The Economics of Competition (Law)

Dr. Johannes Paha

Justus-Liebig-University Giessen

Version 3.0 – April 8, 2013

Johannes Paha is a research associate at the

Chair for Industrial Organization, Regulation and Antitrust (VWL 1)
Justus-Liebig-University Giessen
Licher Straße 62
D-35394 Giessen

email johannes.paha@wirtschaft.uni-giessen.de
phone +49 – 641 – 99 22052
fax +49 – 641 – 99 22059
web http://wiwi.uni-giessen.de/ma/dat/goetz/Johannes_Paha%2C%20M.A./

This reader is work-in-progress. If you detect orthographical, grammatical, or other mistakes, please mark them in this file and send it to johannes.paha@wirtschaft.uni-giessen.de. A freeware for making such comments in pdf-files is, e.g., PDF-XChange Viewer (http://www.tracker-software.com/product/pdf-xchange-viewer).

I would especially like to thank

Tim Brühn and Daniel Herold

for their valuable comments on this reader and the accompanying lecture.

TABLE OF CONTENTS

INTRODUCTION

This introduction is designed to provide you an overview on the topics covered in this reader and the accompanying lecture. Moreover, you will find reasons why it is worthwhile to study the subject of industrial organization and learn more about competition policy.

What is industrial organization?

The subject of industrial organization (IO) is a branch in economics that is concerned with analyzing firm behavior and market outcomes. This is particularly interesting in markets which are not characterized by perfect competition, such as monopolies or oligopolies. Especially the latter are subject to the following examples of strategic interaction among firms:

- What price should a firm optimally set and what quantity supply, given that its competitors make a similar reasoning? What determines the extent of **competition**? What determines the dimensions upon which firms compete? What is the effect of the number of firms on price-cost margins?
- Can firms increase their profits by **coordinating** their market behavior? Should they trust their co-conspirators? How can firms *attain* supracompetitive profits? Is it allowed to attain such a position? If yes, how can firms maintain such increased profits?
- What determines **market structure**? What determines the ease of entry and the response of incumbents to the appearance of entrants? What determines asymmetries among firms?
- What is meant by **market power** or dominance? What determines the creation and sustainability of dominance? Is there increasing or decreasing dominance? Is it good or bad for welfare that an industry is dominated by a few firms?

These questions illustrate that economics (including industrial organization) is not about finding equilibria or the intersection of two curves. Economics is about trying to understand certain phenomena and is defined by the set of questions, not a set of methods. Methods can change but the questions do not.

"One of the first steps in studying industrial organization is to have an idea of what types of industry structure there may be" (Cabral 2000: p. 69). In section A , we start with defining the most extreme and best-known cases with many firms (perfect competition) and a single firm (monopoly). In section B , we proceed to the somewhat more advanced case of industries with few firms (oligopoly). We find that firms' competitive conduct as well as the market outcome depend on the production technology, for example the existence of capacity constraints. Competition is supposed to be intense when every firm in a market would be able to serve the entire demand (Bertrand-

competition). However, competition is less intense when capacity-constrained firms can serve demand only if several firms operate in this market.

We show that these different market structures (monopoly, oligopoly, and perfect competition) differ in their levels of allocative efficiency and welfare. Section C extends this analysis to productive and dynamic efficiency. We find that the characteristics of the production technology used affect the structure of a market. For example, an industry will be characterized by fewer firms if a high proportion of fixed costs of production implies a large minimum efficient scale of production. Moreover, the competitive conditions in an industry affect firms' incentive to innovate, i.e. to create better products or reduce the costs of production.

What is competition policy?

In practice, industrial organization has had an ongoing effect on competition policy. A passionate statement why competition and, thus, **competition policy is important** for everyone was provided by the European competition commissioner Joaquín Almunia in a speech in February 2011[1]:

> *"Ladies and Gentlemen:*
>
> *Competition is an instrument, not an end in itself. But it is indeed a vital instrument in very many respects. Without fair, robust, and effective competition policy and enforcement, I don't see how we Europeans can overcome the crisis rapidly and shape up to compete with the other, dynamic players that are increasingly present on the world scene. Of course, competition is not the only tool we should use to pursue this goal. But we need a vibrant and competitive environment in the single market if we are serious about leading in the information age.*
>
> *We need competition to be equal partners with the US, China, and the other leading global players; we need competition to grow; we need competition to preserve our social model for the benefit of our citizens and of the future generations. Considering our demographic trends and the imperative task of building sustainable and green economic and social models, Europe needs all its resources and resourcefulness.*
>
> *The EU competition system is one of the best, if not the best in the world. My commitment is to use it to the full extent of the law, because I am convinced that this is what I must do within my area of responsibility to contribute to a better future for Europe.*
>
> *Thank you."*

1 http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/11/96&format=HTML&aged=0&language=EN&guiLanguage=en

In this lecture, you will learn why competition is so important for our well-being, why dominant firms can reduce welfare, and what harm is caused by collusive agreements. In short, you will learn the details underlying Mr. Almunia's above conclusions. Moreover, you will learn how measures of competition policy are used to raise welfare. We present an introduction to competition policy in the European Union supplemented by remarks on competition policy in Germany and USA (chapter D). We follow a dual approach, i.e. presenting (European) competition laws with an emphasis being laid on their economic justification and consequences.

In sections A and B we propose that prices charged and quantities sold in a market depend on the structure of an industry and the competitive conduct of the firms. In section E , we elaborate on the relationship between market structure and firms' conduct with regard to the existence of market power. Competition policy defines market power as “the ability to profitably maintain prices above competitive levels for a period of time or to profitably maintain output in terms of product quantities, product quality and variety or innovation below competitive levels for a period of time” (EC 2011: para. 39). A related definition considers market power to be a firm's ability to profitably charge prices above marginal costs. Therefore, we present measures for assessing both the structure of an industry and the existence of market power. Additionally, we present answers to the question: “How can market power persist in an industry?” With respect to firm-behavior we identify the actual or potential entry of competitors into a market as one factor that erodes market power. With respect to consumer-behavior we propose that buyer power can countervail firms' market power. Moreover, we identify switching costs and network effects as elements which allow firms to exercise market power on their customers.

Many of the concepts for measuring market power or industry structure (e.g. market shares, the HHI, or the analysis of substitution patterns) crucially depend on the delineation of the relevant market. Chapter F presents some common concepts for defining the relevant market before we proceed to the four key concerns in competition policy.

Chapter G is concerned with the economics of merger control in the EU. In this context, we deal with the assessment of the pro- and anti-competitive effects of both horizontal and non-horizontal mergers. The earlier refer to a merger of competitors while the latter can be vertical mergers (for example, a manufacturer and a retailer merger) or conglomerate mergers (the merging firms appear to be unrelated with regard to the production and distribution of their goods). Chapter H adds to the discussion of agreements between undertakings by focusing on the assessment of horizontal co-operation agreements and a detailed analysis of vertical restraints. SectionError: Reference source not found is concerned with the competitive conduct of firms who consider the impact of their current decisions on market performance in the future. These dynamic aspects are

relevant in the analysis of horizontal agreements among firms. Such agreements typically aim at an increase in prices which raises firms' profits and reduces consumer surplus. Interestingly, such agreements need not necessarily be explicit and, thus, illegal. They may also consist of an implicit understanding among firms to raise prices (so-called tacit collusion). Tacit collusion can arise in an industry when firms may effectively prevent freeriding behavior of any of the participating firms and when the potential deviators pay sufficient attention to this punishment.

