and then use it for camera storage. This would translate into substantial savings with respect to buying camera storage outright. (In response, MuVo changed its microdrive so it no longer conformed the CompactFlash standard: it could be recognized by digital cameras.)

For all the variety of these three examples, they do have one thing in common: they all refer to actual physical products, in other words, stuff. It's difficult if not impossible to arbitrage services. You cannot buy a men's haircut in Denver (average price $32) and sell it in San Francisco (average price $49). For this reason, the threat of arbitrage is typically less of a problem in services than in physical goods.

This is good news for sellers in the media and entertainment space, where what's being sold is typically a service, not a product. That said, here too sellers must beware of this possibility. For example, by setting a virtual private network (VPN) I may be able to go online as if my address were in a country different from my home country. In this context, as of 2019 HBO Now was selling for $8.69 in Sweden and $14.99 in the US (cf Figure 5.3). It's the same *Game of Thrones* regardless of whether it's streamed from Sweden or from the US. You can see the incentive for viewers to set up a VPN and save about $5 a month by pretending to be Sweden based. (I do not

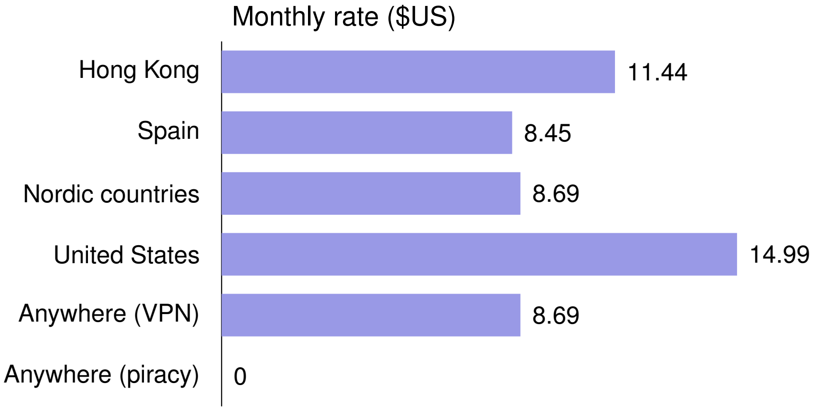


FIGURE 5.3
HBO pricing

include Spain in this calculation because HBO's content in Spain is dubbed into Spanish.)

A recent study compared Netflix rates across countries. It concluded that

> It's clear Colombia is the best value place to subscribe to Netflix. So if you want access to a good number of titles and at an affordable rate, a move to Colombia (or a good VPN) may be in the cards.

In fact, in 2019 Netflix's standard plan cost $17.29 in Switzerland but only $7.21 in Colombia. Moving to Colombia may not be easy for most viewers, but setting up a VPN may not be a bad idea. (Keep in mind, however, that some content may be dubbed into Spanish. Maybe arbitrage is not a big problem for Netflix after all.)

Finally, unlike market research and arbitrage, one thing that does *not* seem to impose any significant restrictions on market segmentation is the law: both the US and the EU are fairly laissez-faire when it comes to sellers setting different prices for apparently the same product or service, even if, strictly speaking, there are laws against the practice.

NAME YOUR PRICE

In all of our analysis so far we've considered the case when sellers set the price. There are, however, alternative ways of organizing a sale.

If, instead of the seller, prices are set by buyers, then we enter the world of auctions. If, by contrast, price offers are made by both seller and buyer then we are in the realm of negotiations. Within auctions we can distinguish between different types of auction, including the English auction, the second-price (a.k.a. Vickrey) auction, the Dutch auction, and so on. (Chapter 6 *Introduction to Industrial Organization* — arguably one of the two best textbooks ever written — includes more details.)

In the world of media and entertainment and at the consumer level, most sales are done at prices set by the seller. However, auctions and negotiations play an important role in some cases. For example, the rights to broadcast the Olympic games are frequently auctioned off; and the compensation paid to movie actors is frequently negotiated.

Finally, while the rule in the media and entertainment space is for prices to be set by sellers, there are exceptions. One such exception is the practice of **pay what you like** (PWYL), which is exactly what the name suggests. To some extent, this system is similar to an auction: one expects that, as in an auction, each buyer pays a price that is greater the greater the buyer's willingness to pay.

Does PWYL really work? Why or why not? Are entertainment goods more or less likely candidates for PWYL? Why or why not? These are some big questions for which we still do not have complete answers. Box 5.1 reports on a particular experiment of PWYL: Radiohead's launch of *In Rainbows*.

FAIRNESS

In the previous chapter we touched on the issue of fair pricing. The relation between an artist and her fans is more than the relation between a seller and a buyer. Setting too high prices may come across as being too greedy, which in turn may turn fans away. A similar dilemma arises in the context of market segmentation: artists must be very careful about the fans' reaction to it.

Box 5.2 discusses the case of Kid Rock, an example of the balance between the revenue-increasing goal of market segmentation and the fairness concern that artists have about their fans.

Box 5.1: Radiohead's *In Rainbows* experiment.

In 2007 Radiohead released their seventh full-length album, *In Rainbows*. The novelty: by October 10, the album was available as a digital download from the band's website and fans were encouraged to "pay what you wish." Gene Simmons of Kiss was one of the many critics of Radiohead's move:

> I open a store and say 'Come on in and pay whatever you want.' Are you on f***ing crack? Do you really believe that's a business model that works?

Fortune magazine went further, listing the *In Rainbows* experiment in its article, "101 Dumbest Moments in Business." In fact, Many fans downloaded the album without paying anything. Moreover, there was a spike in piracy, with more than 400,000 copies swapped in one week, 2 million by the end of the month.

But there were also a lot of fans who did pay for the download. According to Radiohead's publisher, *In Rainbows* made more money before the album was physically released than the total sales of the band's previous album, *Hail to the Thief*. As the MNE Blog aptly put it one decade after the experiment,

> 'In Rainbows' absolutely didn't kill the idea that music should be paid for. What it did do, though, was show that the idea of setting a single, one-size-fits-all price for an album was long overdue a rethink. Not just because a lot of people wanted to pay less or nothing, but because plenty of fans wanted to pay more.

In other words, to a great extent the *In Rainbows* experiment was a strategy of market segmentation.

5.1. SEGMENTATION BY INDICATORS

Aside from perfect price discrimination, there are essentially two types of market segmentation: segmentation by indicators and segmentation by self-selection. Segmentation by indicators refers to the case when the seller can observe the buyer's type, that is, what market segment they belong to. For example, HBO knows that willingness to pay for HBO Now is on average higher for Americans than

Box 5.2: Kid Rock.

Like many other artists, singer Kid Rock is concerned with fans having to pay high prices to see him in concert.

> I don't want to break you by coming to see me. And as somebody who sings songs about working-class people and for them, I thought it was important to stand by, you know, by what I sing about, what I preach.

The problem is that, by selling inexpensive tickets, you are basically feeding scalpers. Jared Smith, President of Ticketmaster North America, guarantees that

> If the Rolling Stones sold a bunch of $10 tickets to their concerts, you know exactly what would happen. The scalpers would come in.

In order to address this problem, Kid Rock devised a multi-pronged market segmentation strategy. First, most tickets are sold at a low $20. Second, Kid Rock concerts are based on paperless ticketing: fans must show their driver's license to have access to the better seats. Third, the desirable front-row seats — platinum tickets — sell for "whatever the market dictates." Fourth, in order to avoid a sense of being pushed out by rich patrons who can pay exorbitant prices for a front-row seat, Kid Rock reserves some of the platinum tickets to be given away by lottery to fans who purchased one of the cheaper $20 tickets.

In some ways, this is the best of both worlds: Kid Rock gets to sell platinum tickets at platinum prices, fans have at least the hope of being near the stage having paid a mere $20 — and equally important, scalpers are kept at bay.

for Swedes. As a result, they set a higher price in the US market than in the Swedish market (cf Figure 5.3). In this case, the indicator is place of residence: you need to show you live in Sweden in order to purchase HBO Now at the Sweden price.

