TOPICS IN MICROECONOMICS

INDUSTRIAL ORGANIZATION, AUCTIONS, AND INCENTIVES

ELMAR WOLFSTETTER
Humboldt University at Berlin

CAMBRIDGE UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK http://www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011-4211, USA http://www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1999

Printed in the United States of America

Typeset in Times Roman 11/14 pt. in LaTeX 2_ε [TB]

A catalog record for this book is available from
the British Library.

Library of Congress Cataloging-in-Publication Data

Wolfstetter, Elmar, 1945–
Topics in microeconomics : industrial organization, auctions, and
incentives / Elmar Wolfstetter.
p. cm.
ISBN 0-521-64228-0 (hb). – ISBN 0-521-64534-4 (pbk)
1. Microeconomics. 2. Competition, Imperfect. 3. Industrial
organization. I. Title.
HB172.W673 1999
338.5 – dc21 99-11535
CIP

ISBN 0 521 64228 0 hardback
ISBN 0 521 64534 4 paperback

Contents

Part I

Imperfect Competition

1
Monopoly

The best of all monopoly profits is a quiet life.
Sir John Hicks

1.1 Introduction

In this chapter we analyze the supply and pricing decisions of a pure, single-product monopolist facing a large number of price-taking buyers. We take the firm's choice of product as given and assume that consumers know all about product characteristics and quality. Moreover, we assume that the monopolist's market is sufficiently self-contained to allow us to neglect the strategic interdependency between markets. The strategic interdependency between markets is the subject matter of the theory of oligopoly with product differentiation.

Monopolies do exist. In the early days of photocopying, Rank Xerox was the exclusive supplier – some still use "xeroxing" as a synonym for "photocopying." Postal and rail services are (or have been) monopolized (things are changing fast in these sectors), and so are public utilities (gas and electricity) and computer operating systems, to name just a few. One can even find inconspicuous products that are subject to monopolization. For example, in Germany matches were exclusively supplied by a single Swedish supplier who had acquired a monopoly license from the German government during World War I, when the German government was hard pressed for foreign currencies. Similarly, gambling licenses are often issued by states to raise revenue. Moreover, there are many local monopolies, like the single hardware store in a small community, the bus line exclusively served by Greyhound, or the flight route, say from Ithaca to New York City, served by a single airline.

As these examples suggest, monopolization has a lot to do with the size of a market, but also with licensing, patent protection, and regulation – supported by law. If entry into a monopolized market is not prohibited, a monopoly has little chance to survive unless the market is too small to support more than one firm. Monopoly profits attract new entrants. And even if entry is prohibited, eventually patent rights expire, or rival firms spend resources to develop similar products and technologies or even to gain political influence to raid the monopoly license. Therefore, a monopoly is always temporary unless it is continuously renewed through innovations, patents, or political lobbying.

Monopolies – Weak and Strong A monopolist has exclusive control of a market. But to what extent a monopoly is actually turned into a fat profit depends upon several factors, in particular:

- the possibility of price discrimination,
- the closeness to competing markets,
- the ability to make credible commitments.

A *strong* monopolist has full control over his choice of price function. He can set linear or nonlinear prices, he can even charge different prices to different buyers. In other words, the strong monopolist can use all his imagination to design sophisticated pricing schemes to pocket the entire gain from trade, restricted only by consumers' willingness to pay. No one will ever doubt the credibility of his announced pricing policy.

In contrast, the *weak* monopolist is restricted to linear prices.[1] He cannot even price-discriminate between consumers.

Monopolists come in all shades, between the extremes of weak and strong. For example, a monopolist may be constrained to set linear prices, but he may be able to price-discriminate between some well-identified groups of consumers. Or a monopolist may be restricted to set a menu of nonlinear prices, like the ones you are offered by your long-distance telephone company and your public utilities suppliers.

In the following pages you will learn more about these and other variations of the monopoly theme. We will not only analyze the monopolist's decision problem under various pricing constraints, but also attempt to explain what gives rise to these constraints from basic assumptions on technology, transaction costs, and information structures.

We begin with the simplest analysis of the weak monopoly, also known as *Cournot monopoly*, in homage to the French economist Antoine Augustin Cournot (1801–1877), who laid the foundations for the mathematical analysis of non-competitive markets. Most of this analysis should be familiar from your undergraduate training. Therefore, you may quickly skim through these first pages, except where we cover the relationship between rent seeking and the social loss of monopoly, the durable-goods monopoly problem, and the analysis of regulatory mechanisms.