In chapter I we present economic principles to be applied in abuse of dominance cases. In all these areas of competition policy two-sided markets have attracted increasing attention recently. Therefore, chapter J describes some economic principles for the assessment of market power when a firm provides services to two types of agents who benefit from the network effects created by the platform. Fourth, chapter K concludes with a brief outlook on the assessment of state aid.

Why should you study industrial organization and competition policy?

Put plainly, the analysis of strategic decision making – as is done in industrial organization – and a profound knowledge of the economics of management are important for anyone who wants to become a successful business manager. In addition to this target group, the demand for economists in the consulting business has increased over the last years. Some important consulting firms are ESMT[2], Frontier Economics[3], NERA[4], Oxera[5], and RBB[6]. Additionally, industrial economists have increasingly been employed by competition authorities such as the Bundeskartellamt[7] and the Directorate General Competition at the European Commission.[8] The service supplied by economists is sometimes called *forensic economics* (Connor 2008: p. 31). For example, economists serve as expert witnesses in competition cases and advise either the judges, a competition authority, the claimant, or the defendant. Moreover, they assist in designing competition laws against welfare-detrimental mergers, cartels, the abuse of a dominant position, or state aid. The scope of their work goes beyond purely theoretical analyses and – to a considerable extent – includes empirical work.

2 http://www.esmt.org/eng/consulting/esmt-competition-analysis/

3 http://www.frontier-economics.com/

4 http://www.nera.com/

5 http://www.oxera.com/

6 http://www.rbbecon.com/

7 http://www.bundeskartellamt.de/

8 http://ec.europa.eu/competition/index_en.html

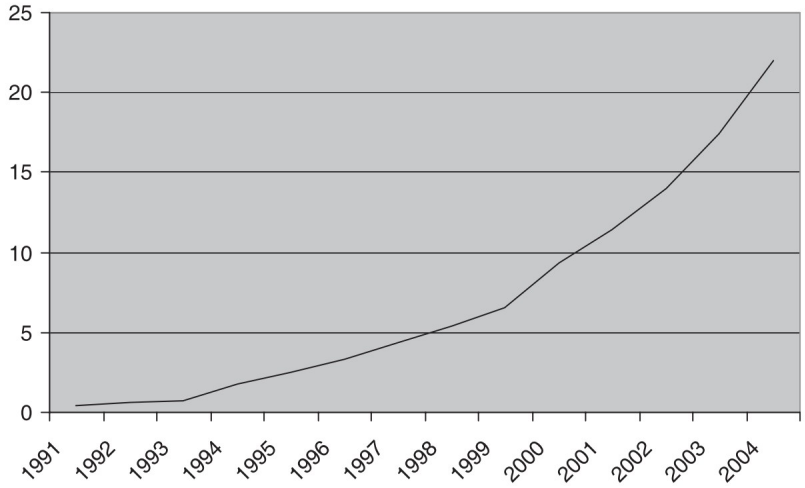


Figure 1: Turnover of economic consultancy firms (in £ million, base year: 2004)

Source: Neven (2006: p. 748)

Neven (2006: p. 748) shows (see Figure 1) that the turnover of economic consultancy firms has drastically increased between 1992 and 2004. Since 2004 competition cases have gained further importance. In Europe, fines imposed on cartels rose from a total of 3.2bn EUR in 2000-2004 to 8.9bn EUR in 2005-2009. The number of cartel cases decided rose from 11 in 1990-1994 to 33 in 2005-2009.[9] For these reasons, one may expect a stable or even increasing demand for competition economists over the following years. In 2010, the Antitrust Division of the US-American Department of Justice filed 60 criminal cases and obtained $555m in fines (Shapiro 2010: p. 1). The evolution of these numbers from 2006-2010 is shown in Table 1.

	2006	**2007**	**2008**	**2009**	**2010**
Total cases filed	34	40	54	72	60
Defendants charged	61	57	84	87	84
Fines obtained	USD 473m	USD 630m	USD 701m	USD 1,007m	USD 555m
Total jail days obtained	5,383	31,391	14,331	25,396	26,046

Table 1: US Department of Justice - Antitrust Division Criminal Enforcement Data

Source: Shapiro (2010: p. 2)

To illustrate the ongoing **demand for forensic economists**, consider that the first known cartel case was reported in Athens 326 BC (Connor 2008: p. 32). Due to military disturbances the

9 DG Comp cartel-statistics as of 9 November 2010: http://ec.europa.eu/competition/cartels/statistics/statistics.pdf

import-price of grain had been extremely volatile. Therefore, grain dealers formed a collusive agreement in form of a bidding ring. They would not have overbid each other in purchasing grain which harmed the sellers of grain. Moreover, the grain dealers restricted sales of grain to the Athenian people in periods of scarcity. This caused an increase in sales prices and harmed the buyers of grain. As a consequence, the dealers' profits increased by 500 percent. The court judgment in this case reflects two aspects that are still important today. First, the grain dealers should not only be punished for the infringement of competition laws. Second, others should also be deterred from breaking those laws in the future.

References

Connor, J.M. (2008). “Forensic Economics: An Introduction with Special Emphasis on Price Fixing.” Journal of Competition Law and Economics. Vol. 4 No. 1, pp. 31-59

Harrington, J. (2011). “Lectures on Industrial Organization.” Unpublished lecture notes

Neven, D.J. (2006). “Competition economics and antitrust in Europe.” Economic Policy. Vol. 21 No. 48, pp. 741-791

Shapiro, C. (2010). “Update from the Antitrust Division.” US Department of Justice. http://www.justice.gov/atr/public/speeches/264295.pdf

A COMPETITION IN STATIC INDUSTRIES: PERFECT COMPETITION AND MONOPOLY

The welfare that is generated by a market depends on the structure of the market. In particular, the number of firms in a market is an important determinant for the existence of market power and, thus, welfare. Subsection A.1 is concerned with defining welfare and allocative efficiency. Moreover, we see that allocative efficiency is achieved when a market can be described by a model of perfect competition (see subsection A.2).

Models are simplified descriptions of reality, i.e. they provide a means for understanding a particular situation or event. We can use models to predict how the market outcome (e.g. welfare) changes in response to changes in, for example, market structure or firms' behavior. Comparing a situation after a change in these variables to the situation prior to this change is called **comparative statics**. The word *statics* implies that we are not predicting the dynamic path that takes us from one equilibrium to the other. In order to see how welfare responds to changes in market-structure we start with analyzing a model for many firms (perfect competition). The subsequent sections are concerned with analyzing welfare when a market is characterized by only one firm (monopoly – see subsection A.3) or a few firms (oligopoly – see section B).