There are many other potential indicators on which to base market segmentation: age (e.g., senior discount), student status, gender (e.g., auto insurance rates are different for men and women), and so forth. Examples of market segmentation by indicators in the entertainment world include Cinemex (poor quarters vs rich quarters of Mexico City); The New York Mets (playing against the Yankees vs playing against the Nationals); a rock band (playing at a state university campus vs playing at a party paid by a Silicon Valley billionair). Can you think of other examples?

In terms of optimal pricing, the trick is to apply the elasticity rule to each market segment — assuming that you have estimated each segment's demand curve. Consider first the case when production cost is positive. We should then set higher prices in less elastic markets, following the *elasticity rule* presented in Section 4.1:

$$\frac{p_i - c}{p_i} = \frac{1}{-\epsilon_i}$$

where ϵ_i is price elasticity in market segment i.

As a numerical example, consider the pricing of *Harry Potter and the Philosopher's Stone*, paperback edition. Suppose the production cost is $1 per copy, the demand elasticity in the US is given by -1.2, and the demand elasticity in the UK is given by -1.4. What are the optimal prices? Applying the elasticity rule we get a price of $6 in the US and $3.5 in the UK. (As of January 2017, actual prices were $6.91 and $4.33, respectively.)

Consider now the case of pricing entertainment goods with zero variable cost. As we saw in Chapter 4, this implies that the price in each market segment i be set such that elasticity in that market is equal to 1 (in absolute value), that is, $|\epsilon_i| = 1$.

This situation is illustrated in Figure 5.4, where we consider two markets: a small size, small valuation market with demand D_L and a large size, large valuation market with demand D_H. The two top panels correspond to the case when the seller does *not* engage in market segmentation, that is, sets the same price in both markets. In this case, revenue in the small market equals 32, whereas revenue in the large market equals 224.

The two bottom panels correspond to the case when the seller does engage in a market segmentation strategy. As can be seen, the seller sets a lower price in the lower valuation market and a higher price in the higher valuation market. Why are these the optimal prices? As we saw in Section 4.1, when demand is linear — as is the case in the present example — the optimal price corresponds to

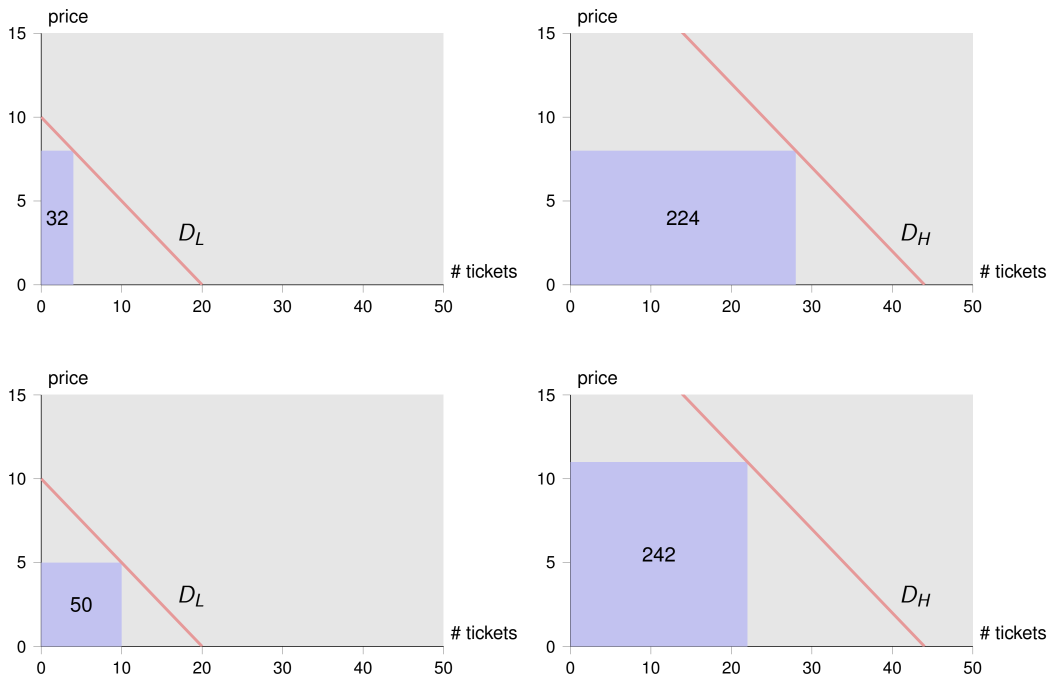


FIGURE 5.4
Market segmentation and profits

the midpoint in the demand curve between axis intercepts. This is the price level that optimally balances a high price and a high quantity demanded. And as can be seen in Figure 5.4, such midpoint is higher in the D_H market than it is in the D_L market segment.

Finally, in the market segmentation case revenues are given by 50 in the small market and 242 in the large market. In other words, the seller manages to increase revenues in both markets.

The point of this example is that, when the seller sets the same price in both markets, the best it can do is to achieve a sort of compromise: not to set a price that is too high for the low-valuation market nor a price that is too low for the high-valuation market. Allowing for market segmentation allows for the seller to move away from this constraint and thus increase revenues in both markets. In the present case, total revenues increase from $32 + 224 = 256$ to $50 + 242 = 292$, an increase of about 14%. Not quite doubling as in the example of perfect price discrimination considered earlier, but still pretty good.

Type	Price ($)	WTP(A)	WTP(B)
16 GB	199	250	300
32 GB	249	290	400
64 GB	299	330	500
128 GB	399	370	700

FIGURE 5.5
iPod versions

5.2. SEGMENTATION BY SELF SELECTION

In many cases the seller knows that there are various market segments and knows the demand characteristics of buyers in each segment, but is unable to identify a given buyer as belonging to one or the other segment. Consider, for example, an 128GB iPod. Suppose that Apple has been able to identify a market segment, call it segment A, of consumers willing to pay up to \$370; and a second market segment, call it segment B, of consumers willing to pay up to \$700. The problem is that Apple is unable to discriminate between types based on place of residence or any other indicator. Is there any hope for market segmentation?

Suppose that, instead of offering one iPod model, Apple offers different versions of the iPod, each containing a different amount of memory. Table 5.5 lists the four versions of iPod on sale, Apple's sale price (in dollars), and the buyer's willingness to pay (WTP) by segment (also in dollars). (In the bottom row we observe the values 370 and 700 mentioned in the previous paragraph.)

Now suppose that each buyer chooses the iPod version that maximizes the difference between willingness to pay (WTP) and price paid, that is, each buyer maximizes value for the money. If you are a type A buyer, you're better off by choosing a 16 GB iPod. At this price your buyer surplus is given by $250 - 199 = 51$, that is, \$51. For the remaining versions, a type A's surplus is given by \$41, \$31, and -\$29 (that is, a type A is willing to pay less for a 128 GB iPod that it costs).

By contrast, if you are a type B buyer, then your surplus form buying an iPod is \$101 for a 16 GB, \$151 for a 32 GB, \$201 for a 64 GB, and \$301 for a 128 GB. It follows that a type B's best choice (value for the money) is a 128 GB iPod.

The strategy of offering different versions of (essentially) the same product is known as — you guessed it — **versioning**. In the present case, versioning is done along the dimension of total memory, but there are many other ways in which it can be done. The versioning strategy is effectively a market segmentation strategy. In fact, type A consumers end up paying a price of $199 for an iPod, whereas type B consumers pay $700. Admittedly, they are not getting the same product, so this is not exactly the same as offering a senior discount for the exact same product. However, to the extent that the difference in production cost between a 16 GB iPod and a 128 GB iPod is not that high (it is not), then we may rightly consider the versioning strategy as one of market segmentation.