Finally, keep in mind that there are really two opposite ways to model pure monopoly. The most common approach – exclusively adopted in this chapter – describes the monopolist as facing a given market demand function and ignores potential actions and reactions by the suppliers of related products. The other, opposite approach faces the strategic interdependency of markets head on and views monopolist pricing as an application of the theory of oligopoly with product

[1] A price function $\mathcal{P}$ is called linear if it has the form $\mathcal{P}(x) := px$, where $p > 0$ is the unit price.

differentiation. While we stick, in this chapter, to the conventional approach, you should nevertheless keep in mind that there are many examples where oligopoly theory gives the best clues to the monopolist's decisions.[2]

1.2 Cournot Monopoly – Weak Monopoly

We begin with the weak or Cournot monopolist who can only set a linear price function that applies equally to all customers. The demand function, defined on the unit price p, is denoted by $X(p)$, and the cost function, defined on output x, by $C(x)$. Both $X(p)$ and $C(x)$ are twice continuously differentiable; also, $X(p)$ is strict monotone decreasing, and $C(x)$ strict monotone increasing. The inverse demand function, defined on total sales x, exists (due to the monotonicity of $X(p)$) and is denoted by $P(x)$. The rule underlying this notation is that capital letters like X and P denote functions, whereas the corresponding lowercase letters x and p denote supply and unit price.

In a nutshell, the Cournot monopolist views the market demand function as his menu of price–quantity choices from which he picks that pair that maximizes his profit. We will now characterize the optimal choice.

At the outset, notice that there are two ways to state the monopolist's decision problem: one, in terms of the demand function:

$$\max_{p,x} \quad px - C(x), \qquad \text{s.t.} \quad X(p) - x \geq 0, \quad p, x \geq 0,$$

and the other in terms of the *inverse* demand function:

$$\max_{p,x} \quad px - C(x), \qquad \text{s.t.} \quad P(x) - p \geq 0, \quad p, x \geq 0.$$

Obviously, the two are equivalent. Therefore, the choice is exclusively one of convenience. We choose the latter. Also, notice that the constraint is binding (the monopolist would forgo profits if he did sell a given quantity below the price customers are willing to pay). Therefore, the monopolist's decision problem can be reduced to the unconstrained program:

$$\max_{x \geq 0} \quad \pi(x) := R(x) - C(x), \tag{1.1}$$

where R denotes the revenue function:

$$R(x) := P(x)x. \tag{1.2}$$

1.2.1 Cournot Point

Suppose, for the time being, that X and C are continuously differentiable on $\mathbb{R}_+$, that revenue $R(x)$ is bounded, and that profit is strictly concave.[3] Then the decision

[2] Of course, also the opposite may hold, where standard monopoly theory gives the best clues to oligopolistic pricing. This is the case when reaction functions are horizontal.

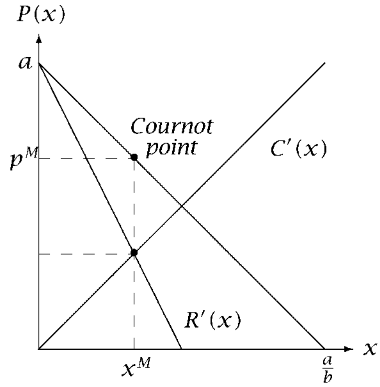

Figure 1.1. Cournot Point.

problem is well behaved, and we know that there exists a unique solution that can be found by solving the Kuhn–Tucker conditions:[4]

$$\pi'(x) := R'(x) - C'(x) \leq 0 \quad \text{and} \quad x\pi'(x) = 0, \quad x \geq 0. \tag{1.3}$$

In principle, one may have a corner solution ($x = 0$). But if $P(0) - C'(0) > 0$, an interior solution is assured, which is characterized by the familiar condition of equality between marginal revenue and marginal cost, $R'(x) = C'(x)$. Denote the solution by x^M, $p^M := P(x^M)$. The graph of the solution is called the *Cournot point* and illustrated in Figure 1.1.

Example 1.1 Suppose $P(x) := a - bx$, $a, b > 0$, and $C(x) := \frac{1}{2}x^2$, $x \in [0, a/b]$. Then profit is a strictly concave function of output, $\pi(x) := ax - bx^2 - \frac{1}{2}x^2$. From the Kuhn–Tucker condition one obtains

$$0 = \pi'(x) = a - 2bx - x. \tag{1.4}$$

Therefore, the Cournot point is $(x^M = a/(1 + 2b),\ p^M = a(1 + b)/(1 + 2b))$, and the maximum (or indirect) profit function is $\pi^*(a, b) := a^2/(2(1 + 2b))$.

Obviously, the monopolist's optimal price exceeds the marginal cost. But by how much? The answer depends upon how strongly demand responds to price. If demand is fairly inelastic, the monopolist has a lot of leeway; he can charge a high markup without suffering much loss of demand. But if demand responds very strongly to a price hike, the best the monopolist can do is to stay close to marginal cost pricing. This suggests a strong link between monopoly power and the price responsiveness of demand.

[3] Concavity of the revenue and convexity of the cost function – at least one of them strict – are sufficient, but not necessary.

[4] In case you are unsure about this, prove the following: 1) strong concavity implies strict concavity; 2) if a solution exists, strict concavity implies uniqueness; 3) the Weierstrass theorem implies the existence of a solution (you have to ask: is the feasible set closed and bounded?); 4) the Kuhn–Tucker theorem implies that every solution solves the Kuhn–Tucker conditions, and vice versa. Consult Appendixes C and D.