A.1 Welfare and Allocative Efficiency

Consumer surplus is the difference between the maximum amount a consumer is willing to pay for a unit of a good and the amount actually paid for that unit p_0. Let the demand function of a good be given by equation (1).

$$q_D = D(p) = q(p) \tag{1}$$

Consumers' willingness to pay $p(q)$ for some quantity q of the good is defined by the inverse demand function (2).

$$p_D = D^{-1}(q) = p(q) \quad . \tag{2}$$

The consumer surplus *CS* is defined as the area between the demand curve and the ordinate in Figure 2, evaluated in the interval between the price paid p_0 and the maximum willingness to pay $\bar{p}$ of the consumers.

$$CS(p_0, \bar{p}) = \int_{p_0}^{\bar{p}} q(x)\,dx \tag{3}$$

Alternatively, the consumer surplus may be calculated as the area between the demand curve and the abscissa in Figure 2 in the interval $[0;q(p_0)]$ minus the amount paid for the quantity bought $q(p_0)$ at price p_0.

$$CS(p_0, \bar{p}) = \left(\int_0^{q(p_0)} p(x)\,dx \right) - p_0 q(p_0) \tag{4}$$

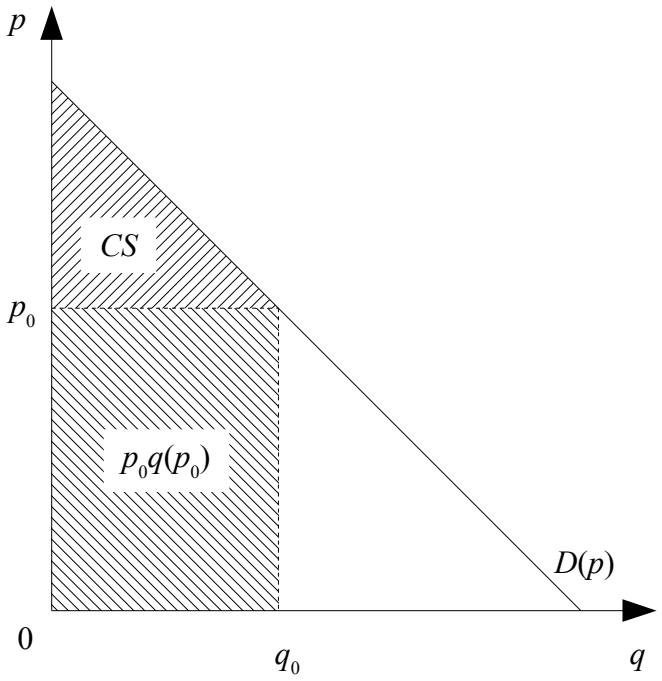


Figure 2: Consumer Surplus

Producer surplus is the difference between the amount a producer receives from the sale of a unit, i.e. its revenues $p_0q(p_0)$, and the amount that unit costs $c(q_0)$ to produce the sold quantity q_0. The variable costs of a firm are defined as the area between the supply curve and the abscissa in the interval $[0;q_0]$. The supply function is defined as

$$q_S = S(p) \quad . \tag{5}$$

The supply curve is defined by the inverse of the supply function and equals the marginal costs c of the most efficient firm in the production of output quantity q.

$$c(q)=S^{-1}(q) \tag{6}$$

Hence, the producer surplus PS can be expressed as in equation (7).

$$PS(p_0)=p_0 q(p_0)-\int_0^{q(p_0)} c(x)\,dx \tag{7}$$

Economic **welfare** $W(p_0)$ (or total surplus) as shown in Figure 3 is defined as the sum of consumer surplus and producer surplus.

$$\begin{aligned} W(p_0) &= CS(p_0)+PS(p_0) \\ &= \int_0^{q(p_0)} p(x)-c(x)\,dx \end{aligned} \tag{8}$$

Choosing q to maximize this expression leads to the first order condition

$$p(q)=c(q) \quad , \tag{9}$$

which occurs precisely at the perfectly competitive equilibrium quantity when demand is downward sloping and marginal costs rise (Jehle and Reny 2000: ch. 4.3.3).

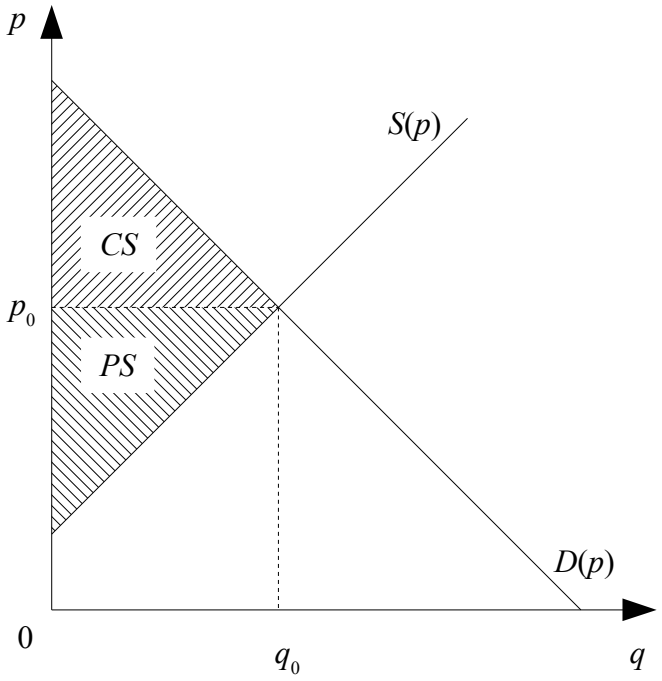


Figure 3: Total Surplus

This situation is (Pareto-)**efficient**. A market outcome is said to be efficient when it is impossible to determine some change in the allocation of capital, labor, goods, or services that would improve the

well-being of one individual in the market without hurting any others (Pepall et al. 2008: p. 35). To see this, suppose the quantity sold was q_1 with $q_1<q_0$ and, thus, $c(q_1)<p(q_1)$. It can be seen from equation (8) that welfare can be raised by increasing the quantity, which is equivalent to lowering the price. This would raise both consumer surplus and producer surplus. This situation is allocatively efficient because resources were allocated to their most efficient use. Hence, total surplus is a measure of allocative efficiency (Cabral 2000: ch. 2.4). The allocatively efficient equilibrium in perfect competition is shown in greater detail in section A.2 .

A.2 Pricing in Perfect Competition

In this section, we present the basic model of perfect competition besides some further relevant issues such as the economic definition of costs. This requires a **definition of competition** (Vickers 1995: p. 4).

> *Competition can be described as a form of rivalry that arises whenever two or more parties strive for something that all together cannot obtain.*

This definition emphasizes the behavioral aspects of competition. However, in economic models such as the model of perfect competition, competition is often treated as a state or a situation. When we say that a market becomes more competitive, this can be the result of

(i) a greater behavioral freedom of rivals (e.g. the freedom to enter an industry – section E.2),

(ii) an increase in the number of rivals (section B.3), and/or

(iii) a move away from collusion towards independent behavior between rivals (section H.2).

Brandenburger and Nalebuff (1996: 18) do not define competition in terms of market structure or performance, i.e. the state of the market, or in terms of firms' conduct. They focus on the products that are supplied by competitors and emphasize that competitors supply substitutable products such as Coca-Cola or Pepsi-Cola. Most chapters of this Reader are concerned with the case where the firms supply homogeneous goods, i.e. perfect substitutes. Differentiated products, i.e. imperfect substitutes, are analyzed in section G .

The Basic Model of Perfect Competition

The model of **perfect competition** is based on five central assumptions (Cabral 2000: ch. 6.1).

1. **Atomicity**: There are many suppliers in the market. Each supplier is so small that its actions (on input and output markets) have no significant impact on other suppliers.
2. **Product homogeneity**: The products of all suppliers are perfectly the same.

3. **Perfect information**: All economic agents know the characteristics of the good and can observe the prices set by all firms.
4. **Equal access to production technologies**: All firms have access to all production technologies.
5. **Free entry**: Any firm may enter or exit the market as it wishes.

