Firm Behavior in Monopolistic Markets 10

This chapter covers . . .

- cognitive, technological, and regulative prerequisites for the existence of monopolies.
- how firms can use their monopoly to develop basic and sophisticated pricing strategies.
- the role of price discrimination in markets with imperfect information about the willingness to pay of the customers, and why the findings help to better understand pricing behavior in, for example, airline, software and hardware markets.
- the role of price discrimination between market segments and why the findings help to understand the debate about international price differences.
- how the informational demand for optimal pricing strategies is related to the optimal organization of firms.
- the economic-policy consequences of the above pricing models.

10.1 Introduction

> Like many businessmen of genius he learned that free competition was wasteful, monopoly efficient. And so he simply set about achieving that efficient monopoly. (Mario Puzo (1969), The Godfather)

The model of firm behavior under perfect competition has shown how a firm's supply is determined, if it takes prices as given. If it produces a positive quantity, a firm's optimal policy is generally determined by the condition "marginal revenues equal marginal costs," which simplifies to "price equals marginal costs," because marginal revenues and prices are identical under perfect competition. Therefore, the firm's supply function is the inverse of its marginal-costs function. The fact that it is willing to sell at marginal costs is also a prerequisite for the Pareto-efficiency of a competitive market.

However, the model has also shown that not all goods can be traded under conditions of perfect competition. One prerequisite is that there are many perfect

M. Kolmar, *Principles of Microeconomics*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-319-57589-6_10

substitutes for the good a firm produces. Another prerequisite is that the firm's production technology has to guarantee that the (long-run) marginal costs are non-decreasing. For different reasons, both conditions cannot be taken for granted. Consequently, one has to ask how markets function, if there is no perfect competition and keep one's focus on imperfect competition on the supply side. A similar logic applies to imperfect competition on the demand side as is, for example, frequently observed in regional labor markets, when there are only a few firms and labor has a low mobility. Another example is public procurement with firms that specialize in public projects. However, because imperfect competition on the supply side is the more commonly analyzed case, I will derive the implications of a supply-side monopoly for the functioning of markets.

One can start by analyzing a situation in which a firm has a monopoly for the supply of some good. The definition of a monopoly as a market with only one supplier of a good seems pretty obvious. However, this definition is not very operational, because it is unclear what exactly is meant by the idea that there is only one supplier. Hence, one first has to get to grips with a more operational understanding of what it takes for a firm to have a monopoly. Then one studies the optimal policy for a monopolist and analyze what this implies for the functioning of markets.

10.2 Conditions for the Existence of a Monopoly

Assume that one wants to bake a cake and needs flour. If one compares different retailers, one will find that each of them has different brands. In this sense, for example Migros is a monopolist for flour sold as "M-Budget Haushaltsmehl" (M-Budget flour), because Migros is the only supplier of this brand in the world. However, does this mean that Migros has a monopoly on "M-Budget Haushaltsmehl" in any meaningful economic sense? This question cannot be answered without further information. The reason is that two conditions have to be met in order to leverage the unique characteristics of a brand onto a monopoly position.

1. The customers have to be able to differentiate the product from other products and this ability to differentiate is reflected in the fact that alternative goods are not perceived as perfect substitutes. If consumers of flour are aware that there are different brands, but if this fact does not influence their decision which one to buy (because, for example, all they care for is the price of the flour), then flour is a homogeneous good sold by different suppliers, irrespective of the different brands. The fact that no other firm sells "M-Budget Haushaltsmehl" does not translate into the ability of Migros to raise prices above those of its competitors. However, if the customers consider the different brands to stand for non-homogeneous goods, then firms can use this willingness to differentiate between brands to charge brand-specific prices and optimal pricing becomes an integral aspect of the optimal firm policy. Indeed, flour seems to be a homogenous good to most customers and one may thus conjecture that the market for flour is, in fact, competitive. However, it is important to note that homogeneity

of goods has nothing to do with the good's *physical* characteristics or the brand name *per se*. It is the customers' willingness and ability to differentiate between goods of different suppliers that is a necessary prerequisite for a monopoly.

The willingness to differentiate can be assessed empirically by estimating the price and cross-price elasticities of demand. Intuitively, elasticities measure the percentage change of a variable that is caused by a one percent change in some other variable. If demand is ordinary and very price elastic, then it reacts strongly to price changes and there is no leeway to set prices actively. Similarly, the cross-price elasticity describes how demand changes, if the price of another good changes. If this elasticity is very large (in absolute terms) and the goods are close substitutes, then there is, again, little scope for price setting. An introduction to the concept of elasticities can be found in the mathematical appendix in Chap. 14.

If the existence of a monopoly position depends on the customers' ability and willingness to differentiate between products, then it must be an integral element of corporate communications and marketing to define and communicate relevant differences to other firms' products or to create them in the first place. From this point of view, even an ordinary product like flour becomes interesting: in recent years, the market has displayed increasingly differentiated products. For example, wheat flour has been differentiated by cultivation method (organic vs. conventional), origin (local vs. from somewhere else), etc. This differentiation has the purpose of transforming a formerly homogeneous product into a set of heterogeneous products for which – if the efforts are successful – differences in the willingness to pay exist that can be exploited by the firms. Two other examples are denim jeans and coffee.

- Denim jeans do not fundamentally differ in their functionality: they protect from weather, have pockets to store and carry small items, and so on. The physical characteristics of jeans seem to suggest that they are a fairly homogeneous product, which is sold on competitive markets. However, this reasoning does not take into account that producers of jeans can use advertising campaigns in an effort to create a specific brand image that adds additional "cultural" content to the product, from which customers can benefit: jeans do not only protect from weather, but customers send a specific social message by wearing a specific brand. The brand's image is transferred onto the customer, allowing the customer to perform a specific societal role; to belong to a specific group whose values are implicitly communicated by the brand name. In our societies, jeans and many other products are sophisticated mediums of communication and the communicative function often dominates, or even replaces, the primary, utilitarian one (think of intentionally ripped jeans). This is why firms often produce cultural narratives in which their products play an important role. If successful, there are many differentiated products with their own differentiated markets, like the markets for Levi's jeans, Diesel jeans, Wrangler jeans, and so on, and the firms have more or less extensive leeway to set prices. One can, for example, buy a

$\mathcal{B}$

pair of H&M jeans for $19.95 and a pair of Tom Ford Jeans for $990 (spring 2015).

- For many firms in the food industry, the wine market is *the* reference point for the development of marketing strategies. Because, for large parts of the population, it is fashionable to be a "wine connoisseur," a plethora of differentiated products exist, such that producers are, to some degree, able to exploit the customers' ability and willingness to differentiate by charging higher prices (which, *ceteris paribus*, translates into larger profits).
  Currently, coffee is probably one of the more interesting products in the food industry, because many producers – inspired by the market for wine – try to escape the dead end of the homogeneous-good market by "educating" the customer to distinguish different types of coffee. Historically, the lion's share of the market sold homogenous quality and customers were very price-sensitive, which implied a high degree of competition between suppliers. The central elements of the "third wave of coffee culture," as it is called, are referencing the origin of the coffee all the way back to the farm where it was cropped: an accentuation of the varieties of tastes of coffee and differences in cultivation methods, from the coffee cherry to the final beverage. This includes the introduction of quality standards, like the "Cup of Excellence" seal, and the training of customers with respect to the flavors and brewing methods. On top of this, the emergence of the third wave as a subcultural phenomenon in Portland, Chicago and San Francisco gives third-wave coffee specific cultural overtones, which makes it an attractive symbol for certain groups of customers (as in the jeans example) and which is important for the evolution of a niche product into the (profitable) mainstream. If this project succeeds, the "third wave" has the potential to change the logic of coffee markets.

2. If the demand for some product is not perfectly elastic (the market demand curve is not flat), then the potential for a firm to create a monopoly exists. However, a somewhat inelastic demand function is not sufficient, but merely necessary for a monopoly. In addition, it has to be impossible for other firms to undercut the privileged position by simply imitating the first firm's products. There are different reasons for why imitation might not be a possibility.
   - The producer has exclusive control over some necessary resource. For example, the "De Beers diamond monopoly" existed because De Beers controlled a large share of raw diamond mines.
   - The producer is a technology leader, such that other firms are not able to imitate the product, because of a lack of skills. A good example is the US-American telecommunications company AT&T. It became the first long-distance telephone network in the USA and made huge investments in research and development that allowed it to acquire crucial inventions. As a result, the company obtained near-monopoly power on long-distance phone services.
   - The state regulates market entry by creating public monopolies or by patents and trademarks. An example for high profit margins that are protected by

patents and trademarks was the Nespresso capsule system (most patents have expired by now). A slightly awkward example for a public monopoly was the German monopoly for safety matches (Zündwarenmonopol) from 1930, which granted exclusive rights to distribute safety matches within the borders of the German Empire to a monopolist (Deutsche Zündwaren-Monopolgesellschaft). It ended in 1983. Historically, most countries had public monopolies in the post and telecommunications industries, as well as for rail transportation. A lot of these industries were, at least partially, privatized and opened up for competition in the 1980ies and 1990ies.

- One has already seen that competitive markets cannot function, if either or both marginal and average costs are decreasing. In industries with such technologies, firms with larger market shares have an advantage, because they can produce at lower marginal and average costs and, hence, can sell at lower prices compared to their smaller competitors. Therefore, a large market share protects firms to some extend against competition. I have already discussed a special case of such a technology in Chapter 6, where I argued that club goods are sometimes called natural monopolies, because they imply decreasing average costs by increasing the number of users.
- Some products are characterized by the fact that the utility and, therefore, the customers' willingness to pay is increasing with the number of customers. This phenomenon is called a positive "network externality." Examples include telephones, word processing software and social media like Facebook and Twitter. Positive network externalities benefit firms with large market shares. Market share protects, to some degree, against market entries and competition, because a competitor with a smaller market share offers a *ceteris paribus* less attractive product.

If one or more of the above conditions holds, then the customers' willingness to differentiate translates into a (possibly temporary) monopoly position. The next subchapter analyzes how firms can make use of such a position.