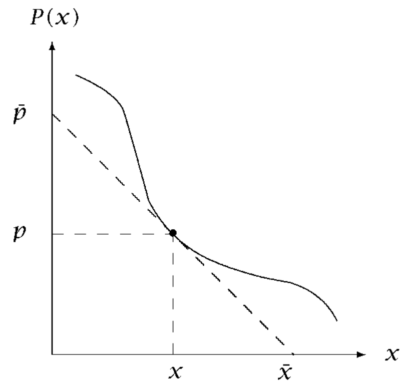

Figure 1.2. Relationship between Marginal Revenue and Price Elasticity $|\varepsilon| = p/(\overline{p} - p)$, $R'(x) = p - (\overline{p} - p)$.

The conventional measure of price responsiveness of demand is the *price elasticity of demand:*

$$\varepsilon(p) := X'(p)\frac{p}{X(p)}. \tag{1.5}$$

We now use this measure to give a precise statement of the conjectured explanation of monopoly power.

As you probably recall from undergraduate micro, marginal revenue is linked to the price elasticity of demand as follows (see also Figure 1.2):[5]

$$\begin{aligned} R'(x) &= P'(x)x + P(x) \\ &= P(x)\left[1 + P'(x)\frac{x}{P(x)}\right] \\ &= P(x)\left[1 + \frac{x}{X'(P(x))P(x)}\right] \\ &= P(x)\left[1 + \frac{1}{\varepsilon(P(x))}\right] \\ &= P(x)\left[\frac{1 + \varepsilon(P(x))}{\varepsilon(P(x))}\right]. \end{aligned} \tag{1.6}$$

Therefore, marginal revenue is positive if and only if demand responsiveness is high, in the sense that the price elasticity of demand is less than -1.

Using this relationship together with the Kuhn–Tucker condition (1.3) for an interior solution, one obtains the following optimal *markup rule*:

$$P(x) = \frac{\varepsilon(P(x))}{1 + \varepsilon(P(x))}C'(x). \tag{1.7}$$

[5] Note, by the definition of the inverse one has $P(X(p)) \equiv p$; hence, $P'(X(p))X'(p) \equiv 1$.

Another frequently used variation of this form is the *Lerner index* of monopolization:

$$\frac{P(x) - C'(x)}{P(x)} = \frac{1}{-\varepsilon(P(x))}. \tag{1.8}$$

These convenient forms should also remind you that the Cournot point always occurs where the price elasticity of demand is less than minus one, that is, where an increase in output raises revenue.

Monopoly and Markup Pricing

In the applied literature on industrial organization it is claimed that monopolistic firms often stick to a rigid markup pricing rule. This practice is sometimes cited as contradicting basic principles of microeconomics. Notice, however, that (1.7) is consistent with a constant markup. All it takes is a constant elasticity demand function and constant marginal cost.

Another issue in this literature concerns the problem of measurement. Usually, one has no reliable data on firms' cost functions. So how can one ever measure even such a simple thing as the Lerner index? As in other applications, a lot of ingenuity is called for to get around this lack of data.

A nice example for this kind of ingenuity can be found in Peter Temin's study of the German steel cartel in imperial Germany prior to World War I (see Temin (1976)). He noticed that the cartel sold steel also on the competitive world market. Temin concluded that the world-market price, properly converted using the then current exchange rate, should be a good estimate of the steel cartel's marginal cost. And he proceeded to use this estimate to compute the Lerner index. Make sure you understand the economic reasoning behind this trick.

Monopoly and Cost-Push Inflation

In economic policy debates it is sometimes claimed that monopolists contribute to the spiraling of cost-push inflation because – unlike competitive firms – monopolists apply a markup factor greater than 1. To discuss this assertion, it may be useful if you plot the markup factor $\varepsilon/(1+\varepsilon)$ for all $\varepsilon < -1$. Notice that it is always greater than 1, increasing in ε, and approaching 1 as ε goes to minus infinity and infinity as ε approaches -1.

Software Pirates and Copy Protection

As a brief digression, consider a slightly unusual Cournot monopoly: the software house that faces competition from illegal copies and in response contemplates introducing copy protection.

Legally, the copying of software is theft. Nevertheless, it is widespread, even among otherwise law-abiding citizens. Software houses complain that illegal copies rob them of the fruits of their labor and pose a major threat to the industry.

Suppose copy protection is available at negligible cost. Should the monopolist apply it, and if so, how many copies should he permit? A copy-protected program can only be copied $N \geq 0$ times, and copies cannot be copied again. Therefore, each original copy can be made into $N + 1$ *user copies*.

To discuss the optimal copy protection, we assume that there is a perfect secondary market for illegal copies. For simplicity, users are taken to be indifferent between legal and illegal user copies, and marginal costs of copying are taken to be constant.

Given these admittedly extreme assumptions, the software market is only feasible with some copy protection. Without it, each original copy would be copied again and again until the price equaled the marginal cost of copying. Anticipating this, no customer would be willing to pay more than the marginal cost of copying, and the software producer would go out of business because he knew that he could never recoup the fixed cost of software development.