A firm in perfect competition acts as a **price taker** on both input and output markets. The price p is not something that the perfectly competitive firm chooses. Instead, that price is determined by the interaction of all the firms and consumers in the market for this good. This implies that the assumption of atomicity does not require infinitely many firms to be in the market. It rather requires the number of firms to be large enough for firms to *think* that their actions will not affect the market price. "An example of a "small" firm would be a wheat farmer in Kansas or, alternatively, a broker on the New York Stock Exchange trading IBM stock" (Pepall et al. 2008: p. 22). As a single firm in perfect competition cannot influence the market price it faces a horizontal residual demand curve. The industry demand curve can, nonetheless, be downward sloping as shown in Figure 2.

The profit $\pi_i(q_i)$ of a perfectly competitive firm i when supplying quantity q_i is defined as the difference between its revenue $R_i(q_i)$ and its total costs $C_i(q_i)$, that can be decomposed into marginal costs $c_i(q_i)$ and fixed costs F_i.

$$\begin{aligned} \pi_i(q_i) &= R_i(q_i) - C_i(q_i) \\ &= p \cdot q_i - \int_0^{q_i} c_i(x)\,dx - F_i \end{aligned} \tag{10}$$

Firm i must decide what optimal quantity $q_{i,opt}$ to supply in order to maximize its profits. The condition for profit maximization

$$max_{q_i} \pi_i(q_i) \tag{11}$$

implies first-order condition (12).

$$\begin{aligned} \frac{d\,\pi_i(q_i)}{dq_i} = \frac{d\,R_i(q_i)}{dq_i} - \frac{d\,C_i(q_i)}{dq_i} &\stackrel{!}{=} 0 \\ p &= c_i(q_{i,\,opt}) \end{aligned} \tag{12}$$

In perfect competition, the marginal revenue $dR_i(q_i)/dq_i$ equals the market price p, which in optimum must equal marginal costs c_i.

In the following, we show that perfect competition is a good situation for two reasons. First, each firm sets the efficient output level, í.e. the output level such that prices equal marginal cost. Second, the set of firms active in the long run is efficient. Because of free entry, firms produce a long-run output such that price equals the minimum average cost. Please note that this refers to

static efficiency, i.e. efficiency at the current point in time. The model is silent about the implications of competition for technological progress (Cabral 2000: ch. 6.1).

Note that the aggregate supply q_s of the n firms in a market is the horizontal sum of each firm's output q_i at price p. This yields the short-run industry supply curve.

$$q_S = S(p, n) = \sum_{i=1}^{n} q_i(p) \tag{13}$$

The **short run** is defined as the period where no entry or exit of firms in this industry occurs. Figure 4 presents the horizontal summation of individual supply curves for an industry with n = 3 firms. For any market to be in equilibrium, first order condition (12) must be satisfied for every firm i. Hence, in equilibrium all firms produce at the same marginal costs equaling the equilibrium price p_0.

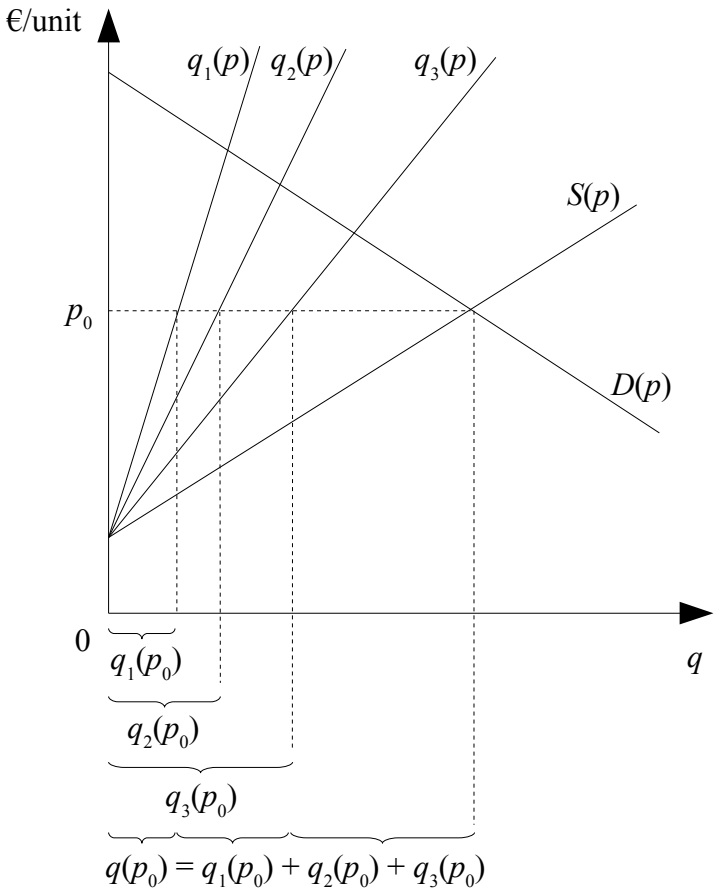


Figure 4: Horizontal Aggregation of Supply-Curves

In the **long run** firms can enter or exit the industry. The above assumption of free entry ensures that in the long run each firm make zero economic profits. If any firm makes positive economic profits, other firms will enter the industry until economic profits of all firms equal zero.

$$\begin{aligned} \pi_i(q_i) &= R_i(q_i) - C_i(q_i) = 0 \\ p &= \frac{C_i(q_i)}{q_i} \end{aligned} \tag{14}$$

Condition (14) denotes the **free-entry equilibrium** because (i) no active firm wishes to leave the market, and (ii) no inactive firm wishes to enter the market. In this case, prices equal average costs. Because the marginal cost curve intersects the curve of average costs at its minimum (see Figure 5), condition (12) for the short run equilibrium is found to apply when condition (14) for the long run equilibrium applies. The minimum of the average cost curve defines the output that can be produced by firm *i* at the lowest unit-costs. This output is also called the Minimum Efficient Scale.

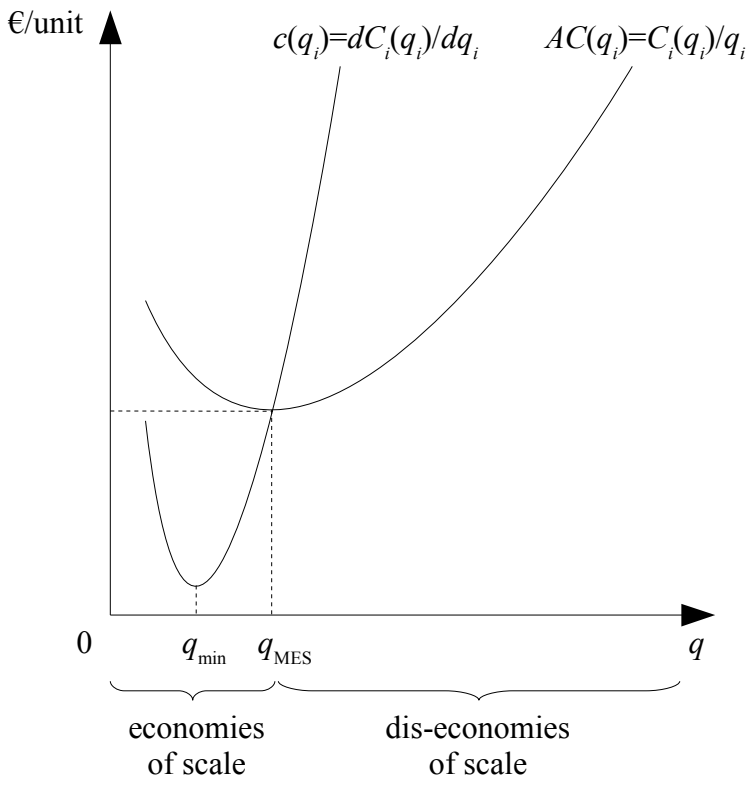


Figure 5: Minimum Efficient Scale

Economic Profits and Costs

The economic profit as defined in equation (10) is not equivalent to the accounting profit as it appears, e.g., in profit and loss accounts. In particular, economic costs do not equal accounting costs because the earlier are defined as opportunity costs. Thus, economic costs include the amount necessary to pay the owners of the firm's capital a risk-adequate, competitive return (**cost of equity**). The opportunity cost for the firm's capital is measured as the rate of return that the capital could earn if invested elsewhere.