The beauty of the versioning strategy is that I do not need to directly observe whether a given consumer belongs to segment A or B. All I need to do is to offer a menu of versions and let consumers self-select into the version intended for their type. In other words, the trick is to offer versions such that a high-valuation buyer optimally chooses a high-priced version, whereas a low-valuation buyer optimally chooses a low-priced version.

Easier said than done. The trick for this to work is that prices and qualities be such that: (a) low-valuation consumers want to buy the low-value version (the so-called **participation constraint**); and (b) the price difference across versions is not so high that high-valuation consumers would prefer to switch to low-value version (the so-called **incentive constraint**). In practice, this is often based on years of experience of what the market will bear.

Continuing with the iPod example. Why does Apple's pricing do the trick? One important feature of buyer valuations is that buyers with higher valuations are also willing to pay more for extra memory. Specifically, a type B consumer is willing to pay an extra $400 for 128GB instead of 16GB, whereas a type A is only willing to pay an extra $120. For this reason, the seller is able to find prices such that type A goes for the low-memory version and type B for the high-memory one. Generally speaking, a self-selection mechanism relies on some correlation of buyer valuations. In this case, the crucial feature is the positive correlation between valuation for extra memory and overall valuation.

The art of versioning comes itself many different "versions". Con-

# days	price per day ($)
1	119.00
2	104.50
3	101.67
4	95.00
5	79.00
6	67.50
7	59.29
8	53.13
9	48.34
10	44.50

FIGURE 5.6
Disney theme parks (Orlando)
Note: Each bar corresponds to specific number of days.
One-day ticket prices vary by date, theme park, age.

sider the case of Disney Orlando theme parks. As Table 5.6 documents, the price per day at Orlando is lower the more days you stay at Disney. In other words, there is a quantity discount. Quantity discounts work as a self-selection market segmentation strategy if the increase in valuation from n to $n+1$ days is higher for high-valuation consumers than for low-valuation consumers, very much like memory in the iPod example.

For example, suppose that a low-valuation buyer is willing to pay at most $125 for a one-day ticket, $130 for a two-day ticket, etc, all the way to $170 for a 10-day ticket. For such a consumer the optimal choice is to buy a one-day ticket. For example a 10-day ticket costs $44.5 per day, for a total of $445, which is way more than the $170 valuation.

Now suppose that a high-valuation consumer is willing to pay $200 for a one-day ticket, $300 for a 2-day ticket, etc, all the way to $1,100 for a 10-day ticket. Such a consumer is better off buying a 10-day ticket. In fact, the surplus from buying a 10-day ticket is given by $1000-445=655$, whereas the surplus from buying a one-day ticket is $200-119=81$.


Kārlis Dambrāns

The iPhone 5C may be considered a case of a “damaged” good.

DAMAGED GOODS

A particularly important type of versioning is what is known as **damaged goods**: starting from a higher-valuation product or service, the seller purposely lowers its quality, frequently at no savings or very little savings — or even at an extra cost —, with the sole purpose of segmenting the market.

One example of damaged goods is the iPhone 5c. Although its production cost was lower than that of the regular iPhone 5, the price difference was considerably greater. Specifically, the iPhone 5c was priced considerably below the regular iPhone so as to attract lower-valuation buyers who could not afford the iPhone 5. Apple did not continue this strategy, which suggests it did not work so well. One possible reason is that the iPhone 5c’s price was so attractive that its launch cannibalized sales of the iPhone 5.

In the world of motion pictures, the movie release window system corresponds to a form of damaged goods. Over time a movie is gradually released, first in theaters, then in premium streaming services, then as a DVD, and eventually it is shown on cable TV.

Similar to movies, books are often first released in a hard cover version and then — maybe two years later — in a paperback version. Although printing a paperback is cheaper than printing a hard cover, the price difference is typically substantially greater than the cost difference.

		Willingness to Pay ($)	
Type	Count (000)	Print	Online
Baby boomers	30	50	0
Gen X	40	30	30
Millennials	35	0	50

FIGURE 5.7
Willingness to pay by consumer type

5.3. BUNDLING

Bundling refers to the practice of selling several goods — normally related goods — as a bundle. A distinction can be made between **pure bundling**, whereby buyers must purchase the bundle or nothing, and **mixed bundling**, whereby buyers are offered the choice between purchasing the bundle or one of the separate parts. In what follows, we consider each possibility separately.

MIXED BUNDLING

Consider a newspaper that is offered in a print and in a digital version. Suppose that there are three types of buyers: Baby Boomers, Gen X and Millennials. Suppose moreover that valuations (that is, willingness to pay) for the print and for the digital editions are as in Table 5.7. For example, there are 30 thousand baby boomers, each willing to pay up to $50 for the print version but nothing for the digital one. And so forth.

In this context, one possible selling strategy is to price each version (print, digital) at $50. This would lead to a revenue of $50 \times 30 + 50 \times 35 = 3250$, that is, $3.25 million. How did we get this value? Well, at a $50 price, only Baby Boomers (30 thousand) are willing to pay for the print version, and only Millennials (35 thousand) are willing to pay for the digital version. Gen X are not willing to pay $50 for either version.

A second possibility is to price each version (print, digital) at $30. This price would attract both Baby Boomers and Gen X to purchase the print version, and both Gen X and Millennials to purchase the digital version. Altogether, this would lead to a revenue of $30 \times (30 +$

$40) + 30 \times (40 + 35) = 4350$, that is, \$4.35 million. This is better than pricing at \$50, but as we will see next the seller can do even better.

Consider now a third possible strategy: to price each version (print and online) at \$50 *and* offer the bundle print+online for \$60. Baby Boomers don't care about the digital version. Therefore, they are not willing to pay an extra \$10 (\$60 as opposed to \$50) to get the bundle. In other words, Baby Boomers will opt to purchase the paper version only. By a similar reasoning, Millennials will opt to purchase the digital version only. Finally, Gen X will optimally purchase the print+online bundle for \$60. The seller's total revenues are now given by $30 \times 50 + 40 \times 60 + 35 \times 50 = 5650$, that is, \$5.65 million.

Magic! Introducing the bundle leads to an increase in revenues from \$4.35 million (the best alternative to bundling) to \$5.65 million.

What is going on here? Offering the bundle, together with the single options, effectively allows the seller to expand the set of options, which in turn allows the seller to better segment the market. Specifically, there is an intermediate demographic — Gen X — whose preferences are between print and digital. They are not willing to pay \$50 for print or \$50 for digital. They are, however, willing to pay \$60 for both. The older and younger demographic segments have very extreme preferences, so the introduction of the bundle does not affect their choice.

PURE BUNDLING

Consider now the case of pure bundling. For example, cable subscribers typically do not have the option of buying only the subset of channels in which they are interested. In other words, cable subscribers are not offered an a-la-carte option. It's all or nothing. It's pure bundling.

For the sake of illustration, suppose that Kabral Kable offers only two channels: K sports and K cartoons. Suppose moreover that there are two types of consumers: Type A consumers, a total of 800 thousand, are willing to pay \$10 for sports, \$3 for cartoons. Type B consumers, a total of 600 thousand, are willing to pay \$3 for sports, \$10 for cartoons. What is the optimal pricing policy?

Consider first the case when Kabral Kable sells channel subscriptions separately. If price per channel is \$3, then both types buy both


CC0 public domain photo

Cable TV is an example of pure bundling.

channels. This leads to a total revenue of 3 dollars times 2 channels times 1.4 million viewers (both types), a total of $8.4 million.

Alternatively, suppose that Kabral Kable sells channel subscriptions separately but charges $10 per channel. In this case Type A only buys the sports channel, whereas Type B only buys the cartoon channel. It follows that total revenues are given by $10 \times 800 + 10 \times 600 = 14000$, that is, $14 million. This is better than pricing at $3: you sell fewer channel subscriptions but sell them for a much higher price.