10.3 Profit Maximization in Monopolistic Markets

The problem of a monopolistic firm is quite complex. As the previous subchapter explained, it has to decide about brand management and product development, needs to develop pricing strategies and has to take into account the political environment to protect and further its interests. In the following subchapter, I will reduce this complexity by focusing on the pricing aspect. Thus, the following analysis will assume that a firm has a monopoly in an already existing market with an established product. To understand optimal pricing policies, different cases have to be distinguished:

- The firm is not able to discriminate prices between customers; each customer buys at the same price. This is the standard model of monopoly theory. It will allow one to understand the important elements of an optimal pricing strategy

better. However, it is not quite realistic, because firms will usually try to differentiate prices between customers. Hence, the standard model alone cannot give one an appropriate picture of monopolistic behavior. The fact that price discrimination is an important tool of a firm's policy will become clear during the analysis.

- The firm is able to discriminate prices.
  - Perfect *(first degree) price discrimination* is possible. This is a theoretical benchmark, where the firm is able to set individualistic prices for each customer and can, in addition, discriminate by the quantity demanded. This model helps one to understand the consequences of price discrimination by bringing it to its extreme. It is, however, not particularly realistic, because it assumes that firms have all the relevant information about their customers and that a legal environment exists that allows them to use this information to charge different customers different prices. The availability of "big data" and the development of sophisticated algorithms that analyze the behavior of individuals on the internet may, however, allow them to move closer in the direction of perfect price discrimination in the future.
  - Price discrimination according to the quantity, quality or time demanded (*second degree price discrimination*). A firm often knows that there are different "types" of customers, who differ in their willingness to pay. From its market research, it may also know the different demand functions of these types. However, it does not know the willingness to pay of a specific customer. Therefore, the firm cannot condition the price on the customer's type directly, but needs to find alternative, indirect ways to skim off the different types' willingnesses to pay by appropriately designing products and prices. Examples are economy- and business-class airline tickets. Flexibility regarding the altering of the booking, leg space and service on board are important quality dimensions and airlines have an incentive to play with these variables to optimize profits.
  - Price discrimination according to specific customer attributes or customer segments (*third degree price discrimination*). The monopolist is able to discriminate between groups of customers, but cannot discriminate within groups. An example is price discrimination of multinational firms between countries.

The following subchapters will discuss these four cases.

10.4 The Optimal Production and Pricing Decision of a Monopoly Without Price Discrimination

The most studied case is the non-price-discriminating monopolist. Before discussing the circumstances under which it may be optimal to dispense with price discrimination, one should analyze how such a market functions. Assume the mo-

nopolist acts as a profit maximizer and produces with a technology that leads to a cost function $C(y)$.

A firm's market-research department estimates a market-demand function $x = X(p)$ for one of the firm's products. This function shows the amount of the good or service that can be sold at a given price p. In perfectly competitive markets, the perceived demand function is perfectly price elastic at the market price. This implies that the only information necessary to determine the optimal output is the existing or expected market price. However, this is no longer the case for a monopoly, where the firm needs to estimate the market-demand function with as much precision as possible. The organization of the firm is hence more complex: while in perfectly competitive markets, firms only need managerial accounting to determine marginal and average costs, but a firm in a monopoly market also needs a market research unit to estimate the demand function, because it is no longer a price taker. $\mathcal{B}$

Digression 25. Measuring Willingness to Pay $\mathcal{B}$

This chapter's analysis shows that an understanding of the likely responses of potential buyers to price changes is of considerable importance for firms. Despite this fact and despite the advances in pricing research, many firms price and develop products without an adequate knowledge of their customers' willingness to pay. Research has shown that only 8% to 15% of all companies use pricing strategies that are based on empirical assessments of buyer responses (Monroe and Cox, 2001), despite the fact that there is empirical evidence that even minor changes in prices can have important effects on profits (Marn et al., 2003). A lot of firms would rather use a strategy that may be dubbed "intuitive" pricing.

Marketing research offers a large variety of different techniques for measuring the willingness to pay. Broadly speaking, they fall into two different categories: "revealed" and "stated" preference models. Revealed-preference techniques infer a customer's willingness to pay from observed data. This can be market data, data that is generated while browsing the internet or data generated in experiments. Stated-preference techniques are based on surveys that are designed to elicit information about the willingness to pay. Examples of these include, among others, expert or customer surveys, and a technique that is called conjoint analysis. Conjoint analysis is a statistical technique where a product is partitioned into different attributes that together generate value for the customer (in the case of a car, these attributes might be mobility, versatility, status, etc.). Customers are then asked to rank or rate different bundles of these attributes. The results are used for the design and pricing of future products.

The different techniques to measure one's willingness to pay have their own strengths and weaknesses and it depends on the specific product and the available budget of the market-research department which method is applied.

At this point, one has to make a decision. One can assume that the monopolist sets a price and passively adjusts the produced quantity, $X(p)$, or one can alternatively assume that the monopolist decides on the quantity and demand determines the price at which the market clears. Both approaches lead to the same result but, since the second is somewhat simpler, it is the one that is usually applied. In order to do so, however, one has to infer the so called "inverse demand function" from $X(p)$, knowing that, for any price (p), the quantity (x) that can be sold is given by $x = X(p)$. Taking the inverse function of this demand function yields $p = X^{-1}(x)$, which determines the price that clears the market for any quantity offered by the firm. The convention from the previous chapters is to denote demand by x and supply by y. Given that one is analyzing the problem from the position of the monopolist who decides how much to supply, it makes sense, therefore, to replace x by y in the inverse-demand function that one denotes as $p = P(y)$.

If π denotes the firm's profit, then one can use this information to express it as revenues minus costs:

$$\pi(y) = P(y) \cdot y - C(y).$$

The problem faced by the firm's manager is to determine the quantity that maximizes profits. This quantity is implicitly defined by the necessary ("first-order") condition $\pi'(y) = 0$. (Assume in the rest of the book that this condition characterizes the global profit maximum. This is guaranteed, for example, if the second derivative of the profit function is globally negative, has a positive slope at $y = 0$ and has a negative slope for $y \to \infty$.) This yields:

$$P'(y) \cdot y + P(y) \cdot 1 - C'(y) = 0.$$

This condition has a straightforward economic interpretation: the first two terms represent the marginal revenues of an additional unit of the good, which can be decomposed into a price effect (first term) and a quantity effect (second term). The quantity effect is known from perfectly competitive markets. It measures by how much revenues increase, if an additional (marginal) unit is sold and the price stays constant. For infinitesimal changes, it is equal to the good's price. The price effect is new, however, and measures the loss in revenues of the firm, if it wants to sell another unit. To be able to sell another unit, the firm has to lower its price a bit to gain more customers. Since price discrimination is not possible, the firm also has to lower the price for those customers who would have paid a higher price. This "loss" can be interpreted as an opportunity cost and is measured by the price effect. The third term represents marginal costs and hence the "marginal revenues = marginal costs" rule holds.

$\mathcal{B}$ The first-order condition can be transformed into an easy rule of thumb, which is of great relevance in the management and pricing literature. Simple manipulations of the first-order condition show:

$$\begin{aligned} & P'(y) \cdot y + P(y) = C'(y) \\ \Leftrightarrow\ & P(y)\left(P'(y) \cdot \frac{y}{P(y)} + 1\right) = C'(y) \end{aligned}$$

$$\Leftrightarrow p\left(\frac{1}{X'(p)\cdot\frac{p}{X(p)}}+1\right)=C'(y)$$
$$\Leftrightarrow p\left(\frac{1}{\epsilon_p^X(p)}+1\right)=C'(y).$$

The transformation from the second to the third row follows from the definition of the inverse-demand function and the fact that demand is denoted by x. The manipulation between the third and fourth rows follows from the definition of the price elasticity of demand: $\epsilon_p^x(p)$.

One can start by making a short plausibility check: when the market-demand function has a negative slope, the price elasticity of demand is negative. Hence, the expression in brackets is smaller than one. Therefore, the condition can only hold if the price exceeds marginal costs. If there are perfect substitutes for the good, then the price elasticity converges to $-\infty$, such that the expression in brackets converges to 1, which leads to an intuitive conclusion: to comply with the above condition the price has to be equal to marginal costs, which is, of course, the case in perfect competition.

If the price elasticity is finite, however, then the optimal price exceeds marginal costs. The difference between price and marginal costs is called the "markup" and this rule of thumb is called "cost-plus pricing." Generally, the optimal markup is the higher, the less elastic market demand reacts to price changes. Thus, a manager who wants to set the optimal price needs information about two things: the marginal costs and the price elasticity of demand.

To further illustrate the optimal pricing decision of a monopolist, I derive the solution for the special case of a linear demand function $p(y) = a - b \cdot y$ and constant marginal costs $MC(y) = c$. In this case, revenues are equal to $R(y) = a \cdot y - b \cdot y^2$, with marginal revenues being $MR(y) = a - 2 \cdot b \cdot y$. Equating marginal revenues and marginal costs and solving for y yields the optimal solution $y^* = (a - c)/(2 \cdot b)$ and a price of $p^* = (a + c)/2$.

The demand function $x(p) = a/b - (1/b) \cdot p$ has a price elasticity of demand:

$$\epsilon_p^x(p) = -\frac{1}{b}\cdot\frac{p}{a/b-(1/b)\cdot p} = -\frac{p}{a-p}.$$

Thus, the absolute value of the price elasticity of demand is equal to:

$$|\epsilon_p^x(p^*)| = \frac{p^*}{a-p^*} = \frac{a+c}{a-c} \geq 1.$$

The fact that the optimal quantity and price are in the elastic part of the demand function is no coincidence: if demand is inelastic, then the monopolist can increase revenues by reducing output, because a one percent decrease in output increases the price by more than one percent. However, in this case, the initial level of output could not have been profit maximizing, since reducing output also reduces costs.