An obvious solution is full copy protection ($N = 0$) combined with the Cournot point (p^M, x^M). However, this is not the only solution. Indeed, the software producer can be "generous" and permit any number of copies between 0 and $x^M - 1$ without any loss in profit. All he needs to do is to make sure that N does not exceed $x^M - 1$ and that the price is linked to the number of permitted copies in such a way that each original copy is priced at $N + 1$ times the Cournot equilibrium price, $(N + 1)p^M$.

Given this pricing-plus-copy-protection rule, each customer anticipates that the price per user copy will be equal to the Cournot equilibrium price p^M; exactly $x^M/(N + 1)$ original copies are sold; each original copy is copied N times; exactly x^M user copies are supplied; and profits and consumer surplus are the same as under full copy protection.

At this point you may object that only few software houses have introduced copy protection;[6] nevertheless, the industry is thriving. So what is missing in our story?

One important point is that copy protection is costly, yet offers only temporary protection. Sooner or later, the code will be broken; there are far too many skilled "hackers" to make it last. Another important point is that illegal copies are often imperfect substitutes, for example because handbooks come in odd sizes (not easily fit for photocopying) or because illegal copies may be contaminated with computer viruses. Instead of adding complicated copy protection devices, the monopolist may actually plant his own virus-contaminated copies in the second-hand market. Alas, computer viruses are probably the best copy protection.

Leviathan, Hyperinflation, and the Cournot Point

We have said that monopoly has a lot to do with monopoly licensing, granted and enforced by the legislator. Of course, governments are particularly inclined to grant such licenses to their own bodies. This suggests that some of the best applications of the theory of monopoly should be found in the public sector of the economy.

A nice example that you may also come across in macroeconomics concerns the *inflation tax* theory of inflation and its application to the economic history of hyperinflations. A simple three-ingredient macro model will explain this link (the classic reference is Cagan (1956)).

[6] Lotus is one of the few large software houses that rely on copy protection.

1. The government has a monopoly in printing money, and it can coerce the public to use it by declaring it *fiat money*. Consider a government that finances all its real expenditures G by running the printing press. Let p be the price index, let M^S be the stock of high-powered money, and suppose there are no demand deposits. Then the government's budget constraint is

$$pG = \Delta M^S \qquad \text{(budget constraint).} \tag{1.9}$$

2. Suppose the demand for real money balances M^d/p is a monotone decreasing and continuously differentiable function of the rate of inflation[7] $\hat{p}$,

$$\frac{M^d}{p} = \phi(\hat{p}) \qquad \text{(demand for money).} \tag{1.10}$$

3. Assume the simple quantity theory of inflation,

$$\hat{p} := \frac{\Delta p}{p} = \frac{\Delta M^S}{M^S} = \frac{\Delta M^d}{M^d} \qquad \text{(quantity theory of money).} \tag{1.11}$$

Putting all three pieces together, it follows that the real expenditures that can be financed by running the printing press are a function of the rate of inflation:

$$G(\hat{p}) = \hat{p}\phi(\hat{p}). \tag{1.12}$$

The government has the exclusive right to issue money, and it can force people to accept this money in exchange for goods and services (this is the origin of the term "fiat money"). However, even though it can set the speed of the printing press, the real expenditures that it can finance in this manner are severely limited. Therefore, the inflation tax is only a limited substitute for conventional taxes.

To determine these limits, simply compute the Cournot-point rate of inflation $\hat{p}^M$, defined as the maximizer of $G(\hat{p})$ over $\hat{p}$. Since the government's maximization problem is equivalent to that of a Cournot monopolist subject to zero marginal costs, it follows immediately that real government expenditures reach a maximum at that rate of inflation where the elasticity of the demand for real money balances ε is equal to -1:

$$\varepsilon(\hat{p}) := \phi'(\hat{p})\frac{\hat{p}}{\phi(\hat{p})} = -1. \tag{1.13}$$

Of course, this revenue-maximizing rate of inflation imposes a deadweight loss upon society, just like any other Cournot monopoly. The socially optimal rate of inflation is obviously equal to zero. However, alternative methods of taxation tend to impose their own deadweight loss, in addition to often high costs of collecting taxes. Keeping these considerations in mind, it may very well be that some inflation is optimal, depending upon tax morale and other institutional issues. Different countries with their different institutions may very well have different optimal inflation rates. Incidentally, these considerations are the background of current discussions on optimal currency areas.

[7] In macroeconomics it is often assumed that the demand for real money balances is a strict monotone decreasing function of the nominal interest rate. The latter is usually strongly correlated with the rate of inflation.

Table 1.1. *German hyperinflation 1923*

Month		Exchange rate, monthly average (Mark/$)
January	1921	64
January	1922	191
January	1923	17,972
July	1923	353,412
August	1923	4,620,455
September	1923	98,860,000
October	1923	25,260,208,000
November	1923	4,200,000,000,000

Source: Stolper (1964).