Now suppose that Kabral Kable offers a bundle of the two channels and no a-la-carte option. How much is each type willing to offer for the bundle? Type A consumers are willing to offer $3 + 10 = 13$, that is, $13 per channel. Similarly, Type B consumers are willing to pay $13. So, if Kabral Kable sells the bundle for $13 its total revenues are $13 \times (800 + 600) = 18200$, that is, $18.2 million. This is a substantial improvement with respect to a-la-carte pricing with individual channels priced at $10.

What is going on here? As we saw at the beginning of the chapter, uniform pricing leaves a lot of money on the table. One specific problem is given by area B in Figure 5.1. This corresponds to consumers who are willing to pay a positive price but less than the price the seller charges. By selling a bundle of both channels, the seller is able to effectively offer all channels to all consumers, thus capturing the area B that would be left on the table under uniform pricing.

Note that, if mixed bundling can be understood as a form of market segmentation by versioning, pure bundling has the opposite effect: it aggregates the various products so that, effectively, there is only one “segment”, in this case the segment of consumers who are willing to pay $13 for the bundle.

One of the notable novel features in the music and home video segments in the past decade or so has been the emergence of the subscription model. Examples include Pandora and Spotify for music, Netflix and Amazon Prime for video, and many others. The subscription model can be seen as a case of pure bundling: instead of paying for the rights to listen to a particular song, I pay a monthly subscription which gives me access to a whole host of content, similarly to what we find in a cable subscription. This is a major departure from the more traditional business model in music, where consumers bought individual recordings. That said, even then we observed some degree of bundling: when you buy a CD you effectively buy a bundle of songs. For a long time, you were unable to purchase individual songs.

In other words, for quite some time we lived in a world of pure bundling (CDs only), then we switched to mixed bundling (CDs or individual mp3 downloads), and now seemingly we're moving back to a sort of pure bundling (subscription service).

5.4. CASE STUDY: PRICING AT THE NY METS

> No strategy will fill more empty seats at Shea than for the team's new manager to guide the Mets to a winning season. By all means, let the quest to win continue. Meanwhile, the team and its fans will gain by having prices that track [demand] more closely.

All baseball games are not equal: It's not the same to play the Yankees on Sunday or the Royals on Wednesday — with no offense to either team or weekday. Such was the idea underlying the New York Mets' 2002 decision to switch from uniform to tiered pricing. "The more we studied it, the more it made sense to tailor pricing to match demand as much as possible," said David Howard, the Mets' senior vice president for business.Was tiered pricing good for the Mets' finances? Did it have any influence on fan loyalty? Is tiered pricing the future of baseball?

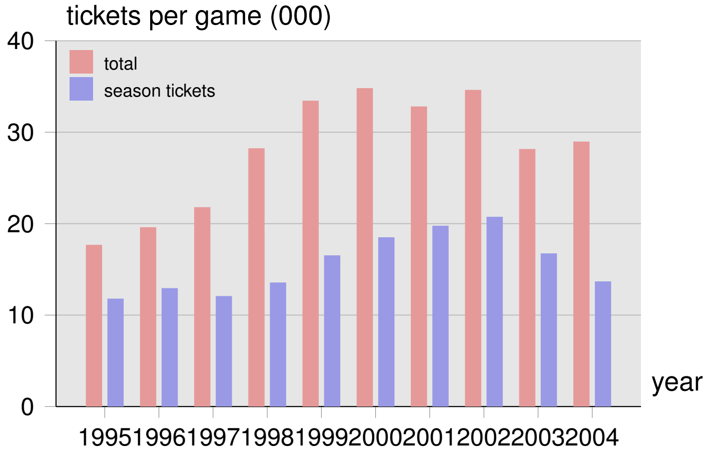


FIGURE 5.8
Ticket sales

THE NEW YORK METS

Originally slated to be one of eight teams in the Continental League (a new baseball league that never got off the ground), the New York Metropolitans ("Mets") instead joined the National League and played their first game on April 11, 1962. Throughout the subsequent decades, the team has only reached the post-season six times, winning World Series championships in 1969 and 1986. During the 1990s, the win-loss percentage hovered around 50%, more frequently below than above (see Figure 5.9).

The Mets have always played "second fiddle" to the New York Yankees dynasty across town. Even when the two teams shared the World Series spotlight in 2000, the Yankees eclipsed the Mets at the gate, outdrawing them by about 4,000 tickets a game.As Dave Howard puts it, "We have to be creative and sell tickets, while the Yankees are able to simply 'take orders' for season tickets."

TICKET OPTIONS

Mets fans can purchase tickets in several ways.

- Season tickets indicate the greatest level of commitment and entitle the owner to the same seat at each of the Mets' 81 home games.

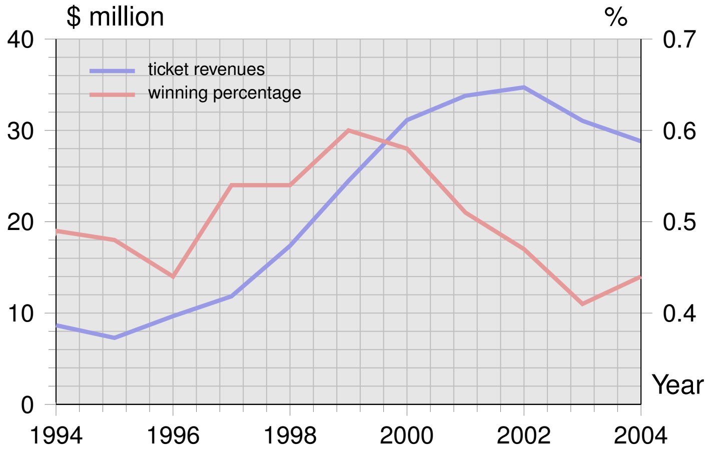


FIGURE 5.9
The New York Mets performance, on the field and at the ticket office: total revenues for ticket sales (left axis) and winning percentage (right axis).

- Fans who wish to see multiple games, but do not want to invest in the entire season, can choose from a variety of ticket plans enabling them to purchase weekend only or weeknight only combinations.
- The next level of commitment is to purchase one of seven different "Six-Packs" of predetermined games.
- Advance ticket sales for individual games begin in late February. The Shea Stadium Ticket Office usually sells as many as 100,000 seats on the first day of advance sales.
- Fans may continue to purchase tickets to individual games throughout the season.

For any of the five options, the most expensive tickets for any game would be those for central, close-to-the-field seats; the least expensive tickets would be for the far-away bleacher and upper-tier seats; and intermediate locations would command intermediate prices. (See Figures 5.10 and 5.11.) There are also a limited number of luxury suites available, primarily purchased by corporations for entertaining clients and employees.

FIGURE 5.10
Shea stadium seating

REVENUE SOURCES

Ticket sales are the main source of revenues during game day, but not the only one. Figure 5.12 shows the values of ticket revenues and concessions revenues from 1999-2004. As can be seen, ticket revenues correspond to about two thirds of total revenues.

Ticket revenues, in turn, can be divided into season tickets and other sales. Figure 5.9 shows how these evolved during the 1992–2004 period. It is interesting to note the relation between team performance and ticket sales, both season tickets and other sales. Notice in particular how success in year t translates into sales in year $t+1$.

VARIABLE PRICING

Notwithstanding the variations in price across different stadium seats, until 2002 each particular seat was priced the same regardless of what game was being played. In the Fall of 2002, the Mets announced their plan to switch to tiered pricing. The idea was not novel to baseball: The Cubs, Colorado, Cleveland and San Francisco all had experienced some variation of tiered pricing.But the Mets were probably the first team to implement a comprehensive tiered pricing strategy ("the New York Mets operate the most complex pricing structure in all of pro sports").