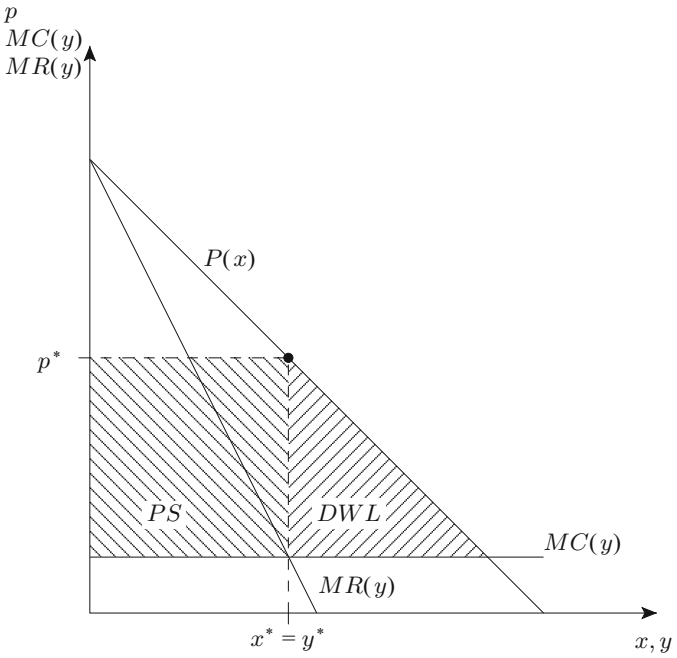

Fig. 10.1 The supply of a monopolist without price discrimination

Figure 10.1 shows the graphical solution to the profit-maximization problem of the monopolist. The optimum is given at the quantity where the marginal-revenue and the marginal-cost curves intersect. The associated price is defined by the value of the demand function for that quantity.

The solution to the linear model looks abstract, but is, in fact, rather intuitive. Assume for the moment that $c = 0$. In this case, profit maximization boils down to revenue maximization and revenues $p(y) \cdot y$ are the rectangular area under the demand function for any output y. Hence, the monopolist maximizes the size of this area. The optimal output is, therefore, given at $y = a/(2 \cdot b)$. If costs are positive, one has to restrict attention to the area where gains from trade are positive. $(a - c)$ are the maximum gains from trade of the customer with the highest willingness to pay, in this case. Thus, what one effectively gets is a "truncated" demand function $\bar{p}(y) = \bar{a} - b \cdot y$ with $\bar{a} = a - c$. The same argument as was used in the case of zero marginal costs also applies to this truncated function.

Digression 26. What Factors Determine Price Elasticities?
There are two factors that determine the price elasticity of demand: the customer's purchasing power (when the good becomes more expensive, customers can *ceteris paribus* afford less of it) and the willingness and possibility to substitute the good with another one. This second determinant makes the model applicable to markets with close but imperfect substitutes, e.g. different brands of jeans. The model's implication for markets with close substitutes is that markups have to be relatively moderate. The markup rule also gives one a first clue about how firms should invest in advertising and public relations: cost-plus pricing indicates that the markup in a market is negatively related to the price elasticity of demand. A marketing campaign should hence aim at making demand less elastic. In order to determine the optimal advertising budget, the firm needs to know the marginal revenues and marginal costs of advertising. The marginal revenues are determined by the change of the price elasticity of demand that is induced by another unit of advertising.

The monopolist's optimal price policy has interesting implications for economic policy: since the optimal price exceeds marginal costs, unexploited gains from trade will remain. There are still customers who are willing to pay a price that exceeds the firm's marginal costs, but at which the monopolist is not willing to sell. Consequently, the sum of consumer and producer surplus is below its maximum and the market is inefficient, a situation that is also called "market failure." The reason for this market failure is easy to grasp: if the monopolist wants to sell another unit of the good, he needs to lower the price by a bit. However, because price differentiation is not possible, the price needs to be lowered for all: not only for the marginal customers, but also for those who would buy the product at higher prices. If the firm wants to sell at marginal costs, the decrease in revenues due to the price effect exceeds the increase due to the quantity effect; therefore, the monopolist prefers to constrain the quantity to keep the price high. The inefficiency is given by the triangular area DWL in Fig. 10.1. DWL stands for *deadweight-loss* and is a measure for the inefficiency of the allocation. Before one can discuss the implications for economic policy one needs to have a better understanding of the causes of this inefficiency. Hence, the next analysis is a monopolist's optimal pricing policy, if price discrimination is possible.

10.5 Price Discrimination

Reasonable charges
Plus some little extras on the side!
Charge 'em for the lice, extra for the mice
Two percent for looking in the mirror twice
Here a little slice, there a little cut
Three percent for sleeping with the window shut
When it comes to fixing prices
There are a lot of tricks he knows
How it all increases, all them bits and pieces
Jesus! It's amazing how it grows!
(Alain Boublil (2013), Les Miserables (based on the novel by Victor Hugo))

10.5.1 First-Degree Price Discrimination

This subchapter covers the problem of a monopolist who is able to discriminate prices perfectly. Although this is not a very realistic assumption, as firms are usually unable to get all the relevant information, it is a useful theoretical benchmark and allows one to better understand the reasons for the above-mentioned inefficiency as well as current trends in firms' pricing strategies.

Perfect price discrimination is easy to analyze. In order to be able to pursue this strategy, the monopolist needs to know the willingness to pay of each individual customer. If this information is available, the firm will charge individualized prices for each customer, which equal that customer's willingness to pay. (It may be necessary to lower the price a bit to induce customers to actually buy the product. The remainder of this book will assume that indifferent customers behave in the interest of the firm. This assumption is innocuous with respect to its implications and simplifies the analysis.) Hence, in such a market, there is no uniform price, but a price function that is exactly equal to the inverse-demand function.

What is the minimal price at which the monopolist will supply the good? His profit increases as long as the price of the last unit exceeds the marginal costs of that unit. Hence, he will expand his supply up until the point where price equals marginal costs. This brings about a surprising result: the resulting market equilibrium is Pareto-efficient and the sum of consumer and producer surpluses are maximized. However, contrary to the case of perfect competition, gains from trade are not shared between the producer and the customers. Instead, the monopolist is able to skim off all the surplus in the market (see Fig. 10.2).

Digression 27. Price Discrimination in the Digital Age
Compared to second- and third-degree price discrimination, first-degree price discrimination has long been seen as a theoretical benchmark without much practical relevance, because the need for customer-specific information that

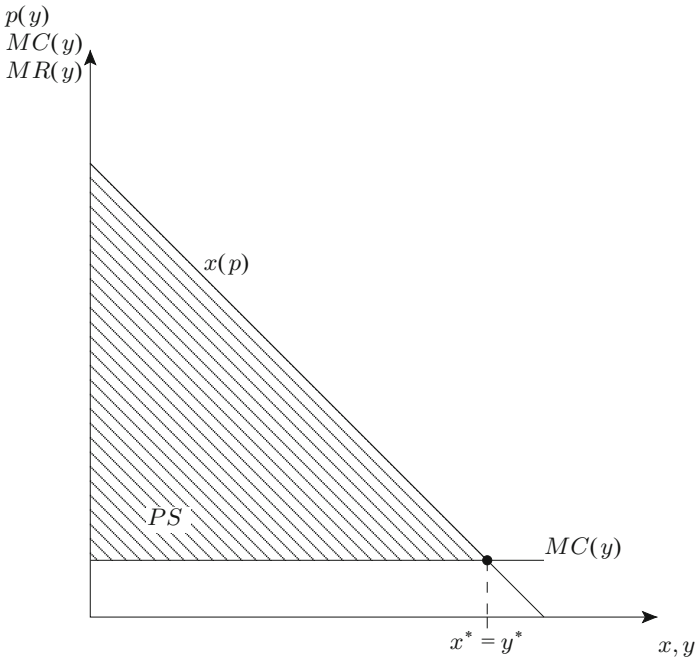

Fig. 10.2 Supply of the efficient quantity with first-degree price discrimination

is necessary to charge personalized prices was considered too extensive. On the other hand, moving in the direction of perfect price discrimination is extremely tempting for firms, because of its obvious consequences for profits. It should therefore not come as a surprise that e-commerce sites experiment heavily with pricing strategies that are based on the tracks people leave while browsing the internet.

A 2015 report in the hands of the President of the United States concludes that "the combination of differential pricing and big data raises concerns that some customers can be made worse off, and have very little knowledge why. [...] [M]any companies already use big data for targeted marketing, and some are experimenting with personalized pricing, though examples of personalized pricing remain fairly limited. [...] [P]roviding consumers with increased transparency into how companies use and trade their data would promote more competition and better informed consumer choice."

Hannak et al. (2014) analyzed the search results of 300 people who visited 16 online retailers and travel agencies from the US. They found that customers were shown different prices or different results for the same searches on nine of these 16 sites. For example, the online-travel company Expedia dis-

criminates prices according to the browsing history stored on the customers' computers. It is unclear, however, which type of browsing history triggers high prices. Another travel-agency, Travelocity, offered hotel rooms that were $15 a night cheaper if viewed from an iPhone or iPad. Home Depot displays higher prices and pricier products for smartphone users than for customers using desktops. In 2012, Wall Street Journal found that Staples discriminated prices according to the location of the device, and Orbitz discriminated prices between Mac and PC users, because data analysis revealed that Mac users are willing to pay higher prices for hotels.

These attempts to discriminate prices are still relatively crude, but the availability of more information and better algorithms may soon change the picture. Calo (2013) concludes that big data and better algorithms will enable companies to profile customers and deliver advertisements in a much more personalized way, also making use of the limited rationality of individuals. For example, Apple and Microsoft have filed patents for so-called "mood-based advertising" and Amazon is developing algorithms that tell them what the customers are likely to want before they place an order. This information is crucial for price discrimination, because it allows them to adjust prices or tweak choices while the customer is still searching. Google, for example, is filing a patent for an algorithm that can decide if a customer is likely to buy something and then to display a high price, while lowering the price for customers who have a low likelihood of buying.