"The reason why this is important is because it makes clear that when a firm earns no economic profit it does not mean that its stockholders go away empty-handed. It simply means that those stockholders do not earn more than a normal return on their investment" (Pepall et al. 2008: p. 22). Likewise, when a firm makes a positive economic profit, its stockholders receive a rate of return on their investment that exceeds the normal rate of return that can be earned from investing their money into a firm or project with a comparable risk structure. An excessive rate of return can result form a firm possessing market power (see section E).

Perloff et al. (2007: 15) name eight problems in calculating economic rates of return correctly from internal or external accounting measures. Such difficulties can well drive a wedge between accounting profits and economic profits (eee, e.g., Paha (2009) for an overview).

1. Capital is usually not valued appropriately because accounting definitions are used instead of the economic definitions. For example, **assets** are frequently valued at historic costs instead of their market value (or: fair value).
2. **Depreciation** is usually measured improperly. For example, a linear depreciation schedule rarely is a good representation of the true, economic (or Hotelling (1925)) depreciation of an asset. The economic depreciation is basically the change in the market value of an asset between period t-1 and period t. This corresponds to the valuation of an asset at its *fair value* (IAS 16).
3. Valuing problems arise for advertising and research and development (R&D) because, as with capital, they have lasting impacts. The money a firm spends on R&D this year may generate benefits next year, just as a plant built this year provides a benefit next year. Therefore, it is difficult to decide whether expenses for research and development constitute an **intangible asset**, which must be recognized in the balance sheet and amortized later on.
4. Rates of return may not be properly adjusted for risk. The issue of **risk adjustment** is important because the rate of return of a firm shall be compared to the normal rate of return of an equally risky alternative investment. This gives an indication whether the firm enjoys

market power or not. If the risk structure of the alternative investment does not match that of the firm, the market power assessment is likely to be biased.

5. The risk associated with an investment also depends on the **ratio of debt to equity**. If a firm is financed by a high share of debt a greater portion of the business risk must be borne by the equity holders as in a firm with a lower debt-equity ratio. Hence, the normal rate of return will be higher for the equity holders of the earlier, highly leveraged firm in order to account for the higher risk. This is the case even if the two firms are exposed to the same risks in the product market.
6. Proper adjustments must be made for **inflation**. The earned rate of return can be calculated as either a real rate of return (adjusted for the effects of inflation) or as a nominal rate of return (excluding the effects of inflation).
7. Sometimes, *goodwill* or *intangible assets* (IAS 38 48-53) are recognized as assets whose value implicitly contains a market power effect, i.e. the firm earns higher profits with these assets because of its **market power** and **value**s the assets accordingly. This higher book value of assets incorrectly lowers the rate of return on assets that is reported for this firm.
8. Firms usually base make decisions based on their after-**tax** return. Therefore, rates of return should be calculated as after-tax values.

A.3 Pricing in Pure Monopoly

Monopolies in Theory

The model of **monopoly** rests on several assumptions.

1. There is a well-defined market with **one single supplier**.
2. The seller faces a **negatively sloped demand** $D(p)$ (see equation (1)).
3. There is **no potential entry** by other firms into this market.
4. Here, we assume that the monopolist charges the same price to all customers, i.e. **no price discrimination** occurs. This assumption can be relaxed in more elaborate models of monopoly.

The profit-function of the monopolist is the same than that of a firm in perfect competition (see equation (10)).

$$\begin{aligned} \pi(q) &= R(q) - C(q) \\ &= p(q)\cdot q - \int_0^q c(x)\,dx - F \end{aligned} \tag{15}$$

The only difference between the profit-function of a firm in perfect competition and a monopolist is that the monopolist does not take the price p as given. Instead, the price depends on the quantity produced (i.e. $p(q)$). The assumption of negatively sloped demand implies that the sustainable price is the lower the higher a quantity the monopolist chooses.

Hence, by increasing its output from q_0 to q_1 in Figure 6 the monopolist lowers the market-price from p_0 to p_1. As a result, it looses area L in its revenue but gains the areas G and g. This marginal revenue $dR(q)/dq$ is also shown in Figure 6. The firm's profit[10] rises from $\pi_0=A+L$ to $\pi_1=A+G$. The area g shows the additional costs that are incurred by increasing output.

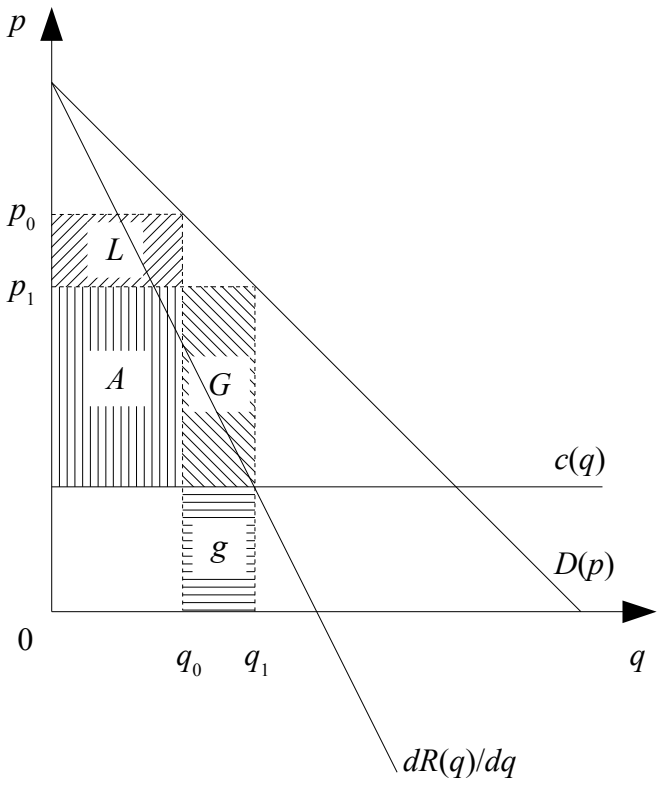


Figure 6: Pure Monopoly

10 We assume that fixed costs F are zero and marginal costs $c(q)$ are constant in output.

Determining the profit-maximizing quantity q_{opt} implies first order condition (16). Notice that, because price and output are related by the demand function, it is the same thing to choose the optimal output or to choose the optimal price. In the following, we assume the monopolist to set an optimal quantity.

$$\begin{aligned} max_q \pi(q) \quad \rightarrow \quad \left(\frac{dp(q)}{dq} \cdot q + p(q)\right) \quad - \quad c(q) \quad &\stackrel{!}{=} \quad 0 \\ \frac{dR(q)}{dq} \quad - \quad \frac{dC(q)}{dq} \quad &= \quad 0 \\ \frac{dR(q)}{dq} \quad &= \quad c(q) \end{aligned} \tag{16}$$

We find that in optimum the marginal revenue $dR(q)/dq$ of a monopolist equals its marginal costs. In Figure 6 this is the case when the monopolist chooses quantity q_1. Since $dp(q)/dq<0$, the marginal revenue of selling one additional unit of output is lower than the current price $p(q)$, i.e. the additional output can only be sold if the price declines. Given the above assumption that no other firm may enter the market, condition (16) is the condition for the short-run *and* the long-run equilibrium.