	2003–04				01–02
Seating Area	Gold	Silver	Bronze	Value	
HP Club Gold	215	205	195	185	195
HP Club Silver	161	153	145	137	145
Baseline Club	161	153	145	137	145
Metro Club Gold	76	70	64	58	64
Baseline Inner	76	70	64	58	64
Metro Club	72	66	60	54	60
Press Level Club	72	66	60	54	60
DVS	72	66	60	54	60
Hcap — HP Club	72	66	60	54	60
Inner Field	53	48	43	38	43
Inner Loge Box	53	48	43	38	43
Baseline Outer	53	48	43	38	43
Middle Field Box	46	42	38	34	38
Mezzanine Box	39	36	33	30	33
Outer Field	39	36	33	30	33
Outer Loge Box	39	36	33	30	33
Press Box	36	33	30	27	30
Loge Reserved	33	31	29	27	29
Picnic Area	32	30	28	26	28
Hcap — Field	27	25	23	19	23
Mezzanine Reserved	27	25	23	19	23
Upper Box	27	25	23	19	23
Loge Reserved, Back Rows⋆	16	14	12	8/5	12
Mezzanine Reserved, Back Rows⋆†	16	14	12	8/5	12/11
Upper Reserved⋆	16	14	12	8/5	9
Hcap — OF	14	12	10	8/5	10

⋆ The price of value games in 2003 was $8 and in 2004 was $5.

† The price was 12 in 2002 and 11 in 2001.

FIGURE 5.11

Prices before and after the introduction of variable pricing ($/ticket).

Year	Ticket revenues	Concession revenues
1999	61.17	33.76
2000	77.82	33.40
2001	84.48	30.05
2002	86.78	32.66
2003	77.62	23.41
2004	72.02	26.96

FIGURE 5.12
Ticket revenues and concessions revenue. Source: New York Mets

David Howard, the Mets' senior vice president for business, explains that

> The three principal factors that determine why a fan goes to a game are time of year, day of week and the opponent. In the summer months, attendance rises with school out, then we see a difference in attendance for weekends than midweek, and there's a different demand for Yankees series than any other.

In 2003 the Mets initiated a variable ticket pricing plan. Each of the 81 home games was assigned to one of the four pricing tiers: value, bronze, silver, and gold. For example, many of the weekday games in April and September were the least popular and were placed in the "value" category ranging from $8 to $38 per ticket depending on location in the stadium. By the same token many of the weekend games during the summer months or those games against competitors such as the Yankees and Cardinals were categorized as "gold" and priced from $16 to $53.

THE SIMPLE ECONOMICS OF VARIABLE PRICING

The idea of variable pricing is rather simple: to the extent that different games are subject to different demand patterns, setting different prices for different games allows for a better adjustment of pricing to demand, leading to higher total revenues. Consider the example

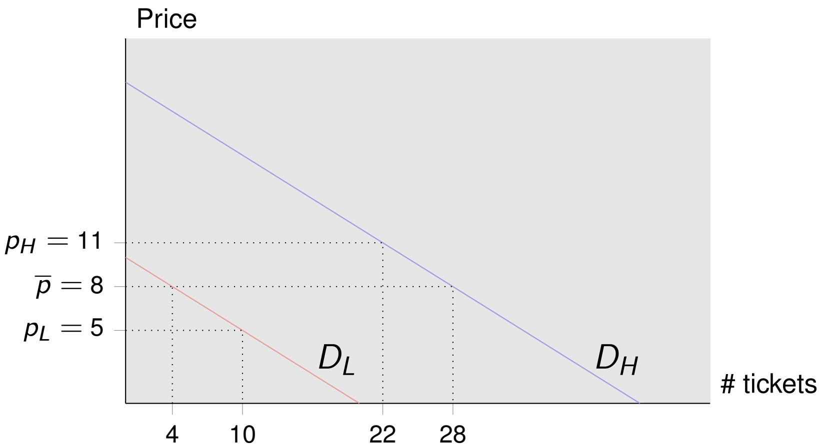


FIGURE 5.13
Variable pricing, demand and revenue

depicted in Exhibit 5.13, where two games have two demand patterns, D_L and D_H. By setting a uniform price, say $\bar{p} = 8$, total revenue is given by $8 \times (4 + 28) = 256$. However, by setting the optimal individual prices, $p_L = 5$ and $p_H = 11$, total revenue is given by $5 \times 10 + 11 \times 22 = 292$, an increase of 14%.

The second thing to notice is that increased variability in pricing should be associated to decreased variability in attendance. In the example depicted in Exhibit 5.13, uniform pricing leads to attendance levels $q_L = 4$ and $q_H = 28$, whereas variable pricing leads to attendance levels $q_L = 10$ and $q_H = 22$, a considerable lower cross-game variation in attendance.

DEMAND DETERMINANTS

In order to go from theory to practice, we need to know what the demand for each game is (or is expected to be). As a preliminary issue, we must decide what data to analyze. Exhibit 5.14 shows the number of tickets sold per game for three selected seating sections. Games are ordered on the horizontal axis (81 per season) and the number of tickets sold is shown on the vertical axis. (The term "attendance" in the Exhibit's title is not entirely correct as there may be tickets sold that do not correspond to actual attendance. However, for simplicity, if with some abuse of terminology, we will refer to attendance and tickets sold indiscriminately.)

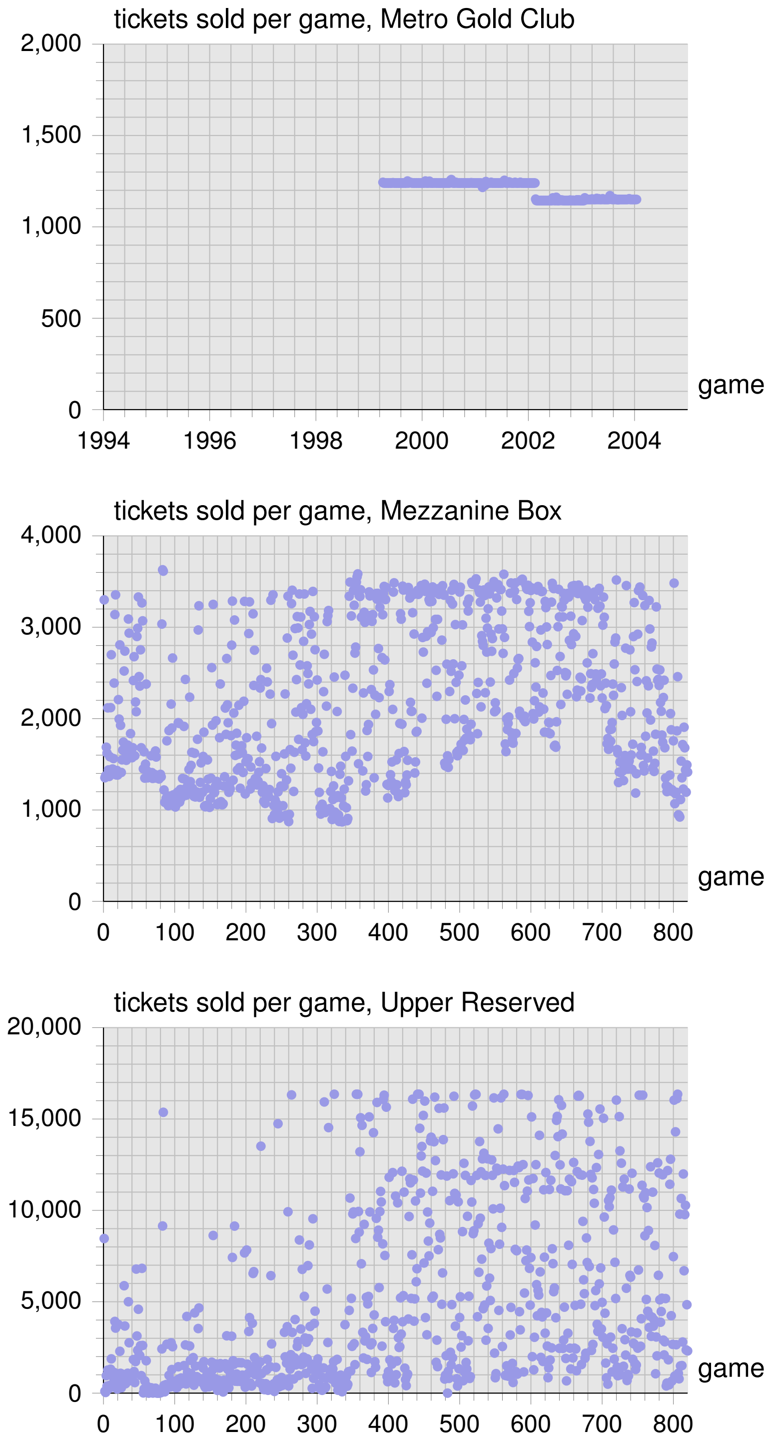


FIGURE 5.14
Attendance by game and selected seating sections.