Shiller (2014) studies the effects of including more information into pricing strategies on profits in the case of Netflix. He found that, compared to standard second-degree price discrimination, using the full set of information about web-browsing behavior increases variable profits by 1.39%, compared to 0.15% if pricing strategies are based on demographics alone. This may not sound like much but, compared to net profit margins of 2.34% in the US online retail industry, it makes a big difference.

What are the consequences of this discovery for economic policy and the regulation of monopoly markets? If one compares the case of a non-price-discriminating monopolist with that of a perfectly price-discriminating one, one can see that the monopolist will always choose to discriminate prices, if he can. Hence, the inefficiency in the market with a non-price-discriminating monopolist is caused by the inability to discriminate prices. There are three reasons why this instrument may be infeasible:

1. Price discrimination is illegal. The monopolist is then forced to charge the same price for any customer. In this case, it is the regulation of the market that causes the inefficiency. Market failure is not a result of some inherent tendency of the monopolist to be inefficient, but by a failed regulation of the monopoly (if the objective of regulation is to achieve efficiency).

An example for efforts to impede price discrimination by legal action is the "Robinson Patman Act," specifically Title 15, Chapter 1 §13 of the United States Code titled "Discrimination in price, services, or facilities". It is worthwhile studying the first paragraph of the act. "It shall be unlawful for any person engaged in commerce, in the course of such commerce, either directly or indirectly, to discriminate in price between different purchasers of commodities of like grade and quality, where either or any of the purchases involved in such discrimination are in commerce, where such commodities are sold for use, consumption, or resale within the United States or any Territory thereof or the District of Columbia or any insular possession or other place under the jurisdiction of the United States, and where the effect of such discrimination may be substantially to lessen competition or tend to create a monopoly in any line of commerce, or to injure, destroy, or prevent competition with any person who either grants or knowingly receives the benefit of such discrimination, or with customers of either of them: Provided, That nothing herein contained shall prevent differentials which make only due allowance for differences in the cost of manufacture, sale, or delivery resulting from the differing methods or quantities in which such commodities are to such purchasers sold or delivered: Provided, however, That the Federal Trade Commission may, after due investigation and hearing to all interested parties, fix and establish quantity limits, and revise the same as it finds necessary, as to particular commodities or classes of commodities, where it finds that available purchasers in greater quantities are so few as to render differentials on account thereof unjustly discriminatory or promotive of monopoly in any line of commerce; and the foregoing shall then not be construed to permit differentials based on differences in quantities greater than those so fixed and established: And provided further, That nothing herein contained shall prevent persons engaged in selling goods, wares, or merchandise in commerce from selecting their own customers in bona fide transactions and not in restraint of trade: And provided further, That nothing herein contained shall prevent price changes from time to time where in response to changing conditions affecting the market for or the marketability of the goods concerned, such as but not limited to actual or imminent deterioration of perishable goods, obsolescence of seasonal goods, distress sales under court process, or sales in good faith in discontinuance of business in the goods concerned."

2. The monopolist cannot prevent the resale of his products. In this case, resale markets evolve and so-called arbitrageurs specialize in buying and reselling the monopolist's products. For example, if there are two customers and one of them can buy at a high price only while the other can buy at a low price, it is worthwhile for both to trade at a price that is somewhere in between the two monopoly prices. Under ideal conditions, this process continues until only a uniform price prevails in the market.

   Why should a monopolist be unable to prevent the emergence of resale markets? To answer this question one needs to take a closer look at the types of contracts a monopolist can use because, from a legal perspective, the sale of a product or service is a transfer of a bundle of rights that is (explicitly or

$\mathcal{L}$

implicitly) specified in the underlying contract. If the monopolist can freely choose and constrain these rights, he can prohibit the resale of his products. He grants his customers the right of usage, but not the right of resale. The formation of resale markets can be precluded, if such contracts are legal and enforceable. However, in reality, it is often the case that courts do not enforce such contracts. They are sometimes legal in insurance markets, where insurance policies cannot be traded freely whereas, in traditional, consumption-goods markets, such contracts are usually illegal (for example in the European Union). However, the picture is more complicated for digital products, where complicated arrangements exist that regulate user rights. If resale were possible, then one would have to conclude that the source of market failure is, once again, an inefficient regulation of the market.

3. The monopolist does not have the information that is necessary to discriminate prices. The next chapter will cover the implications for the monopolist's profit maximization in more detail.

The preliminary conclusion that one can draw at this stage is somewhat surprising, because one cannot make a case against monopolies that is based on efficiency arguments. In light of the two models that have already been covered, one has to conclude that market inefficiencies are a result of an insufficient regulation of the market, not of the monopoly as such. This conclusion is, however, at odds with the intuitive feeling that most people and also most economists share, which states that monopolists are inherently inefficient. There are two ways to align this idea with the realizations discussed above:

First, it is, indeed, possible that a monopolist, who can perfectly discriminate prices, can lead to an efficient market outcome. Still, society might have goals that go further than efficiency. For example, distributive justice is a goal that many societies pursue. However, since monopolies are owned by individuals, who are also customers, customers will, in the end, receive the monopolist's profits. Therefore, one cannot judge the distributive properties of monopoly markets without further knowledge of the distribution of property shares among the population. However, there is empirical data about asset ownership in different countries. The demand for products and services is usually widely scattered, whereas property is concentrated in the hands of relatively few, rich individuals such that, from a more egalitarian point of view, a tradeoff between efficiency and distributive justice can exist. This may explain why some inefficiency is seen as the necessary price for a more egalitarian society. However, then the question arises of why the problem of distributive justice is not addressed more directly, for example by redistributive taxation, which would be more in line with the idea of the Second Theorem of Welfare Economics.

Digression 28. Pricing and Bounded Rationality
Finding ways to more effectively discriminate prices is a key topic in many industries. Strategies to discriminate go under names such as "dynamic pric-

ing," "power pricing," or "yield management." The basic problem behind all of these strategies is the same: how can a firm segment its customers into groups, which differ in their willingness to pay, and charge group-specific prices? Such strategies can actually lead to win-win situations between firms and customers, if there are close substitutes to the offered products (customers do not have to accept the offer, which is why they have to be better off if they accept it) and customers economize on search costs (e.g. finding an appropriate hotel for the planned trip to Vienna).

A related problem has to do with irrational or boundedly rational behavior. Based on findings from behavioral economics, some legal scholars criticize pricing strategies that systematically exploit customers' behavioral biases. Research in this field is still in its infancy.

Here is an example: assume a health club or gym uses a two-part tariff with an upfront-payment of L and a per-visit charge of p. If p equals marginal costs and L contributes to the financing of the club's fixed costs, then the contract is efficient. There is a lot of evidence, however, that customers overestimate the number of times that they will go to the club. This form of irrationality can be used by the club by charging p below its marginal costs and increasing L, which widens the gap between the surplus that the customer expects to receive from accepting the contract and what she actually receives. The customer finds this contract more appealing, but may end up with a negative consumer surplus.

Another example is a pricing strategy that is based on the *anchoring effect* discussed in Chap. 5. The rule of thumb on how to sell a good, for which customers have an unclear willingness to pay, is to place it right next to a similar, but much more expensive good. Williams-Sonoma added a $429 breadmaker next to their $279 model. The consequence was that sales of the cheaper model doubled, even though practically nobody bought the $429 machine (Ariely 2008): in this case, the expensive option acted as a price anchor.

A similar effect occurred in a study on purchasing patterns for beer (Poundstone 2011). In the first test, subjects had the choice between a regular beer for $1.80 and a premium beer for $2.50; 80% chose the premium beer. In the next test, a smaller and cheaper ($1.60) option was added. No one chose the cheap option, but orders for the premium beer dropped to 20%. In the final test, the cheap option was replaced by a large, expensive ($3.40) option. In this case, orders for the premium beer rocketed to 85%. This experiment shows that customers react to the pricing brackets in which products are displayed. Most people go for the "middle" option, which gives firms a lot of leeway in manipulating choices by developing adequate contexts for their products.

It might also be the case that the intuitive problems many people have with monopolies are not adequately grasped by the model. It is possible that the reason for a lack of efficiency of a monopoly is inherently dynamic, for example, because an ironclad monopoly position decreases the incentives to innovate. Such an argument, however, suggests that a completely different model is necessary to tackle the problem.

10.5.2 Second-Degree Price Discrimination

One central problem a monopolist faces when trying to discriminate prices is his lack of information about customers. There are two ways to solve this problem: investing in better information or using the given information to discriminate prices with maximum possible effectiveness. This subchapter will analyze the latter strategy.

In order to keep the problem simple and manageable, assume that there are two groups of customers, which can be differentiated by their willingness to pay. The firm knows each group's willingness to pay and also the respective group sizes, but cannot identify a customer as a member of one group or the other. An example for this situation is an airline that offers a flight from Zurich to Frankfurt, which is frequented by both business and leisure travelers. Business travelers have a higher willingness to pay for the flight and, in particular, for altering bookings flexibly. The airline knows the respective willingness to pay, as well as the groups' relative sizes, but cannot distinguish between individuals at the ticket counter (or on their homepage for that matter).

If the firm had all the relevant information, it would charge each customer according to his or her willingness to pay, such that both groups would receive their respective optimal offers, as in the case of first-degree price discrimination. From the point of view of the firm, the problem with asymmetric information is that a customer with a high willingness to pay may prefer the offer that is being provided for customers with a low willingness to pay. Their "own" offer gives the customer zero consumer surplus, whereas the offer provided for the other group not only differs in the quantity or quality of the good, but is also sold at a different price (both lower). Hence, buying a lower quality or quantity at a lower price might be profitable for the customer, if the lower price compensates for the loss in quality or quantity. In that case, all customers choose the offer that was designed for the group with the lower willingness to pay and the other offer remains a shelf warmer. This observation begs the question of what a firm's optimal reaction should be.