Another interesting application of the inflation tax concerns the theory of hyperinflations, like the one in Weimar Germany in 1923 (or most recently in Serbia, after the breakup of Yugoslavia). There, a government was in desperate need for funds, due to a fatal combination of events, from the exorbitantly high demands for reparations imposed by the Versailles treaty (aggravated by the French occupation of the Ruhr area in 1923) to a parliament torn between cooperation and conflict. Unable to finance its expenditures to any significant degree by explicit taxes, the government took recourse to the printing press. But the faster it set its speed, the fewer real expenditures it could finance in this manner. The result was a rapidly exploding rate of inflation, reflected in the catastrophic devaluation of the mark relative to the dollar reported in Table 1.1, and a complete breakdown of government financing.[8]

Some Comparative Statics

How does the Cournot point change if marginal cost or demand function shift? As always, such questions are meaningful only if uniqueness of the Cournot point is assumed. This is one reason why comparative statics is always pursued in a framework of relatively strong assumptions.

As an example, suppose C is a continuously differentiable function of a cost parameter α in such a way that higher α represents higher marginal costs: $C''_{x\alpha}(x, \alpha) > 0$. Also assume that the profit function is *strongly* concave in output and that the Cournot point is an interior solution.[9] Then, the optimal output is a differentiable function of α, described by the function $x^*(\alpha)$. And we can pursue comparative statics using calculus.

We now show that the monopolist's supply is strict monotone decreasing in α: $x^{*'}(\alpha) < 0$. For this purpose, insert the solution function $x^*(\alpha)$ into the Kuhn–

[8] At one point, the Reichsbank employed 300 paper manufacturers and 2,000 printing presses, day and night.

[9] Recall that *strong concavity* is strict concavity plus the requirement that the determinant of the Hessian matrix of the profit function (which is here simply the second derivative of this function) does not vanish. Strong concavity is always invoked if one wants to make sure that the solution functions are differentiable in the exogenous parameter, which is a prerequisite for the calculus approach to comparative statics.

Tucker condition (1.3), and one obtains the identity

$$\pi_x'(x^*(\alpha),\alpha) := R'(x^*(\alpha)) - C_x'(x^*(\alpha),\alpha) \equiv 0. \tag{1.14}$$

Differentiating it with respect to α gives, after a bit of rearranging,

$$x^{*'}(\alpha) = \frac{-\pi_{x\alpha}''(x^*(\alpha),\alpha)}{\pi_{xx}''(x^*(\alpha),\alpha)} = \frac{C_{x\alpha}''(x^*(\alpha),\alpha)}{R''(x^*(\alpha)) - C_{xx}''(x^*(\alpha),\alpha)} < 0. \tag{1.15}$$

This proves that the monopolist's optimal supply is strictly monotone decreasing in the marginal cost parameter, as asserted.

Two Technical Problems

We close the analysis of the Cournot point with two slightly technical problems. The first one concerns the existence of the Cournot point in the face of plausible discontinuities of demand or cost functions. The second explains how you should proceed if the profit function is not strictly concave. If you are in full control of your undergraduate micro, you may skip this exposition and move directly to Section 8.8.2.

An Existence Puzzle Suppose demand is unit elastic ($\varepsilon = -1$ for all $x > 0$), and the cost function is strictly convex with positive profits at some outputs. Then the profit function is strictly concave. Yet, the monopolist's decision problem has no solution.

The explanation is very simple. First, notice that revenue is constant for all positive x whereas cost is strictly increasing. Therefore, profit goes up as x is reduced (less output means higher profit), except if x is reduced all the way down to $x = 0$. Second, notice that there is no smallest positive rational number (there is no smallest positive output). Combine both observations, and it follows that there is no profit-maximizing choice of x. So which of our assumptions has failed?

As you check the assumptions one by one, you will see that almost all of them are satisfied. The only exception is the continuity of the revenue function, which is violated at precisely one point ($x = 0$).[10] This seemingly minor deviation changes it all.

The discontinuity of the demand function at $x = 0$ is something that one would not like to rule out. For example, applied economists often work with constant elasticity demand functions, all of which share this discontinuity property.

Another frequently encountered discontinuity that should not be excluded concerns the cost function. Recall, costs are usually decomposed into fixed and variable, where fixed costs are defined as $\lim_{x\to 0} C(x)$. Some fixed costs are *reversible* (or quasifixed), and some are irreversible or *sunk*. Whenever some fixed costs are reversible, one has $C(0) < \lim_{x\to 0} C(x)$, so that the cost function has a discontinuity at $x = 0$. In the face of it there is always a reasonable chance that the corner point $x = 0$ may be optimal. Therefore, watch out for a corner solution.

So, what shall you do if the demand or the cost function has such a discontinuity, and how can one assure existence of the Cournot point even in these cases? As in

[10] This discontinuity rules out the application of the Weierstrass theorem, which was invoked in the proof of existence of the Cournot point sketched in footnote 4.

other applications, a safe procedure is to break up the search for a solution into three steps: 1) search for a solution in the restricted domain $\mathbb{R}_{++}$ (an interior solution); 2) evaluate profit at the corner point $x = 0$; 3) choose the solution (either corner or interior) with the highest profit.