Take for example the Metro Gold Club section. As can be seen from the top panel in Exhibit 5.14, almost all games have the same number of tickets sold. The reason is that most (if not all) of the tickets sold for this section are sold as season tickets. This in turn implies that the data cast no light on the issue of ticket demand on a game-by-game basis. Moreover, variable pricing has little effect on demand other than the effect it has on the overall cost of a season ticket (which in the particular case at hand was relatively small).

Consider now the case of Mezzanine Box tickets. The problem with this data is that, prior to 2003, we find a high density of points close to the maximum number of tickets sold (section capacity). This likely reflects capacity constraints, situations when ticket demand exceeds ticket supply. From an estimation point of view, this creates difficult problems. For example, the data suggests that the dispersion in the number of tickets sold increased with the introduction of variable pricing. However, the pre-2003 values are likely to underestimate the degree of demand dispersion, since many values are censored by Mezzanine Box capacity constraints.

Finally, the Upper Reserved case appears to be free from the problems raised by the Metro Gold Club and Mezzanine Box sections. First, the fact that there are so many games with so few tickets sold suggests that season tickets do not play an important role. Second, the low density of points near the top suggests that capacity constraints are not very important. For these reasons, we will henceforth concentrate our analysis on Upper Reserved ticket sales.

Based on thirteen years of historical data (1994–2004), Exhibit 5.15 presents the results of a simple OLS regression where the dependent variable is the number of tickets in the Upper Reserved section sold for a given game. Several other explanatory variables were considered but their effect is not significant. These include the Mets' win percentage at the time of the game, the opponent's win percentage at the time of the game, the starting pitcher (a baseball expert classified starting pitchers on a 1-2-3 scale); the weather forecast (temperature and precipitation).

The explanatory variables left from this elimination process are all 0-1 variables; the regression coefficient is therefore easy to interpret. For example, a game played on a weekend sells on average 1078.63 more tickets than a game played on a weekday; and so forth. Notice that, statistically, all coefficients are highly significant. From an eco-

Dummy variable	Coefficient	St. Dev.	z	p
Weekend	1078	402	2.68	0.01
Evening	-905	391	-2.31	0.02
Season opener	8196	1373	5.97	0.00
July	2410	411	5.86	0.00
August	1425	415	3.43	0.00
September	1555	464	3.35	0.00
October	3774	1176	3.21	0.00
Yankees	9169	1002	9.15	0.00
Constant	401	634	2.21	0.03

Year dummies included. $N = 651$. $\bar{R}^2 = 0.44$

FIGURE 5.15
Upper Reserved ticket demand

nomic point of view, the main determinants of ticket sales are, on a first level, playing against the Yankees or playing the season opener (about 9,000 extra tickets — in a section holding about 16,000); and at a second level playing in July or in October (about 3,000 extra tickets sold).

Finally, notice that the regression's $\bar{R}^2$ is only 44%, that is, the model only explains a little less than half of the total variation.

☐ **The effect of price on demand.** The most notable absence from the above list of explanatory variables is price. The main reason is that there is not enough variation in prices to obtain a reliable estimate of the price elasticity of demand. For a given stadium section and until 2002, all games during a given season were sold at the same price. Presumably, seasons where the team was better — or expected to be better — lead to higher prices. In 2003 and 2004, different games were sold at different prices, but the choice of price was endogenously determined by what the Mets expected would be higher demand games. In summary, prices are endogenously determined based on expected demand, which is a function of observable variables as well as variables that the analyst cannot observe (at least not this analyst).

To put this in a more dramatic way, consider Exhibit 5.16, depict-

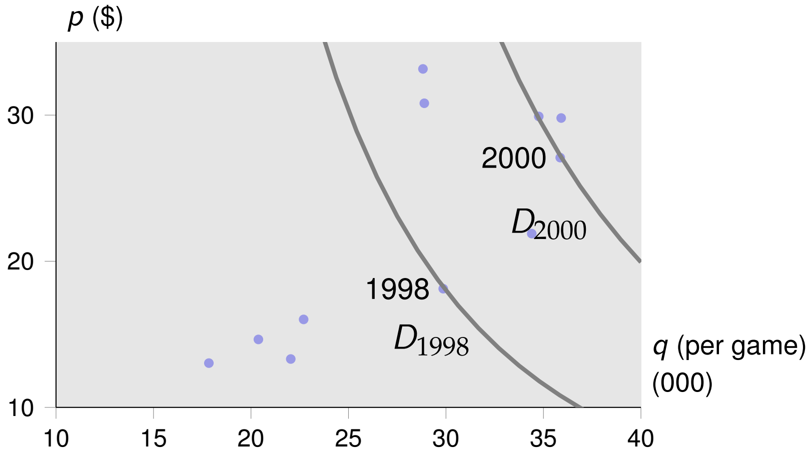


FIGURE 5.16
Average price and total annual ticket sales

ing total annual sales and price for each season from 1994 to 2004. If we were simply to regress quantity on price, we would obtain a positive coefficient, which would be contrary to the law of demand and common sense.

While it is difficult to control for the factors that influence demand and prices (and thus control for the endogeneity of price formation), an event took place in 2004 that may allow us to say something about the effect of price on demand. With respect to the the 2003 season, in 2004 the Mets maintained all of their prices except the price of Value games in the cheaper seats (Out Field, Upper Reserved, Mezzanine Reserved, Loge Reserved). Specifically, the price dropped from \$8 to \$5.

By looking at the evolution of relative demand (Value with respect to Bronze, Silver and Gold), we may get an estimate of the effect of price on demand. The advantage of measuring variations in relative demand is that this insulates many of the unobservable factors that influence demand. Specifically, we may obtain an estimate of the demand elasticity by measuring the percent change in the ratio of demand for Value tickets with respect to other tickets; and dividing this by the percent change in price. Alternatively, we may measure the difference in log values:

$$\epsilon = \frac{\log(r_2) - \log(r_1)}{\log(5) - \log(8)}$$

Based on this method, we obtain a demand elasticity estimate of approximately $-.35$.

THE RISKS OF PRICE TIERS AND PRICE CHANGES

Changing ticket prices in sports can be a tricky business; it's more than just a matter of regression analysis and algebra: the risk of losing fan support is potentially significant — especially when a team has performed poorly. To account for this possibility, The Mets' new plan maintained the same price for 27 games and cut the price of 16 other games. However, anyone wanting to watch the Mets play the Yankees during the 2003 season had to pay 30 to 40% more than in the previous year.

Tiered pricing involves an additional risk: what if a Gold game opponent turns out to have a lousy season, or a Value game turns out to be a crucial playoff decider? One possibility is for the team management to change prices as the season progresses and more information arrives. However, David Howard guaranteed that the Mets would not raise or lower ticket prices as a function of team success.

A sports economics analyst claims that one reason why teams would not employ a differential pricing strategy is that they "would rather not acknowledge that one team's entertainment value is higher than another's."Then there is also the risk of upsetting opponents who find themselves in the "value" category. The Pirates were one such team, but General Manager Dave Littlefield did not seem very concerned: "I don't worry about those types of things. We have to spend our energies improving our club. Whatever games they choose for discounting is their decision."It was not clear whether the Pirates' players shared the same feeling.