In order to answer the question, one should give the problem a more formal structure and analyze it graphically. In the following figures, one can see the quantity or quality of a good along the abscissa and the customers' willingness to pay along the ordinate. "Quantity" or "quality" can thereby be interpreted as an attribute, for which there are differences in the willingness to pay. In the airline example, quality can be interpreted as the flexibility to alter a booking, how much leg space there is or the level of service provided. If the good is a printer, quantity could refer to the

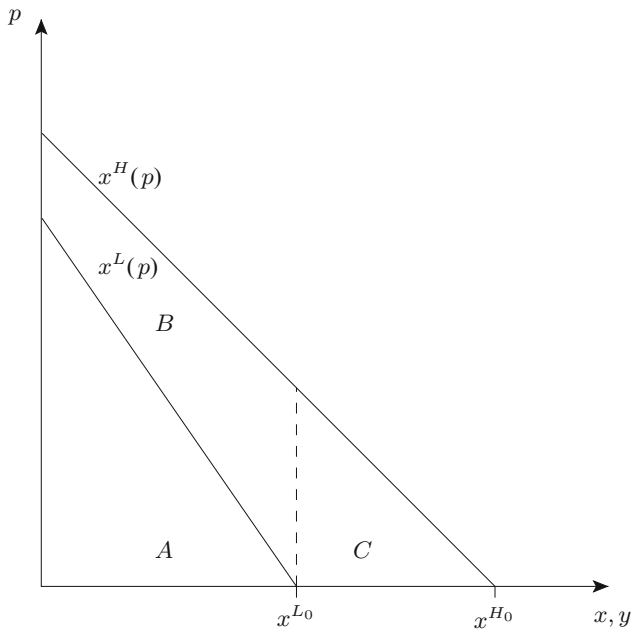

Fig. 10.3 Two types of customers, L (dashed line) and H (solid line)

number of pages the printer can print per minute. Depending on the specific context, it may be hard to distinguish between the quantity and the quality of a product. I will use the convention to talk about quantity in the following example.

In Fig. 10.3, one can see the inverse demand functions $p^H(x)$ and $p^L(x)$ of an individual with a high (H-type, solid line) and low (L-type, dashed line) willingness to pay. In order to simplify matters, assume that there are as many H-types as there are L-types and that there is only one individual of each type. The monopolist has marginal costs of zero, such that the efficient quantities supplied are equal to the maximum demand levels x_o^H and x_o^L. An individual's aggregate willingness to pay for a quantity x is equal to the area under the demand function, $P^H(x) = \int_0^x p^H(x)dx$, $P^L(x) = \int_0^x p^L(x)dx$. Her respective willingness to pay for the efficient quantity is, therefore, given by the areas $P^L(x_o^L) = A$ and $P^H(x_o^H) = A + B + C$.

In the preceding chapters, one has implicitly assumed that a firm sets a price per unit of the good and the customers choose how much to buy. For effective price discrimination with asymmetric information, the firm needs to restrict the customers' sovereignty by offering pre-specified quantity-price bundles. For example, $\{y, P^H(y)\}$ is a possible offer where the monopolist offers quantity y at the maximum price the H-type is willing to pay. An arbitrary pair $\{y, P\}$ is also called a "contract." P is the price for y units, not for one unit of the good, as it was before.

It is immediately clear that, with perfect information and, therefore, price discrimination, the monopolist will offer the efficient contracts, $\{x^L, P^L(x^L)\} = \{x_o^L, A\}$, $\{x^H, P^H(x^H)\} = \{x_o^H, A + B + C\}$. (In the following analysis, assume that a customer is willing to purchase a contract, if he is indifferent between buying and not buying, and that he is also willing to purchase the contract designed for him in case that he is indifferent between two contracts. This assumption simplifies the analysis and is without relevance for the qualitative results.) The monopolist's profit is then $2A + B + C$, and the consumer surplus is zero, $CS^H(x_o^H, P^H(x_o^H)) = CS^L(x_o^L, P^L(x_o^L)) = 0$.

However, with asymmetric information, these contracts are not enforceable. A H-type individual would prefer to buy the contract of the L-type, because it leads to a higher consumer surplus of B. The L-type would never buy the H-contract, because she has no willingness to pay for the additional quantity and, therefore, has no willingness to pay the higher price.

How will the firm react to this problem? In order to answer this question, one first needs to understand whether it is possible to change the contracts in a way that increases profits. If nobody buys the H-contract, because the H-types prefer the L-contract, then the firm's profits are $2A$. In order to induce the H-type to buy "his" contract, the firm can decrease the price of the H-contract until the H-type is indifferent between the two. Because his consumer surplus is B, this is achieved when the H-contract is $\{x_o^H, A + C\}$ (see Fig. 10.3). Because altering the contract in that way increases profits to $2A + C$, it is always profitable. The profit is smaller than it would be with perfect price discrimination, but larger than $2A$ is.

Is this the profit-maximizing pair of contracts? The answer is no, because the firm has another policy parameter that it can use to increase the effectiveness of price discrimination, namely the quantity of the product. If it is reduced in the L-contract, then both types' willingness to pay for this contract decreases. Therefore, the firm has to complement this change with a decrease in the price for this contract in order to be able to sell it. This seems like a bad idea, because it decreases the profit from the L-contract. The fact that a change of the contract in this direction can increase profits can be seen once one takes into account that the H-type has a higher marginal willingness to pay for additional quantity; he is willing to pay more for the last unit than the L-type is. This fact has the following consequence: the reduction of the quantity that is offered in the L-contract does not only decrease the profit from selling to the L-type, but can also be used to increase the price for the H-contract. The H-types' implicit "threat" to choose the L-contract becomes weaker given that $\{x^H, P^H(x^H)\} = \{x_o^H, A + C\}$. The L-contract becomes less attractive for both, but this effect is stronger for the H-type, whose willingness to pay for additional quality is higher than the L type's is. Thus, this quantity reduction can be used as an instrument for type selection. In the limit, as the L-contract's quantity goes to zero at a price of zero, it becomes possible to increase the H-contract's price to $A + B + C$ again.

Contracts that make the H-type indifferent between both contracts fulfill the so-called *self-selection constraint*. Figure 10.4 shows the possible and necessary price adjustments accompanying a change of the quantity of the L-contract from x_o^L to $x_o^L - dx_o^L$.

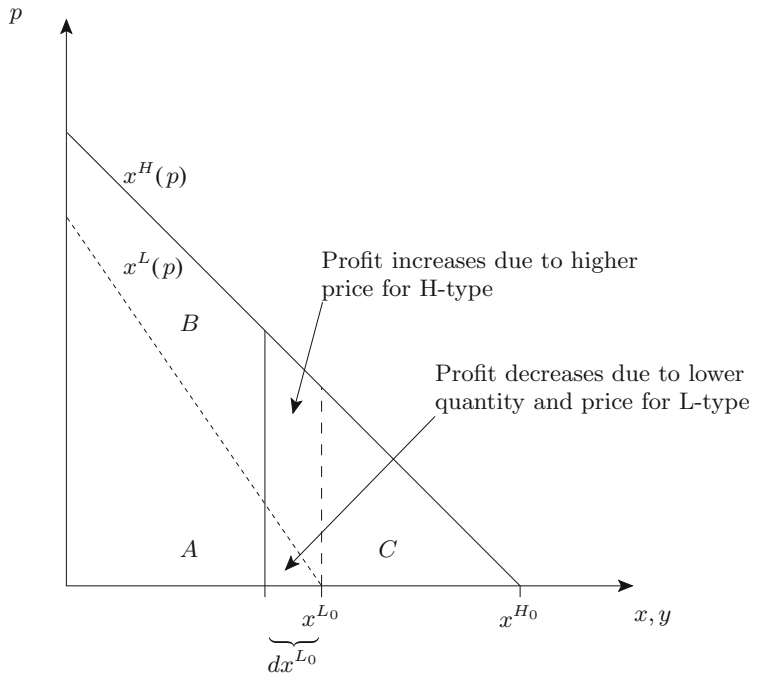

Fig. 10.4 Effect of a reduction in L-quantity

The adjustment of the contract stops when the marginal increase in profits, due to the increase in the price for the H-contract, equals the marginal loss in profits caused by the reduction of the price for the L-contract. Graphically, this means that the line segment between the H-type's and the L-type's demand functions have to be of equal length as the line segment between the L-type's demand function and the abscissa. This situation is depicted in Fig. 10.5.

(In Fig. 10.5, both contracts offer positive amounts of the good. This finding results from the assumption that both types have equal frequencies in the population and the specific demand functions. If there are either very few L-types, or their willingness to pay differs significantly from the H-type's, then it can be the case that the monopolist prefers not to sell to the L-types at all.)

The previous analysis has revealed some general characteristics of optimal price discrimination with asymmetric information.

- The H-type always consumes his optimal quantity, unlike the L-type. This property is also called *no distortion at the top*, because it is a general property of models with asymmetric information to not distort the allocation of the "best" type.
- With such contracts, the H-type always receives positive consumer surplus, if the L-quantity is positive. The L-type always gets a surplus of zero.