Re-arranging (16) shows that in its profit-maximum a monopolist chooses a price-cost margin, which equals the inverse of the price elasticity of demand η. This is the well-known **Amoroso-Robinson relation**.

$$\begin{aligned} \frac{p(q)-c(q)}{p(q)} \quad &= \quad -\frac{dp(q)}{dq} \cdot \frac{q}{p(q)} \\ &= \quad \frac{1}{\eta} \end{aligned} \tag{17}$$

The left-hand side of equation (17) is also known as **Lerner-index** (see section E.1). We find that a monopolist may charge a higher optimal markup on marginal costs when consumers are relatively insensitive to changes in price, i.e. when their price-elasticity of demand is low. Similarly, if demand was perfectly elastic (horizontal demand curve) with a willingness to pay at the level of marginal costs, the monopolist would have to charge a price equaling marginal costs.

Welfare Effects of a Monopoly

Now, we use Figure 7 to examine the welfare-effects of a monopoly in comparison to the base case of perfect competition (Motta 2004: ch. 2.2.2). In the perfectly competitive short-run equilibrium, firms would equalize the price to marginal costs (see condition (12)) and set p_c. This is a long-run equilibrium, too, because firms make zero economic profits. Welfare W_c equals consumer surplus CS_c and encompasses the areas A, B, and C. As we have seen above, a monopolist would optimally sell quantity q_1 at price p_1. Welfare W_1 would equal the sum of consumer surplus CS_1 (= area A) and

producer surplus PS_1 (= area B). Hence, in comparison to perfect competition the monopoly causes a redistribution of rents (= area B) from consumers to producers because of higher prices. Moreover, the increase in price comes along with a reduction of the quantity sold. This causes a **deadweight loss** in welfare (= area C). A welfare loss occurs not just for the monopoly price but for any price above marginal costs. One may see from Figure 7 that the deadweight loss caused by market power is the larger the higher the market price p.

Please note that total welfare in the monopoly-case is smaller than welfare in perfect competition. However, producer surplus in monopoly is higher than that in perfect competition. The monopoly-situation is inefficient, because by lowering the price one could increase the quantity sold and make consumers better off. This is not a **Pareto improvement** (i.e., not everybody is better off), since the producer surplus shrinks with respect to the monopoly case. However, it would be possible to redistribute rents such that the profit of the monopolist is not reduced (Pepall et al. 2008: p. 39).

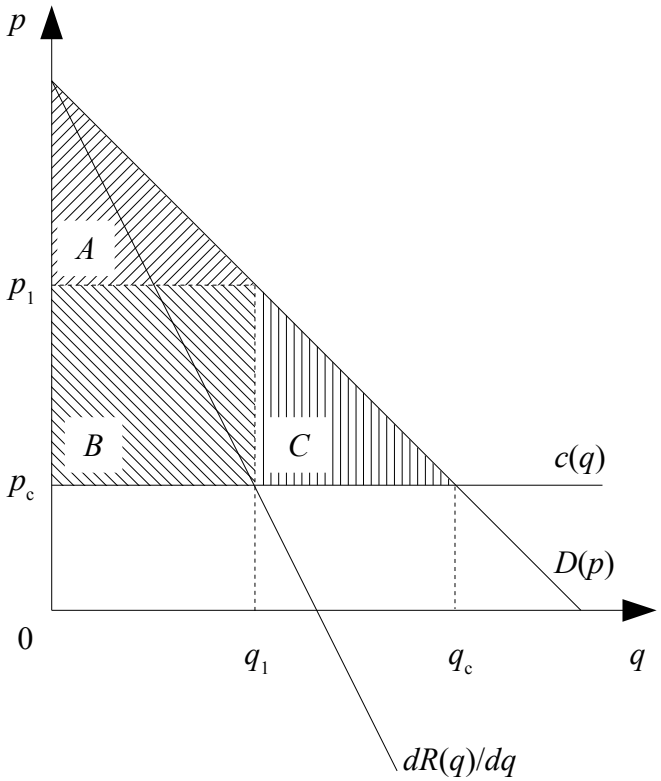


Figure 7: Welfare in Monopoly

In addition to the above deadweight loss, monopolies may create an additional welfare loss because of **rent-seeking activities** (Motta 2004: ch. 2.2.3). To see this, recall that a monopolist makes producer surplus B in Figure 7 as compared to a producer surplus of zero in perfect competition. Therefore, it would be individually profitable for a firm to invest an amount up to B in,

e.g., lobbying activities or the creation of entry barriers in order to acquire or maintain monopoly power. This investment does not necessarily have any social value and, thus, constitutes an additional welfare loss.

Monopolies in Reality

Reasons for the existence of monopolies can be lower costs, higher quality (or better reputation for quality), or network effects (network industries (see section E.2), natural monopolies (see section C)). **Examples of pure monopolies** are rare. One rather finds industries where one firm commands a high market share while its competitors are rather insignificant. Examples of such industries are the mainframe computer industry in the 1960/70s with IBM as a dominant firm, or the industry for photographic films with Kodak as a dominant firm in the late 20th century (Cabral 2000: p. 71). Pepall et al. (2008: 29) provide the following example of a monopoly:

> *"It is not always easy to find examples of the classic monopoly behavior described in economics textbooks. However, Tyco International's control of the plastic hanger market in the late 1990s may have come pretty close. Retail firms such as J. C. Penny and K-Mart use only plastic hangers to display their clothing goods. Starting in about 1994, Tyco used mergers and acquisitions of rival firms to gain control of 70 to 80 percent of the market for plastic hangers. In a number of geographic regions, Tyco became the only plastic hanger firm available. In 1996, Tyco acquired a Michigan-based hanger firm, Batts, that was one of the largest suppliers to the Midwest region. Immediately thereafter, Tyco raised prices by 10 percent to all its customers. Some clients grumbled but most accepted the higher prices. Others though, such as K-Mart and VF (makers of Lee and Wrangler jeans) informed Tyco that they had an alternative hanger supplier, namely a company called WAF. For a brief moment, Tyco appears to have backed off raising the price. Yet the firm's underlying strategy soon became clear. In the fall of 1999, Tyco bought the WAF Corporation. Within a few months, it not only raised prices to all its customers again but, this time, it also added in a new delivery charge. Tyco also pursued an aggressive repurchase program so as to corner the market on used hangers. If it did not control the supply of this alternative to new hangers, Tyco would have faced increasing difficulty in charging a high price."*

This example highlights a few points, that are not obvious from the above theoretic model of a monopolized industry. First, real industries are frequently characterized by the existence of a (small) number of firms rather than a single firm. This requires modeling the interaction among

firms as is shown in sections B.2 and B.3 Second, the relevant market must be well-defined (see assumption 1 above). This requires identifying the relevant substitutes of a product, such as for example used hangers. Moreover, one has to identify the geographic scope of the relevant market. These issues are addressed in greater detail in section F . Third, only after the relevant market has been defined one can engage in attempts to forecast the likely price-increase of a merger like that between Tyco and Batts. This is illustrated more closely in section G .