TWO SEASONS OF VARIABLE PRICING

Based on data from two seasons of variable pricing, what can we say about the predictions from economic theory and the success of the Mets' strategy?

☐ **Assigning games to tiers.** The independent variables in the demand regression described in Exhibit 5.15, in addition to all taking values of 0 and 1, all correspond to information that is available to the Mets

Period	# obs.	Mean	Std Dev	Min	Max
1994–2004	804	5420.6	5079.9	0	16356
2001–2002	156	7444.8	4983.2	582	16355
2003–2004	153	7154.6	5011.2	364	16356

FIGURE 5.17
Upper Reserved section ticket sales (1994–2002). Descriptive statistics.

	"V"	"B"	"S"	"G"	Σ
V	8	7	10	1	26
B	11	19	18	12	60
S	4	23	8	5	40
G	3	11	4	9	27
Σ	26	60	40	27	153

	"V"	"B"	"S"	"G"	Σ
V	12	14	0	0	26
B	10	22	22	6	60
S	2	15	15	8	40
G	2	9	3	13	27
Σ	26	60	40	27	153

FIGURE 5.18
Upper Reserved ticket sales (2003–2004):
Left panel: actual tier (row) and tier by observed sales (column)
Right panel: tier by model prediction (row) and tier by observed sales (column)

at the beginning of the season. This leads to an interesting question: based on the information from the regression analysis, was the assignment of games to each of the tiers optimal? How different an assignment would one make taking into account the regression results?

It is difficult to answer the question, if not impossible. However, recall from our previous discussion that optimal pricing should lead to a reduction in the variation of attendance: attendance to high-demand games decreases (given higher prices) and attendance to low-demand games increases (due to lower prices). Despite this compression in attendance levels, we expect the ordering of attendance to remain as before price changes: high-demand games continue to be high-sales games.

How does the ordering of games implicit in the Gold-Silver-Bronze-Value classification reflect on actual demand: are Gold games those with higher sales and are Value games those with lowest sales. Similarly, suppose we use the regression model to assign games to tiers. How would the assigned tiers relate to sales in this alternative world?

Figure 5.19 provides an answer to the above questions. The left panel relates the actual tier assignment to actual demand. Each row corresponds, respectively, to Value, Bronze, Silver and Gold games. Of the 26 Value games (top row), we see that 8 had sales corresponding to the lowest group, 7 to the next lowest group, 10 to the second highest group and one to the very highest group. In other words, of the 26 games market out (and priced out) as Value games, one had actual sales at the level of a "Gold" game.

The right panel corresponds to a similar exercise, with the difference that, instead of the actual assignment of games to tiers, we use the assignment that would follow from the regression model: the games with lowest predicted demand are chosen as Value games, etc.

If the assignment of games to tiers were perfect (and prices chosen so as to maintain the ordering of demand) we would expect positive values along the main diagonal and zeros on all of the off-diagonal cells. This is clearly not the case, which may happen for a variety of reasons. First since we are setting the same price for all games in each tier, it is likely that the lowest Gold games will have lower demand than the highest Silver games, and so forth. Second, it may be that demand is such that actual sales given optimal prices are lower for Gold games than for Silver games.

Comparing the two panels in Figure 5.19 we notice that the right panel has greater concentration of values along the main diagonal. However, the differences are rather small, especially compared with the overall level of variation.

Finally, despite the level of variance, both panels suggest that the assignment of games to tiers is positively correlated to actual demand: even with price differences, lower-tier games sell less than higher-tier games.

☐ **Is there a labeling effect?** One tantalizing possibility is that the mere assignment of a game to a given tier influences demand beyond what the game's characteristics would imply (opponent, time of day, etc). There are at least two reasons why such a signalling effect might be at work. First, fans who are unaware of the quality of the game (e.g., tourists) may take tier assignment as a signal of game quality and adjust demand accordingly. Second, to the extent that the fun of watch-

		Predicted		
Tier	# obs.	$\epsilon = 0$	$\epsilon = -.35$	Actual
Gold	27	9815	8907	8610
Silver	40	7307	6947	6192
Bronze	60	6489	6510	7572
Value	26	5693	7141	6161
All tiers	153	7155	7155	7155

FIGURE 5.19
Upper Reserved section ticket sales per game (2003–2004). Predicted and observed values.

ing a game depend on how many other people watch the game, tier assignment may serve as a "coordination" device: fans avoid Value games because they expect other fans to avoid such games.

In order to explore the possibility of such signalling and coordination effects, we compare actual demand with predicted demand based on observable characteristics. By grouping games by tier we are then able to test whether there is a tier effect at work. Exhibit **??** presents the results. First, we compute predicted demand based on the model used to produce Exhibit 5.15 (that is, without considering price effects. Next, we consider the effect of price to be uniformly captured by the demand elasticity $\epsilon = -.35$ estimated from the 2004 "experiment." Finally, we rescale all values such that average attendance matches the actual observed values.

The results are somewhat inconclusive. Actual attendance at Value games is considerably lower than the model predicts, suggesting that there may indeed be a signaling effect (that is, being labeled a Value game reduces demand, everything else constant). Notice that the same is true for the opposite tier, Gold. However, the difference is much smaller. Moreover, one explanation for the difference is that capacity constraints are active in many of the Gold games (that is, demand is greater than actual sales). Regarding Silver and Bronze games, the difference between predicted and actual attendance is contrary to what a signaling theory would predict. In fact, if the theory were true than the results suggest that Silver is a negative signal, whereas Bronze is a positive signal. In any event, it is well to recall that the statistical model only accounts for a little less than one half of the total variation.

THE JURY'S STILL OUT

During the 2002 season, the Mets raised $87 million dollars in ticket sales. The next season, the first season with variable pricing, the same figure dropped to $78 million. Admittedly, the relevant comparison would be $78 million versus the 2003 total had the Mets stuck to their previous strategy. Still, the early experience does not bode well for variable pricing. However, sports and economics experts such as Dan Migala, executive editor of Team Marketing Report, are optimistic. "I think it's the wave of the future because it makes economic sense. It maximizes revenues yet it doesn't alienate fans who don't have a huge wallet to draw upon. In years past, you saw price increases running across the board, but now there's a lot more sophistication."

UPDATE

In 2005, the Mets continued tweaking with its pricing structure, this time adding a fifth tier, "Platinum." "Opening Day, the Yankee games, we could get basically any price we want for those," said Dan DeMato, director of ticket operations. He added that "we're leaving money on the table if we don't," and "we look at this the same way as the scalpers and brokers and eBay are looking at it."

As of 2009, about half of Major League Baseball's teams have some form of variable pricing.The New York Yankees are one of the hold-out teams. "People come to see the Yankees and not the visiting club," said a spokesperson, adding that "they want to see the [new] stadium too. It's just a different view."(Note: in addition to the sources linked above, the *NY Mets* case draws on a number of pieces in various sources that cannot be hyperlinked. See this version for a complete list of references.)

	1st year	renewal
Classic Print	198	396
All Access Digital	198	396
Print & Digital	222	444

FIGURE 5.20
Wall Street Journal

5.5. OTHER EXAMPLES

Market segmentation is prevalent in the media and entertainment space. In this section we cover a few examples. As we will see, although in theory we can classify segmentation strategies into nicely defined categories — perfect price discrimination, segmentation by indicators, segmentation by self-selection —, in practice we observe a combination of various features in any given firm's pricing strategy. (Another Yogi Berra adage that applies here: "In theory there is no difference between theory and practice. In practice there is.")

Consider first Table 5.20, reflecting the pricing of the *Wall Street Journal*. Their pricing is a case of mixed bundling (cf Section 5.3): readers can purchase an individual version (print, digital) *or* the bundle combining the two versions. As mention in Section 5.3, this can be interpreted as a versioning strategy, that is, segmentation by self-selection.