Since this procedure is cumbersome, one would of course like to know in which case existence of an interior solution is guaranteed so that the procedure can be stopped after round 1). A simple and often used sufficient condition is the following:

$$\lim_{x \to 0} [R'(x) - C'(x)] > 0, \qquad \lim_{x \to \infty} [R'(x) - C'(x)] < 0. \tag{1.16}$$

Make sure that you understand why this condition is indeed sufficient.

Example 1.2 Suppose the demand function has a constant elasticity $\varepsilon < -1$. Then it must have the form $X(p) = ap^{\varepsilon}$ (show this), so that the inverse demand function is $P(x) = (x/a)^{1/\varepsilon}$. Twice differentiate the revenue function $R(x) := P(x)x$, and you see that the revenue function is strictly concave. Now add the assumption that the cost function is convex and that condition (1.16) holds. Then the Cournot point has a unique interior solution.

The Cournot Point without Concavity Let us get another technical problem out of the way: characterizing the Cournot point if the profit function is not strictly concave. Concavity, as a local property, assures that a stationary point is indeed a maximum, and strict concavity, as a global property, assures uniqueness. But concavity is far too strong a requirement.

A popular weaker requirement is quasiconcavity. But, as a global property, quasiconcavity is often difficult to confirm or reject. As in other optimization problems, if a maximization problem is not concave (that is, if either the objective function is not concave or the constraint set is not convex), it is often a better procedure to look for some transformation of variables that leads to a concave problem. The trouble is, however, that there are no simple rules and that you have to be imaginative to find a transformation that does the job.[11]

In many applications one can safely assume that the cost function is convex. But one may feel less comfortable assuming concavity of the revenue function. So you may wonder whether one could not assume instead that the revenue function is quasiconcave and then obtain a quasiconcave profit function, which is really enough for a well-behaved decision problem. The answer is no. Just recall that the sum of a concave and a quasiconcave function need not be quasiconcave. Consult Appendix C if you are not entirely sure about this matter.

So what shall you do if the profit function is continuous but not quasiconcave in output or in any conceivable transformation of this variable? Well, you cannot avoid the tedious job of checking out all stationary points and all corners. Of course, only those stationary points can qualify where the profit function is (locally) concave. Therefore, you need only consider those stationary points at which the *second-order*

[11] An example of such a transformation of variables is spelled out in detail in Section C.8 of Appendix C.

or local concavity condition

$$\pi''(x) := R''(x) - C''(x) \leq 0 \qquad \text{(second-order condition)} \tag{1.17}$$

is satisfied. But you may still be left with fairly extensive computations to compare the profits at the remaining stationary and corner points.

Example 1.3 Suppose the cost function is S-shaped (strictly concave for low and strictly convex for high outputs), and suppose demand is linear. Then the profit function has two stationary points. But the profit function is only locally concave at the one point with the higher output. Therefore, only one stationary point survives the second-order or local concavity condition. However, this point need not be a profit maximum either. Indeed, if fixed costs are sufficiently high, it is always optimal to close down the firm and choose the corner point $x = 0$.[12] Draw a diagram to illustrate this case.

1.2.2 Deadweight Loss of Monopoly

Compared to a competitive firm, the Cournot monopolist earns higher profits (if he did not, the price would have to be equal to the marginal cost at the Cournot point). This shows that monopoly power redistributes welfare from buyers to sellers. But redistribution alone does not indicate any loss of social welfare, in the sense of the Pareto criterion. However, since the Cournot monopolist can only extract more from the consumers' willingness to pay by charging a higher unit price, the monopolist reduces welfare, unless demand is completely inelastic.

If the unit price rises above the competitive level, the consumers who continue to buy at the now higher price suffer a loss in consumer surplus that is however exactly offset by the seller's gain. However, those who quit buying at the higher price suffer a loss not offset by any gain to the seller. This *deadweight loss* of Cournot monopoly is illustrated in Figure 1.3, by the shaded area D.

As always, a deviation from the welfare optimum suggests that, with a bit of imagination, one can design Pareto-improving trades. For example, starting from the Cournot point, the monopolist could propose to his customers to supply an additional $x^* - x^M$ units in exchange for an additional payment equal to the cost increment (measured by the area under the marginal cost function between x^M and x^*), plus some small bonus. Miraculously, both buyers and sellers would be better off. However, the weak Cournot monopolist cannot take advantage of these gains from trade, because he is restricted to simple linear pricing schemes, for reasons that will have to be explained from basic assumptions concerning technology and information structures.

Conceptually, the deadweight loss of monopoly is the same as the deadweight loss of taxation. In the Middle Ages, it was popular to tax real estate on the basis of the size of windows. Due to the conspicuously high price of glass, the size of

[12] It is useful to distinguish two cases: 1) Suppose the average cost is higher than the price at the qualifying stationary point. Then the corner point $x = 0$ is definitely optimal if fixed costs are reversible (not sunk). 2) Suppose the average variable cost is higher than the price at the qualifying stationary point. Then $x = 0$ is optimal even if fixed costs are irreversible (sunk).