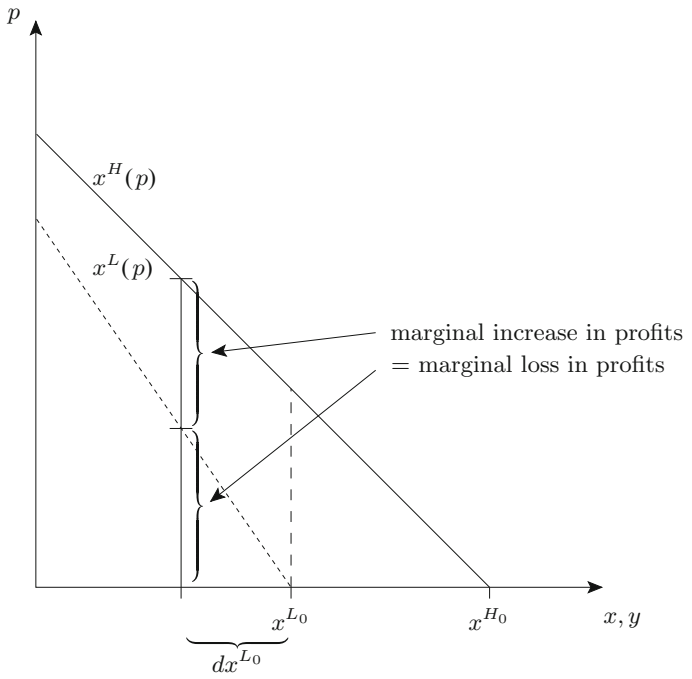

Fig. 10.5 In the optimum, the marginal decrease equals the marginal increase in profits

While these characteristics of optimal contracts sound quite abstract, they are very useful for understanding real-world pricing decisions. In the aforementioned example of airline pricing, business travelers usually have a higher willingness to pay for flexibility than leisure travelers do. In order to apply the model, one can interpret x as a variable that measures the flexibility of a ticket. The results of the model can then be interpreted in the following way: the airline should discriminate between the two groups of customers by offering economy and business class tickets. Business class tickets offer the optimal flexibility and comfort to business travelers, but economy tickets come with less flexibility and comfort than economy customers would like (and are willing to pay for). This reduced flexibility and comfort of economy tickets is the reason that business travelers choose 'their' higher-priced tickets, because they get a larger surplus from doing so. This can lead to strange incentives for airlines that may, for example, make economy seating purposely uncomfortable.

This logic can be applied to many other markets, for example ones in which customers can be grouped as "professional" and "private" users (software, computer hardware, ...). The strategy to play around with product quantity and quality in a way that makes sure that market segments are kept separate, in order to prevent demand spillover from high-priced segments to low-priced segments, is also called

price fencing. Price fences are very important for effective market segmentation and, therefore, for profit maximization. As in the above example, they are designed such that customers who can afford and are willing to pay higher prices are not tempted by the lower-priced versions.

Looking at the pricing of the MacBook Pro with Retina display (US prices, 2013), there were three versions of the computer, priced $1,299, $1,499 and $1,799. The only difference between the $1,299 and $1,499 versions was the RAM and flash capacity. Assume that one prefers the cheapest version, but would like to have more flash capacity. One option would be to customize the cheapest version by increasing the flash capacity. However, this option was not available. The only way to get more flash capacity was to buy the $1,499.00 version. The only option offered was to increase RAM from 4 GB to 8 GB for $100 more. It could be argued that the lack of such an option resulted from a technical problem, but it is not because the option to upgrade flash capacity was offered for the non-Retina version of the MacBook Pro. Hence, constraining options for customers must be seen as a purposeful strategy to "fence them in."

Sometimes, firms even have an incentive to incur costs to make products worse. Manufacturers of printers, for example, standardly equip their printers with soft- and hardware that is designed for fast printing speed and then equip a series of these printers with (costly) additional hard- and software to slow them down and sell them at a lower price. Thus, if a private user wonders at times why firms do not offer the perfect products for her needs, this may be the answer: from the point of view of a firm, the marginal costs of production are not the only ones it has to take into consideration. Additionally, there are opportunity costs that exist because customers from a different group may buy a version of the product that is not designed for them. These opportunity costs are relevant for the firm, but not for society, which is why there are externalities in the resulting equilibrium.

A related strategy for profit maximizing firms, which should briefly be mentioned, is called *bundling*. The underlying problem for a firm is to decide on the number of characteristics or features of a product. A car manufacturer can, for example, include features like driver alert, adaptive cruise control or other safety features in the standard package, or sell them separately. A flower shop can sell bouquets or separate flowers, a computer can come with word processor, spreadsheet and presentation program, or the soft- and hardware can be sold separately, and so on.

If one thinks of different features of a complex product as separate, simpler products, then the problem of bundling is to determine which products should be included in a bundle and which ones should be sold separately. There are several rationales for bundling. There could, for example, be complementarities between products, which is why shoes are usually sold as pairs. Alternatively, bundling might economize on costs, because bundles are more efficient to produce or distribute.

A subtler reason for bundling results from the fact that it allows the producer to skim off the willingness to pay in situations where it would otherwise be impossible

Table 10.1 An example for product bundling

Customer type	WP	SS	Sum
N	120	100	220
A	100	120	220

to do so. This is why it makes sense to discuss bundling in the context of asymmetric information and second-degree price discrimination.

Here is an example. Assume there are two types of customers in the market, who are interested in two different products: word-processing (WP) and spreadsheet (SS) software. One type of customer has a high willingness to pay for WP and a low willingness to pay for SS (say, a novelist, N), and the other has a high willingness to pay for SS and a low willingness to pay for WP (an accountant, A). Table 10.1 gives an example for the two types of customers' willingness to pay for the two different products. Type N is willing to pay up to CHF 120 for WP and CHF 100 for SS, and type A is willing to pay up to CHF 100 for WP and CHF 120 for SS. Hence, the total willingness to pay for both products is CHF 220 for both types.

Assume that the firm that sells the software knows that these two types of customers exist (one of each type), but that it cannot verify the identity of a customer when she buys the software. Assume further that the marginal costs of an additional software license are zero. Now, consider two pricing strategies for the firm: the unbundled and bundled selling of the two products.

What happens if the firm sells both products separately? I start with WP and denote its price by p^{WP}. If $p^{WP} \leq 100$, the firm can sell two licenses. If $p^{WP} \in (100, 120]$, it can sell one license and, for all prices $p^{WP} > 120$, the number of licenses sold drops to zero. Hence, the profit-maximizing price is $p^{WP} = 100$, yielding a profit of $\pi = 200$. The same calculation applies for SS. Therefore, total firm profits with unbundled selling are $\pi = 400$.

What happens if the firm decides to bundle the products? Denote the price of a bundle by p^B. Demand for each bundle is two, if $p^B \leq 220$, and drops to zero for higher prices. Hence, the profit-maximizing price for each bundle is equal to $p^B = 220$, yielding profits of $\pi = 440$.

Compared to the unbundled selling, bundling increases profits by $40 = 440 - 400$. What has happened? The underlying rationale is that differences in the willingness to pay "average out" by bundling. If the licenses are sold separately, the minimum willingness to pay becomes decisive in the example. This effect cancels out, if the two products are bundled and sold as a "package."

This result is robust and especially relevant for digital products that are produced at almost zero marginal costs: bundling large numbers of unrelated goods makes it easier to predict the customers' valuations for a bundle than their valuation for an individual good does when it is sold separately. This "predictive value of bundling" makes it possible to increase sales and profits. Examples are cable television, an internet site's content (e.g. the New York Times), or copyrighted music (for example Spotify).

To conclude this subchapter, it is important to note that the above findings also have implications for the optimal organization of firms. As seen, there are different contracts for different groups of customers. The two dimensions of the optimal contracts, which are price and – depending on the specific interpretation – quantity or quality, are not independent from each other, but can only be understood in combination. Hence, to take these important interdependencies into account, the responsibility for the different customer groups should not be given to different, independent product managers whose responsibility is to maximize profits for their departments (profit centers). This system ignores the fact that modifications of the contracts cause externalities in the other departments thereby leads to a situation where each manager maximizes the profits of his profit center, but not the total firm profits.

10.5.3 Third-Degree Price Discrimination

The last case is third-degree price discrimination. This variant is characterized by the firm's ability to discriminate between different segments of customers, but not within each segment. A prominent example is price discrimination between national markets, which is often practiced by internationally operating firms. Especially the pharmaceutical industry repeatedly makes it into the headlines for selling the same active ingredients at higher prices in Switzerland than in, for example, the European Union. However, prices for ordinary consumption goods are also discriminated in this way and Swiss customers quite often pay more for a good than others do. Apple Inc., for example, makes extensive use of international price discrimination. The average price for a song in the iTunes store was $1.29 in the US and CHF 2.30 in Switzerland in 2011 (the average exchange rate in 2011 was about US$ 1.10 per 1 CHF). However, there are many other forms of third-degree price discrimination, for example according to age group or status (student or senior discounts, discounts for military members in the US) or according to gender ("Ladies' night" in nightclubs or at the dry cleaners, which typically charge higher prices for women's clothes).

Assume that a firm produces a given product at a given production facility (say in China) and sells it to two countries, Switzerland (country 1) and France (country 2). The respective quantities are y^1 and y^2, and the production and logistics costs depend on the total quantity produced, $C(y^1 + y^2)$. The market-research department estimates the demand functions in the two countries as $P^1(y^1)$ and $P^2(y^2)$. Consequently, total profits are given by:

$$\pi(y^1, y^2) = P^1(y^1) \cdot y^1 + P^2(y^2) \cdot y^2 - C(y^1 + y^2).$$

From the manager's point of view, the problem is to choose the quantities supplied to the different markets in order to maximize profits. The optimal decision is char-

acterized by the following necessary conditions:

$$\frac{\partial \pi(y^1, y^2)}{\partial y^1} = \frac{\partial P^1(y^1)}{\partial y^1} \cdot y^1 + P^1(y^1) \cdot 1 - \frac{\partial C(y^1 + y^2)}{\partial y^1} = 0,$$
$$\frac{\partial \pi(y^1, y^2)}{\partial y^2} = \frac{\partial P^2(y^2)}{\partial y^2} \cdot y^2 + P^2(y^2) \cdot 1 - \frac{\partial C(y^1 + y^2)}{\partial y^2} = 0.$$

If one looks at the two conditions in isolation, the result is not very surprising: as in the model without price discrimination, the firm chooses the quantity that equalizes marginal revenues with marginal costs for each market. Only if one takes into account the fact that marginal costs are identical irrespective of the market where the products are sold (production takes place in the same factory) can one learn something new. Then one can establish the following relationship between the two markets:

$$\frac{\partial P^1(y^1)}{\partial y^1} \cdot y^1 + P^1(y^1) = \frac{\partial P^2(y^2)}{\partial y^2} \cdot y^2 + P^2(y^2).$$

Thus far, the above condition only states that marginal revenues are equal in both markets. However, rewriting the equation to transform it into the rule of thumb that was developed before, one gets:

$$p^1 \left(\frac{1}{\epsilon_{p^1}^{y^1}(p^1)} + 1 \right) = p^2 \left(\frac{1}{\epsilon_{p^2}^{y^2}(p^2)} + 1 \right).$$

Further assuming that demand is falling with respect to price in both markets (which implies that the elasticities are negative), one ends up with:

$$p^1 \left(1 - \frac{1}{|\epsilon_{p^1}^{y^1}(p^1)|} \right) = p^2 \left(1 - \frac{1}{|\epsilon_{p^2}^{y^2}(p^2)|} \right).$$

In order to understand the economic reasoning underlying this condition, assume that the price elasticity in market 1 (Switzerland) is lower than the price elasticity in market 2 (France), $|\epsilon_{p^1}^{y^1}(p^1)| < |\epsilon_{p^2}^{y^2}(p^2)|$ (Swiss demand is less elastic). This implies that the expression in brackets is smaller in market 1 than in market 2. Hence, the condition can only be fulfilled if $p^1 > p^2$: the good is sold at a lower price in the market with the higher price elasticity.