Lessons Learned

After reading this section you should be able to answer the following questions.

1. What is the difference between the demand function and the inverse demand function?
2. Define the following economic concepts: producer surplus, consumer surplus, and welfare.
3. Show that the market outcome in perfect competition is allocatively efficient.
4. Determine the industry-supply curve when the marginal cost of each firm in an industry is $c_i(q_i)=4q_i+8$. Assume the number of firms in this industry to be $n = 80$ (Pepall et al. 2008: 23).
5. Explain why the amount of economic profits is below the amount of accounting profits.
6. Show that allocative efficiency in a monopoly is lower than in perfect competition.

References

Brandenburger, A.M. and Nalebuff, B.J. (1996). "Co-opetition." Doubleday: New York

Cabral, L.M. (2000). "Introduction to Industrial Organization." The MIT Press: Cambridge

Hotelling, H. (1925). "A General Mathematical Theory of Depreciation." Journal of the American Statistical Association. Vol. 20 No. 151, pp. 240-353

Jehle, G.A. and Reny, P.J. (2000). "Advanced Micoeconomic Theory." 2nd edition. Addison Wesley: Boston

Motta, M. (2004). "Competition Policy – Theory and Practice." Cambridge University Press: Cambridge

Paha, J. (2009). "Using Accounting Data in Cartel Damage Calculations – Blessing or Menace?" Forthcoming: European Journal of Law and Economics

Pepall, L. and Richards, D. and Norman, G. (2008). "Industrial Organization – Contemporary Theory and Empirical Applications." 4th edition. Blackwell Publishing: Malden, MA

Perloff, J.M., Karp, L.S., Golan, A. (2007). "Estimating Market Power and Strategies." Cambridge University Press: Cambridge

Vickers, J. (1995). "Concepts of Competition." Oxford Economic Papers. Vol. 47 No. 1, pp. 1-23

B STRATEGIC INTERACTION IN STATIC INDUSTRIES: OLIGOPOLISTIC COMPETITION

Above, we were concerned with the two extremes of market structures, i.e. monopoly with just one firm and perfect competition with many firms. Most markets in the real-world are in between these two extremes with a few firms being active in these markets. The situation in which there are a few competitors is called **oligopoly** (see, e.g., Cabral (2000: ch. 7) or Motta (2004: ch. 8.4)). The distinguishing feature of such oligopolies is the fact that firms in these industries take into account the (expected) behavior or their competitors when making their own strategic decisions. This is called the **strategic interdependence** of competitors. Decisions in such an interactive setting are called **strategic decisions**, and **game theory** is the branch of social science that formally analyzes and models strategic decisions. Therefore, section B.1 provides an introduction to game theory.

In case of the monopoly-model, it is easy to show that the market-outcome is the same when firms either set prices or quantities. However, once we leave the world of monopoly the equivalence of price and output strategies vanishes. In oligopolistic markets it matters very much whether firms compete in terms of quantities, as in Cournot, or in terms of price, as in Bertrand (Pepall et al. 2008: p. 224). In section B.2 , we present the Bertrand oligopoly-model, and turn to Cournot-competition in section B.3 .

B.1 Introduction to Static Game Theory

Game theory is divided into two branches: **non-cooperative** and **cooperative** game theory.[11] The essential difference between these two branches is that in non-cooperative games, the unit of analysis is the individual decision-maker or **player**, e.g., the firm. The non-cooperative setting means that each player is concerned only with doing as well as possible for herself, subject to the rules of the game. By contrast, cooperative game theory takes the unit of analysis to be a group or coalition of players, e.g., a group of firms. In section B , we concentrate on non-cooperative, static game. The term **static** means that firms only care about their payoffs in the current period. In sectionError: Reference source not found, we turn to cooperative, dynamic games. The term **dynamic** means that firms care about their payoffs in the current and in future periods.

Firms' strategic interaction can be **sequential** or **simultaneous**, i.e. a firm may, for example, decide on the price for its good after observing the price set by its rival, or it may set a price at the same time as its rival. In this lecture, we concentrate on simultaneous games.

Two basic assumptions underlie the application of non-cooperative game theory to

11 Section B.1 is based on Pepall et al. (2008: ch. 9) and often quotes from this textbook.

oligopoly. The first is that **firms are rational**. They pursue well-defined goals, principally profit maximization. Moreover, players' rationality is **common knowledge**. An event E is defined to be common knowledge if all players know E, all players know that all other players know E, all players all players know that all other players know that all other players know E, ad infinitum. The second basic assumption is that firms apply their rationality to the process of **reasoning strategically**. That is, in making its decision, each firm uses all the knowledge it has to form expectations regarding how other firms will behave. For example, an oligopolist can anticipate the response of its rivals by asking itself "What would I do if I were the other player?"

Each player's decision or plan is called a **strategy** (or: **action**), i.e. a firm, e.g., chooses a price p or a quantity q. A strategy is a fully-specified decision rule as to how to play the game. A **strategy set** S_i is the set of feasible strategies for player i. A list of strategies showing one particular strategy choice for each player is called a **strategy combination** (or: strategy profile). Any given strategy combination determines the **outcome** of the game that describes the **payoffs** or final gains earned by each player. In other words, a payoff function assigns a real number to each strategy profile. In the context of oligopoly theory, payoffs are often interpreted as firm i's profit π_i.

For a game to be interesting, at least one player must be able to choose from more than one strategy so that there will be more than one possible strategy combination, and more than one possible outcome of the game. Yet while there may be many possible outcomes, not all of these will be **equilibrium** outcomes. By equilibrium we mean a strategy combination that no firm has an incentive to change the strategy it is currently using given that no other firm changes its current strategy. This is called the **Nash equilibrium concept**. Hence, in the Nash-equilibrium each player's action is the **best response** to the actions of all other players. More formally, in a game with n players, denoting with S_i the set of actions available to player i (with $i = 1, ..., n$), and with player i's payoff $\pi_i(\sigma_1, ..., \sigma_i, ..., \sigma_n)$, the n-tuple $(\sigma_1^*, ..., \sigma_i^*, ..., \sigma_n^*)$ is a Nash equilibrium (Motta 2004: p. 543) if

$$\pi_i\left(\sigma_1^*, \ldots, \sigma_i^*, \ldots, \sigma_n^*\right) \geq \pi_i'\left(\sigma_1^*, \ldots, \sigma_i', \ldots, \sigma_n^*\right), \forall i = 1,2,\ldots,n \text{ and all } \sigma_i \in S_i \quad . \tag{18}$$

Note that every player is solely interested in choosing the action (subject to the actions chosen by the other players) that maximizes his own payoff, i.e. the player is rational because his behavior maximizes his payoff given his beliefs over unknown variables such as the other players' strategies. In the words of the US-American financier and political consultant Bernard Baruch (1870-1965) this objective may be stated as follows: You don't have to blow out the other fellow's light to let your own shine (cited according to Brandenburger and Nalebuff 1996: 4). Hence, competition among firms, which is a central concept of this lecture, does not emerge from firms' interest to harm their rivals but from their interest to do as well as possible for themselves.