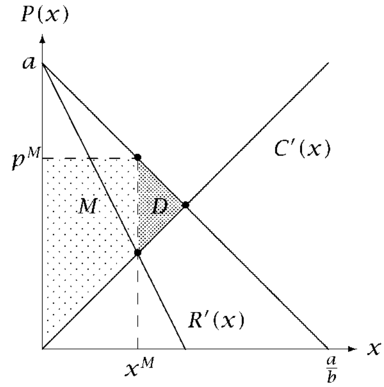

Figure 1.3. Deadweight Loss of Monopoly.

windows was correlated with wealth. Just as consumers reduce their demand when a monopolist raises the unit price, medieval citizens responded to the window tax by reducing the size of windows. In the end, they paid their taxes in any case. But on top of the direct reduction of wealth due to taxation, they sat in the dark – a visible example of the deadweight loss of taxation.

Can a government reduce or even eliminate that deadweight loss by means of corrective taxes? If the government has complete information about cost and demand functions, the task is easily accomplished. For example, a simple linear subsidy based on output – a negative excise tax – will do the job. The intuition is simple. An output subsidy smoothly reduces the effective marginal cost. By result (1.15), it follows immediately that the subsidy increases the Cournot equilibrium output. Therefore, one only needs to set the subsidy at the right level, and the monopolist is induced to produce the socially optimal level of output.

An obvious objection is that such a subsidy makes the monopolist even richer. However, this side effect of the output subsidy scheme can easily be eliminated by adding an appropriate lump-sum tax into the package.

To compute the appropriate subsidy rate and lump-sum tax, you should proceed as follows. In a first step, solve the monopolist's decision problem, given a subsidy rate s per output unit and a lump-sum tax T. Of course, the lump-sum tax does not affect the optimal output, but the subsidy does. Then, impose the requirement that the optimal output be equal to the competitive output x^M, implicitly defined by the condition $P(x^M) = C'(x^M)$. After a bit of rearranging the first-order condition you will find that the subsidy rate has to be set as follows:

$$s = \frac{-C'(x^M)}{\varepsilon(P(x^M))} > 0. \tag{1.18}$$

Finally, make the subsidy self-financing by setting the lump-sum tax equal to

$$T = sx^M. \tag{1.19}$$

It is as simple as that.[13]

[13] An even simpler mechanism is to impose sufficiently high penalties on any deviation from marginal cost pricing. This just shows that the regulation of monopoly is a trivial task if the regulator has complete information.

However, in most applications the regulation of Cournot monopoly is considerably more difficult. The main reason is that monopolists usually have private information about costs and sometimes even about their demand function. This raises a challenging mechanism design problem. We will address this issue in some detail in Section 2.3 of the next chapter.

Another problem has to do with the fact that monopolies are often the product of government regulation. It is hard to imagine that those agencies that restrict entry and thus permit monopolization will also tightly monitor these monopolies and direct them toward maximizing social welfare. And indeed, many economists are inclined to view regulation as industry-dominated and directed primarily to the industry's benefit. As Stigler (1971) put it: "... as a rule, regulation is acquired by the industry and is designed and operated primarily for its benefit."

1.2.3 Social Loss of Monopoly and Rent Seeking

The deadweight loss of monopoly D in Figure 1.3, however, tends to underestimate the social loss of monopoly. As Posner (1975) observed:

> The existence of an opportunity to obtain monopoly profits will attract resources into efforts to obtain monopolies, and the opportunity costs of those resources are social costs of monopoly too.

Under idealized conditions, the additional loss of monopoly is exactly equal to the monopoly profit measured by the area M in Figure 1.3. Therefore, the social cost of monopoly is the sum of the deadweight loss D and the monopoly profit M.

The additional loss may easily outweigh the traditional deadweight loss. For example, if consumers are identical and demand is perfectly inelastic, the deadweight loss vanishes, but the monopoly profit is as large as consumers' aggregate willingness to pay. This suggests that the additional cost component deserves close scrutiny.

The key assumption underlying the proposed inclusion of the monopoly profit as part of the social loss of monopoly is that obtaining a monopoly is itself a competitive activity. Even though there is perhaps no competition *in* the market, there is almost always competition *for* the market. The contestants spend resources to such an extent that, at the margin, the cost of obtaining the monopoly is exactly equal to the expected profit of being a monopolist. For if a monopoly could be acquired at a bargain, others would try to take it away until no net gain could be made. As a result, monopoly profits tend to be transformed into costs, and the social cost of monopoly is made equal to D plus M.

A simple argument illustrates this point. Suppose n identical firms spend resources, each at the level of the expected value z, to obtain a lucrative monopoly with the monopoly profit $\pi^M > 0$. Then each firm has a $1/n$ chance to win π^M. In equilibrium n and z are such that the expected value of the profit from participating the contest is equal to zero:

$$\frac{1}{n}\pi^M - z = 0. \tag{1.20}$$

Therefore, the monopoly profit is exactly equal to the overall cost of competition for the market,

$$\pi^M = nz, \tag{1.21}$$

as asserted.[14]

Assuming competition for the market is reasonable in many applications. For example, if monopoly is based on patents, many firms can enter the patent race for this monopoly.[15] Or if monopoly is based on public licensing, many firms can enter into the political lobbying or perhaps even bribery necessary to obtain a license or raid an existing one.[16]

1.2.4 Monopoly and Innovation

Schumpeter (1975) pointed out that in perfectly competitive markets the benefits of innovation are not cashed in by the innovator, but are immediately competed away by other firms. Therefore, competitive firms have no incentive to invest in innovations. In contrast, a monopolist can easily translate cost reductions and quality improvements into higher profits. This suggests that some degree of monopoly is a prerequisite for innovation.