To further illustrate this condition, assume that demand in both markets is linear, $p^i(y^i) = a^i - b^i \cdot y^i, i = 1, 2$, and that marginal costs are constant and equal to $c > 0$. In this case, one knows from above that:

$$y^{i*} = \frac{a^i - c}{2 \cdot b^i}, \quad p^{i*} = \frac{a^i + c}{2}, \quad \epsilon_{p^i}^{y^i}(p^{i*}) = \frac{a^i + c}{a^i - c}, \quad i = 1, 2.$$

Comparing the elasticities between both markets reveals that:

$$\epsilon_{p^1}^{y^1}(p^{1*}) > \epsilon_{p^2}^{y^2}(p^{2*}) \Leftrightarrow \frac{a^1 + c}{a^1 - c} > \frac{a^2 + c}{a^2 - c} \Leftrightarrow a^1 > a^2.$$

However, this is the case if and only if $p^{1*} = \frac{a^1+c}{2} > \frac{a^2+c}{2} = p^{2*}$: the price on the less elastic (in equilibrium) market is higher.

This result gives an important hint as to why prices in Switzerland are generally higher than abroad. The willingness to pay (as reflected by a low price elasticity) is higher in Switzerland than elsewhere, which implies that firms sell their products at higher prices. One can use this theoretical result to test the theory empirically. All one needs to do is to estimate the price elasticities in different markets and compare them with prices. If the hypothesis of the model cannot be rejected, then one has a valid explanation for an important empirical phenomenon.

Without going into the analytical details, it is time to contemplate the consequences of a regulation that forbids price discrimination between markets. Such a regulation might, for example, prevent price discrimination directly, or it might allow the emergence of resale markets that make profit out of price arbitrage between markets. Such a "single-price philosophy" can, for example, be found in the European Union with its "Single European Market," which is enforced by the European Commission.

Taking the theoretical results from above as a point of departure, the monopolist, who is no longer in a position to discriminate prices, needs to determine the new, aggregate demand function for the joint market. This new demand function results from adding up the individual market demand functions, $X(p) = x^1(p) + x^2(p)$, and it follows that the new inverse demand function is $P(x) = X^{-1}(x)$. The resulting problem is equivalent to the problem of a monopolist who cannot discriminate prices. Hence, even without a formal analysis of the situation, one can determine the differences between the new and the old situation. There are three different constellations possible:

1. The demand structure is similar in both countries and markets are of approximately the same size. The monopolist will sell to customers in both markets and the new price will be in between the prices that would be charged with price discrimination. The redistributive consequences between the customers in both markets are easy to determine: customers in the previously high-price market profit, because they can buy at lower prices, and more customers buy, while those in the previously low-price market lose, because they are paying higher prices and fewer customers buy.
2. The country with less elastic demand is relatively large or the price difference between countries is large, or both. In this case, it can happen that the monopolist will not serve the smaller market anymore. The reason is that, in order to sell to the smaller market, he must lower the price in the larger market to an extent that makes it rational to not serve the more elastic, or smaller market, at all. The effect is that nothing changes for the large country, but the situation in the

smaller market deteriorates, because its customers are excluded from the consumption of the product. Therefore, prohibiting price discrimination in different markets may lead to inefficiency and exclude customers from consumption.

3. The country with higher prices is relatively small or the price difference is relatively small, or both. In this case, the price in the more expensive country might decrease almost to the level of the other country. There is no change for the formerly low-price country, whereas customers in the formerly high-price country benefit.

This qualitative analysis shows that no clear prediction about the consequences of market integration can be made without further information about the relative size of the markets and the relative willingness to pay in each market. Only if this kind of information is available, is it possible to make a reliable prognosis about the effects of such a policy change.

£ **Digression 29. Parallel Imports in the European Union (EU)**

Firms that sell in different markets have an incentive to discriminate prices according to the market-specific elasticities of demand. However, the creation of a common market within the EU has made it possible for parallel imports to move freely across the EU. Parallel imports are sales by authorized or unauthorized distributors to another country without the permission of the initial property owner.

Parallel importers use price differences between markets to make a profit out of price arbitrage. This puts pressure on high prices and, thereby, creates a tendency towards uniform prices within a common market. The only industry-specific exemption from the general competition principles is the automotive industry. The purpose of the so-called "block exemptions" is to restrict competition between car dealers. Nevertheless, even with these special agreements, EU rules require the car dealers to sell their products to any EU citizen regardless of where they live.

This regulation is, of course, a thorn in the side of car manufacturers, who try to find ways to limit competition due to parallel imports. The European Commission fined Volkswagen an amount of €102 million (later reduced to €90 million) for preventing Austrian and German customers from buying cars in Italy. It also fined PSA Peugeot Citroen €49.5 million. Peugeot Netherlands tried, for example, to incentivize its franchise dealers to restrict sales to other countries by withholding bonus payments and limiting the supply of Peugeot cars.

10.6 Monopolistic Competition

I have argued that the models of monopolistic behavior apply to situations where a firm faces a demand function that is not perfectly elastic. This situation allows the firm to charge prices that are above marginal costs. The associated producer surplus is higher than in a situation of perfect competition. This has two consequences.

First, if production involves fixed costs, the firm can stay profitable in situations where firms in competitive markets would have to leave the market, as long as prices are above average costs. Such a situation is illustrated in Fig. 10.6, where it is assumed that the monopolist can not discriminate prices.

In this figure, average costs are above marginal costs, but below the price of a non-price-discriminating monopolist. This leads to a situation where a competitive firm makes a loss equal to area A, whereas a monopolist makes a profit that is equal to area B.

Second, positive profits in a monopolistic market, as in Fig. 10.6, make it attractive for other firms to develop similar products. Even if these firms are legally or otherwise prevented from simply imitating the profitable product, they can try to develop and sell similar ones. Such a situation is called *monopolistic competition* and there are several examples for such industries:

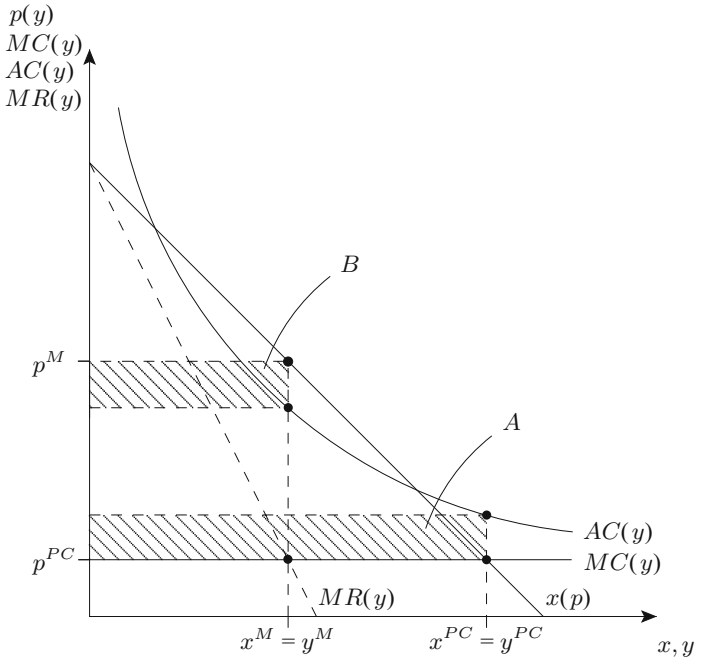

Fig. 10.6 Profits in monopolistic and perfectly competitive markets with fixed costs

- Cars of a given type from different manufacturers (like SUVs from Audi, Mercedes, BMW, Volkswagen, etc.).
- Books or music that are variations of the same topic (romantic novels or textbooks in economics) or style (Jazz, Pop, Classical Music).
- Smartphones, tablets or notebooks from different manufacturers.
- Pubs and restaurants in a city.

The above list illustrates that monopolistic competition is a very prominent market structure, especially in an economy where brands are important for customers (which is the same as saying that they are willing to pay for a specific brand). This is why it makes sense to understand the functioning of this type of a market. The different varieties of similar goods are called *differentiated products*.

The main question is about the number of similar products that exist in such an industry. When one compares the total number of different SUVs with the total number of different romantic novels, one sees that there are huge differences. Are there any patterns that allow one to explain why some industries produce a relatively small number of variants, whereas others produce far more?

The basic idea for answering this question is to blend the analysis of a single monopoly with the idea of market entry: assume there is free market entry and exit and that a monopolist makes a profit with a product, say an SUV from a given manufacturer A. Profits exist, if the price exceeds the average costs, $p^A > AC^A(y^A)$.

These profits encourage another firm, B, to enter the market and to sell a similar product. The availability of this additional product increases the choice of the customers. They still consider the products to be different, but the existence of another model in the SUV market makes the first one less exclusive and manufacturers A and B have to somehow share the market. The effect is that the demand for SUV A is likely to shift leftward and to become more price elastic, which reduces profits. With free entry and exit, additional firms will enter the market as long as profits are still positive.