There is an additional feature in *Wall Street Journal*'s pricing: first-time subscribers pay lower prices. This is a common feature in many services. We expect new readers to be less hooked on to the WSJ, thus having a higher price elasticity of demand (and/or lower willingness to pay). Since we can observe whether a given reader is or is not a first-time reader, we can effectively set a different price for those readers. This is therefore a case of segmentation by indicators.

Consider next the case of Netflix. As of 2020, it provided its consumers with three primary monthly pricing plans, as follows:

- Basic, $8.99 per month. This plan doesn't provide high definition viewing and its programs can only be watched on one screen at a time.

- Standard, \$12.99 per month. This plan offers HD videos and allows for two simultaneous viewings.
- Premium, \$15.99 per month. This plan includes the ability to watch four screens at the same time. It's also the only item on the Netflix that offers a 4K viewing option.

This is clearly a case of versioning (similar, for example, to the iPhone case considered earlier in the chapter). Notice that there are various instances of damaged-product versioning going on here: Netflix does not save any costs by not allowing its customers to watch videos in HD. However, it's important that the basic plan prevent viewers from doing so. Otherwise, the medium-valuation viewers might be tempted to switch from the Standard to the Basic plan.

Consider next the case of Spotify. You can create a free account, but this is ad supported, that is, your will have to listen to ads between song plays. As of 2020, Spotify offered a number of ad-free plans:

- Individual. \$9.99/month after offer period (one month free). As the name suggests, this corresponds to one account only.
- Family. Up to 6 accounts. \$14.99 per month after offer period (one month free).
- Student. \$4.99/month after offer period (one month free). This is very similar to the individual account, but also includes a Hulu (ad-supported) plan as well as SHOWTIME.

Here we find multiple instances of segmentation by indicators. First, as in the case of the *Wall Street Journal*, we see that a discount is offered to first-time subscribers, in the present case in the form of zero price during the first month (cf Chapter 4). Second, students get a lower price — half price — for essentially the same product (individual subscription).

We also see here a case of segmentation by self-selection: The family rate corresponds to a sort of quantity discount which effectively leads different customers to paying different prices by self selection (in other words, Spotify does not need to check that you are a family in order to sell you a family subscription).

Finally, consider the case of Manchester United's non-member ticket prices, as shown on Figure 5.21. This is a clear case of segmentation by indicators, in the present case by age. As I mentioned

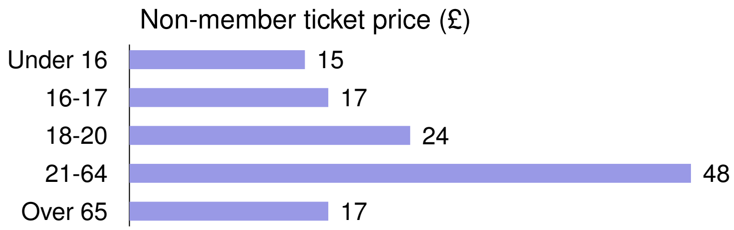


FIGURE 5.21
Manchester United 2018-2019

earlier, the idea of segmentation by indicators is to set a higher price in segments where consumers have higher willingness to pay; or, equivalently, in segments where demand elasticity is lower (in absolute value). Normally we do so by looking at some key correlation. In the present case, we expect there will be a positive correlation between disposable income and willingness to pay. And disposable income tends to be increasing with age up to retirement age (approx. 65) and then decreasing. Accordingly, we observe the same pattern in prices.

At this stage it might be helpful to recall that economists frequently use the term price discrimination in reference to market segmentation. The latter terminology, though less frequent among economists, is probably a better one. In particular, it helps distinguishing from practices such as age discrimination in the labor market, which is illegal. (In other words, treating an applicant or employee less favorably because of his or her age violates (in the US) the Age Discrimination in Employment Act (ADEA). By contrast, as of 2020, treating different customers differently according to age is not considered ilegal. In fact, setting different prices to customers of different genders is also not considered illegal (e.g., women often pay lower auto insurance rates).)

KEY CONCEPTS

market segmentation

Perfect price discrimination

personalized pricing

cookies

big data

market research

arbitrage

resale

gray markets

harvesting

pay what you like

versioning

participation constraint

incentive constraint

damaged goods

Bundling

pure bundling

mixed bundling

REVIEW AND PRACTICE PROBLEMS

■ **5.1. Market segmentation.** Explain the main types of market segmentation. Illustrate with specific examples.

■ **5.2. Secondary markets.** Many entertainment products (concerts, Broadway shows, sports events) sell out quickly, leading to a secondary market with hight ticket prices. Why don't sellers simply increase prices? What strategies would you suggest to address this problem? Provide concrete examples.

■ **5.3. Internet and market segmentation.** Does the Internet make price discrimination easier or more difficult?

■ **5.4. Harry Potter.** Consider the pricing of a *Harry Potter* paperback. What challenges would you face to implement the pricing policy derived on page 113?

■ **5.5. iPod.** Consider the following list of price and willingness to pay (WTP) for each version of Apple's iPod show in Table 5.5. Explain how versioning helps effectively selling the iPod for different prices to different user types.

■ **5.6. Market segmentation** Explain the main types of market segmentation. Illustrate with specific examples.

■ **5.7. HBO.** A subscription to HBO "over the top" costs \$14.99 in the US, \$8.99 in Nordic countries, \$11.44 in Hong Kong, \$8.45 in Spain. What kind of market segmentation strategy is this? What limitations might there be to its implementation?

■ **5.8. Movie tickets.** Figure 5.22 shows the cross-country relation between monthly wage and the price of several goods. How does this relate to the this chapter's theme?

Figure 5.22 shows the cross-country relation between monthly wage and the price of several goods. The idea is that, the greater wages are, the more people are willing to pay and the less sensitive to price changes they are. Thus we expect a positive relation between

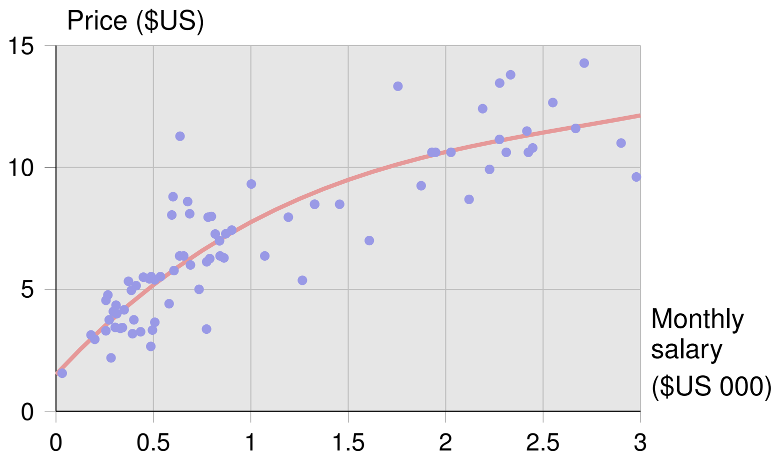


FIGURE 5.22
Movie ticket prices around the world

wages and prices.

■ **5.9. Optimal pricing and demand elasticity.** Consider the bottom panels in Figure 5.4. What are the values of the price elasticity of demand at the optimal price levels?

■ **5.10. Netflix.** When Netflix's business model was to lend DVDs, their pricing policy was to charge a monthly fee as a function of the number of DVDs the customer was allowed to have a given time. Specifically, the rates were as follows.

# DVDs	Fee ($)
1	8.99
2	13.99
3	16.99
4	23.99
5	29.99
6	35.99
7	41.99
8	47.99

What type of market segmentation strategy does this correspond to?

CHAPTER 6

FIRMS AND CONTRACTS

CHAPTER 7

STRATEGY AND GAMES

CHAPTER 8

TOPICS AND TRENDS