Although monopoly induces innovation, it does not give rise to socially optimal innovations. Consider an innovation that reduces marginal cost by \$1. The maximum social benefit of that innovation is reached if and only if the price is lowered by \$1. However, by doing this, the monopolist would give away the entire gain to his customers. Therefore, the monopolist's price reduction must be less than \$1, which in turn destroys some of the potential social benefits of innovation.

Patents

The failure of competitive markets to provide incentives to innovate gives some justification for a system of patents that grants the innovator a transferable monopoly for a certain period of time. However, Arrow (1962) observed that if one firm is awarded a patent, the incentives to innovate are not the same as in an ideal world.

In order to see this, consider a competitive industry with product price p_0 equal to marginal cost C_0'. A new technology is discovered that reduces the marginal cost to $C_1' < C_0'$. The innovation is patented, and the owner of the patent becomes a monopolist for all prices below p_0. Of course, the new price cannot be above p_0 because of potential competition from firms equipped with the old technology. Therefore, if the unconstrained Cournot monopoly price p^M is above p_0, the patent holder will set the new price just marginally below p_0, leading to a monopoly profit equal to the shaded area in Figure 1.4, whereas if p^M is below p_0, the patent holder

14 Typically, the underlying "rent seeking" game has an equilibrium in mixed strategies, which is why z was interpreted as an expected value.

15 Incidentally, Plant (1934) criticized the patent system precisely on the ground that it draws greater resources into inventions than into activities that yield only competitive returns.

16 The case of bribery poses an intriguing problem. At first glance, one is inclined to argue that bribery is purely redistributional and therefore cannot qualify as a social loss component. However, if a political office is the recipient of substantial bribes, it is itself a lucrative monopoly, subject to its own competition for office. As a result, people will spend resources, for example in education, to be put in office and stay in office. Ultimately, it is these costs associated with competition for office that represent the social loss of monopoly.

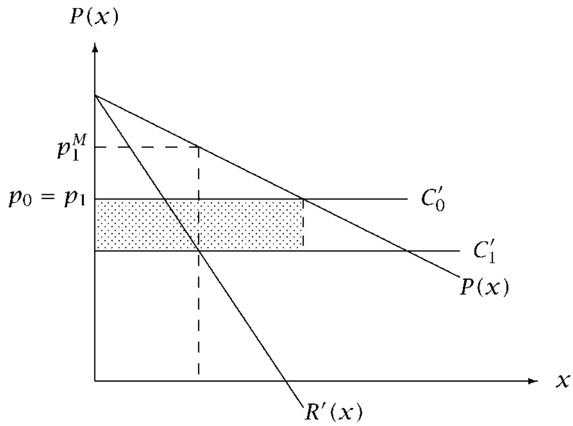

Figure 1.4. Innovation with Patent if $p_1^M > p_0$.

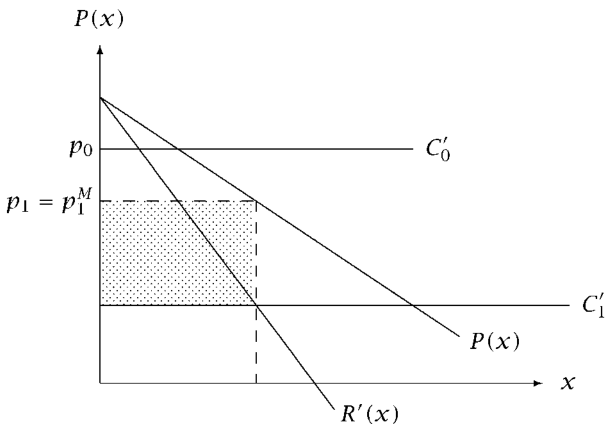

Figure 1.5. Innovation with Patent if $p_1^M < p_0$.

will lower the price to p^M and the monopoly profit is given by the shaded area in Figure 1.5. In either case, the incentive to innovate is clearly lower than in an ideal world, simply because the patent holder is unable to capture the entire increment in consumer surplus.

One can also see from the examples of Figures 1.4–1.5 that if a firm has a monopoly already at the beginning, it has less incentive to innovate than if the firm starts out in a competitive market. This is due to the fact that a competitive firm that acquires a patent not only captures the benefit of a cost reduction, but also gains market power by becoming a temporary monopolist.

To cover one of the two cases in greater detail, consider the example of Figure 1.4, which is elaborated in Figure 1.6. There, the following holds: If the initial state is already a *monopoly*, the gain from innovation G_m is measured by the area (see Figure 1.6)

$$G_m := B + C + D - (A + B) \qquad \text{(monopoly)},$$