By the same token, if the number of different products is so large that the firms (for example i) are making losses, $p^i < AC^i(y^i)$, some of them will have to leave the market. The effect is that the number of products from which the customers can choose decreases. This effect likely shifts the demand rightward and makes it less price elastic.

The long-run equilibrium must, therefore, be a situation where the prices of the products equal average costs, $p^i = AC^i(y^i)$, because at this point firms make zero profits. This situation is illustrated in Fig. 10.7.

This is the situation where no further competitor is willing to sell another similar product and no existing competitor is willing to leave the market. The figure reveals two properties of the long-run equilibrium in such a market.

First, if firms can develop similar products, there is no escape from the zero-profit equilibrium in the long run. As in the case of perfect competition, competitors will react to positive profits by entering the market with differentiated products.

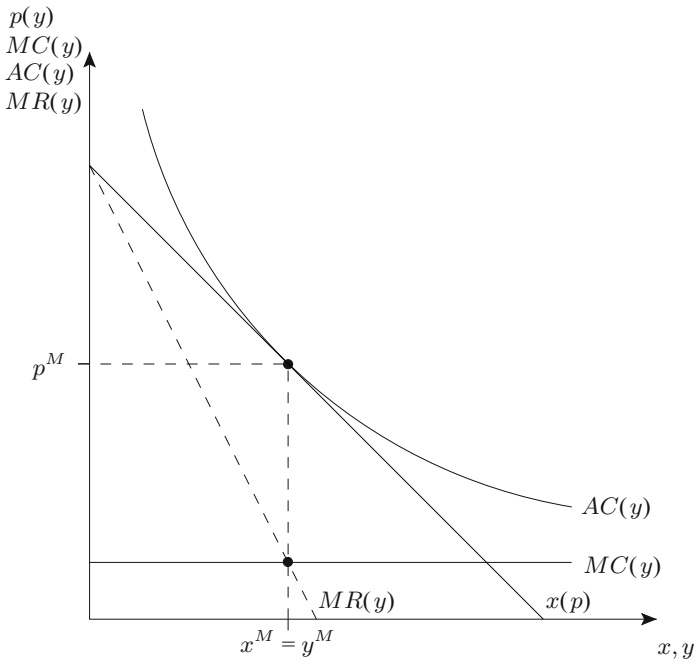

Fig. 10.7 Long-run equilibrium in a market with monopolistic competition with free entry and exit

Second, even though long-run profits are zero, the resulting allocation is not Pareto-efficient. The single firm is still facing a downward-sloping demand function for its product, so prices in this industry will be above marginal costs.

In addition, one can say something about the number of differentiated products that can survive in such an industry. Figure 10.6 shows a situation where the monopolist has profits π_M. Profits depend on the relationship between the demand and the average-cost curve. The bigger the difference is, the higher the profits are. Now, assume that this firm is the first one to sell a new type of product (the first SUV), such that Fig. 10.6 refers to a situation where no other firm has entered the market with a similar product.

Profits depend on fixed costs. An increase in fixed costs shifts the average-cost curve upwards, which implies that profits decline. In the extreme case, profits are equal to zero without any competition from differentiated products. If a second firm enters the market with a similar product, any leftward shift of the demand function that is caused by market entry implies losses for the firm. If profits had been positive but relatively small, the shift in the demand function reduces them, but not necessarily to zero, which would imply that there is room for a second firm selling a similar product. The pattern that becomes visible here is a general one:

the more fixed-cost-intensive the production of the product is, the fewer firms can succeed in the market; the number of differentiated products is inversely related to the fixed costs of production.

The above argument was pretty loose, so it makes sense to develop it formally. In order to do so, assume that there are n differentiated products in a market and that demand for a single product is given by the demand function $y^i = Y(1/n - b(p^i - \bar{p}))$, where $\bar{p} = \sum_{j=1}^{n} p^j/n$ is the average price level and $Y = \sum_{j=1}^{n} y^j$ is the total output in the industry. Here, b represents the responsiveness of a firm's output to its price. The demand function implies that the n different firms share the market equally, if they charge identical prices: $p^i = \bar{p}$ implies $y^i = Y/n$. When we solve the model we will see that this function implies that individual and market demand are not absolutely fixed so that we have a degree of freedom in determining the equilibrium. Assume further that all firms produce with identical cost functions $C(y^i) = c \cdot y^i + FC$ and maximize profits. These assumptions may not be particularly realistic, but they simplify the analysis considerably without changing the qualitative insights.

From the point of view of a single firm, the inverse demand function is given as $P(y^i) = 1/(b \cdot n) + \bar{p} - y^i/(b \cdot Y)$, which leads to profits as a function of output:

$$\pi(y^i) = P(y^i) \cdot y^i - c \cdot y^i - FC = \left(\frac{1}{b \cdot n} + \bar{p} - \frac{y^i}{b \cdot Y} \right) \cdot y^i - c \cdot y^i - FC.$$

The profit-maximizing output of firm i is, again, characterized by the first-order condition of the profit function. (In general, Y and $\bar{p}$ are functions of y^i. Assume that the firm neglects this effect.) If one does the math, one ends up with the following price and output of product variant i:

$$y^{i*} = \frac{bY\,(1/(b \cdot n) + \bar{p} - c)}{2}, \quad p^{i*} = \frac{1/(b \cdot n) + \bar{p} + c}{2}.$$

The solution has intuitive economic properties: the profit-maximizing output and price of variant i is decreasing as the number of variants n increases. This property illustrates the effect of competition on the market for product variant i: the larger the number of similar products, the fewer the number of products of a given variant that can be sold and the lower the price level for this variant.

What is even more interesting is if one can say something about the number of differentiated products that can be supplied in this market. In order to gain insight into this question, assume that all firms charge equal prices in equilibrium, $p^{i*} = \bar{p} = p^*$. If one uses this assumption, one can solve for the equilibrium price level in this industry for a given number of differentiated products n:

$$p^* = \frac{1/(b \cdot n) + p^* + c}{2} \quad \Leftrightarrow \quad p^* = \frac{1}{b \cdot n} + c.$$

This finding nicely illustrates that the markup rule still applies: the firm is able to sell the variant above marginal costs c, but the markup is the lower, the higher the competitive pressure is that results from the number of similar products n.

If we use the information that all prices in the industry are the same, we get back to the above-mentioned property that individual and market demand are not fixed in absolute terms such that we have a degree of freedom that we can use to fix one of these variables and solve the rest of the model relative to it. We use $y^{i*} = 1$ as normalization because it is easy to solve, but any other convention would work as well:

$$y^{i*} = \frac{bY\,(1/(b \cdot n) + 1/(b \cdot n) + c - c)}{2} = \frac{Y}{n} = 1.$$

Hence, the total supply of variant i is equal to 1, and industry output is, therefore, equal to n, the number of firms in the industry. One last step is missing to determine how many differentiated products exist in the long run. Given the equilibrium outputs and prices, profits of an arbitrary firm i are equal to

$$\pi^{i*} = \left(\frac{1}{b \cdot n} + c\right) \cdot 1 - c \cdot 1 - FC.$$

This equation, again, shows the effects of competition: equilibrium profits are decreasing with an increasing number of differentiated products.

One knows that free entry and exit into this industry drive profits down to zero,

$$\left(\frac{1}{b \cdot n} + c\right) \cdot 1 - c \cdot 1 - FC = 0.$$

This information can be used to finally determine the long-run number of differentiated products:

$$\left(\frac{1}{b \cdot n} + c\right) \cdot 1 - c \cdot 1 - FC = 0 \quad \Leftrightarrow \quad n = \frac{1}{b \cdot FC}.$$

This result confirms the intuitive conclusion that I have discussed before: there is a negative relationship between the fixed costs of an industry and the number of differentiated products that exist in a long-run equilibrium. This result sheds light on the question of why there are fewer SUVs than romantic novels on the market: product categories differ with respect to fixed costs. Writing a beach novel is far less expensive than developing a new car. (Both, the costs of writing the novel and the development costs, are part of the technological fixed costs, because they occur independently of the number of copies or cars sold.)

The result also contains an important message for managers in an industry with differentiated products, because profits for a given number of competitors and the long-run number of competitors allows estimations as to whether market entry is still profitable and how many variants a product can survive in the long run.

$\mathcal{B}$

References

Ariely, D. (2008). *Predictably Irrational*. Harper Collins.
Boublil, A., & Michel-Schonberg, C. (2013). *Les Misérables*. Wise Publication.
Calo, R. (2013). Digital Market Manipulation. *George Washington Law Review* 2013(27), 995–1051.
Executive Office President of the United States (2015). Big Data and Differential Pricing. Report.
Hannak, A. (2014). Measuring Price Discrimination and Steering on E-Commerce Web Sites. In *Proceedings of the 2014 Conference on Internet Measurement Conference*.
Marn, M., Roegner, E., & Zawada, C. (2003). Pricing New Products. *The McKinsey Quartely*, *3*, 40–49.
Monroe, K., & Cox, J. (2001). Pricing Practices that Endanger Profits. *Marketing Management*, *10*(3), 42–46.
Poundstone, W. (2011). *Priceless: The Myth of Fair Value (and How to Take Advantage of It)*. Simon and Schuster.
Puzo, M. (1969). *The Godfather*. Signet.
Shiller, B. (2014). *First-Degree Price Discrimination Using Big Data*. Working Paper.

Further Reading

Allen, W. B. (2009). *Managerial Economics Theory, Applications, and Cases*. Norton.
Frank, R. H. (2008). *Microeconomics and Behavior*. McGraw-Hill.
Haisley, E., Mostafa, R., & Loewenstein, G. (2008). Subjective Relative Income and Lottery Ticket Purchases. *Journal of Behavioral Decision Making*, *21*, 283–295.
Pindyck, R., & Rubinfeld, D. (2001). *Microeconomics*. Prentice-Hall.
Tirole, J. (1988). *The Theory of Industrial Organization*. MIT Press.
Varian, H. R. (1992). *Microeconomic Analysis*. Norton.