A core objective of game theory is to **solve games**, i.e. to determine their Nash equilibria. In other words, our objective is to describe what strategies players will choose and what the resulting outcome will be. Sometimes Nash equilibria are easy to determine. This is because some of a firm's possible strategies may be dominated. For example, suppose that we have two firms *A* and *B*, in a market whose set of possible strategies is S_i (with $i = A, B$). Suppose that one of *A*'s strategies $\underline{\sigma_i}$ is such that it is never a profit-maximizing strategy regardless of the choices (σ_j) made by *B*, i.e. there is always a strategy σ_i' such that

$$\pi_i\left(\sigma_i', \sigma_j\right) > \pi_i\left(\underline{\sigma_i}, \sigma_j\right) \forall\, \sigma_j \text{ with } \left(i \neq j\right) \tag{19}$$

applies. Then we say that the strategy in question is **strictly dominated**: rationally speaking, it will never be chosen. Dominated strategies cannot be part of the equilibrium outcome and can be eliminated one by one. Similarly, a **dominant strategy** is one that outperforms all of a firm's other strategies no matter what its rivals do. That is, it leads to higher profits than any other strategy the firm might pursue regardless of the strategies selected by the firm's rivals.

This section B shows that competing firms rarely end up with the highest payoffs that they could theoretically make. Brandenburger and Nalebuff (1996: 10) suggest that you "can play the game extremely well, and still fare terribly. That's because you're playing the wrong game: you need to change it." The remainder of this lecture is concerned with an analysis of the ways firms attempt to change the game. Some of these strategies are legal and benefit consumers such as research and development that creates more efficient production technologies or better products (section C). Other strategies can be illegal like attempts of a dominant firm to exclude its rivals from the market (section I.3). Sometimes, governments attempt to change the game by granting aid to domestic firms at the expense of foreign firms (section K). In Europe, such state aid can be illegal if the positively and negatively affected firms do business in European member states such that the aid would have adverse effects within the European Union.

B.2 Pricing in Bertrand-Competition with Homogeneous Products

In Bertrand-competition, firms reason what price to choose in order to sell the output produced. Cournot-competition takes this analysis one step back and asks what output capacity-constrained firms should optimally produce. Although the Bertrand model was published about 50 years after the Cournot model (see section B.3), we present both models in reverse order because the Bertrand model shares more similarities with perfect competition.

The Basic Bertrand Model

The characteristic feature of the Bertrand model is its assumption that firms use price as strategic variable. The model rests on the following assumptions.

1. The market consists of ***n* identical** firms (set of players). The marginal costs of firm i are constant in output and equal those of firm j ($c_i=c_j \; \forall \; i,j$). The firms do not incur fixed costs.
2. The firms supply a **homogeneous product**, i.e. the firms' products are perfect substitutes.[12]
3. The firms face a **downward-sloping demand** $D(p)$, which is continuous and bounded. Demand is perfectly observed by all firms. The consumers demand from the firms with the lowest price. If there is more than one firm with the lowest price, demand is equally divided among them.
4. Firms' strategic variable is price. They **simultaneously set prices**. The strategy set for each firm is $[0,\overline{p}]$ with $D(p)=0$ if $p \geq \overline{p}$.
5. The firms are **not capacity-constrained**, i.e. each firm would be able to supply the entire market.[13]
6. The firms play a **one-shot game**, i.e. they are only interested in the profits of the current period.

Suppose, the number of firms in the market is $n=2$. Every firm is concerned with determining a best price response given its expectation of the other firm's price. For determining this best response, consider that the firm which sets the lowest price may supply the entire demand. This is because the homogeneity of products (assumption 2) implies that customers always buy at the cheapest offer. If both firms set the same price p, the market is split evenly and both firms receive half of the demand $D(p)/2$. Assumption 5 (no capacity constraints) ensures that every firm would be able to serve the entire market. What is the best strategy in this context?

(i) If firm 1 conjectures that firm 2 sets a price above the monopoly price p_m, firm 1 should set the monopoly price. With this strategy, it gets all of the demand and receives the maximum possible (i.e. monopoly) profits π_m while firm 2 makes zero profits.

(ii) Now, firm 1 conjectures that firm 2 sets a price p_2 below the monopoly price but above marginal costs. If firm 1 sets $p_1=p_2$ it receives half of the demand at this price $D(p_2)/2$, as shown in Figure 8. Therefore, firm 1 should set a somewhat lower price, $p_1^*=p_2-\varepsilon$. With this strategy, it receives the entire demand at this price and makes profits π_1 while firm 2 makes

12 The assumption of homogeneous products is relaxed in section G.2 .

13 Further below in this section the case of constrained capacities is analyzed.

zero profits. This is (almost) a doubling in profits as compared to setting the same price as firm 2. Firm 1's demand is discontinuous as shown by the solid line in Figure 8 (Pepall et al. 2008: p. 225).

$$D_1(p_1, p_2)=\begin{cases} D(p_1), & \text{if } p_1<p_2 \\ D(p_2)/2, & \text{if } p_1=p_2 \\ 0, & \text{if } p_1>p_2 \end{cases} \tag{20}$$

(iii) If firm 1 expects firm 2 to set a price below marginal costs c it should set a price at the level of marginal costs in order to avoid losses $\pi_1<0$.

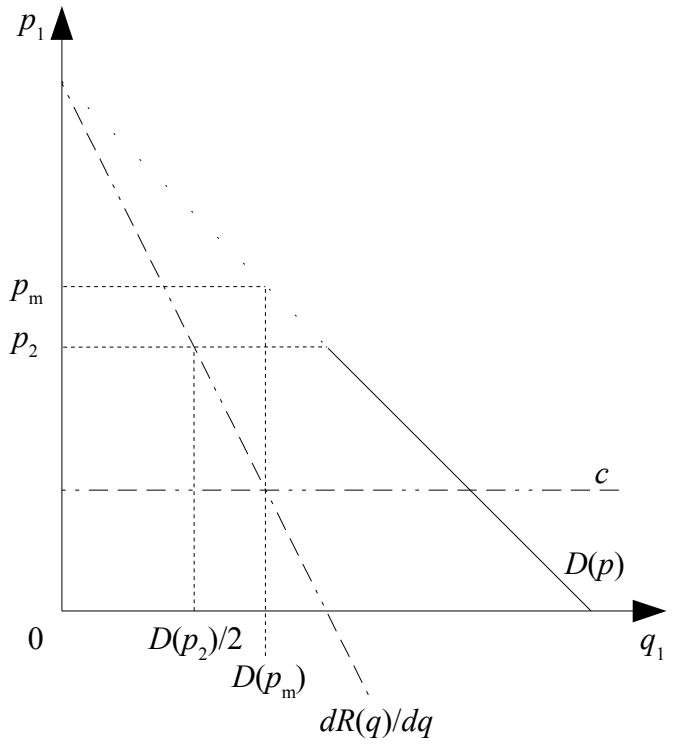


Figure 8: Demand Curve in the Bertrand Model

The **best responses** (or reaction functions) of firms 1 and 2 are shown in Figure 9. A reaction function is a function $p_i^*(p_j)$ that shows firm i's optimal price for each price of firm j. A **Nash equilibrium** is a pair of strategies (here: a pair of prices) such that no firm can increase profits by unilaterally changing price (Motta 2004: ch. 8.4.1).

$$\pi_i(p_i^*, p_j^*) \geq \pi_i'(p_i, p_j^*) \tag{21}$$

In Bertrand competition, the market is at equilibrium when both firms charge a price that equals marginal costs.

$$p_1^*(p_2)=p_2^*(p_1)=c \tag{22}$$

The equilibrium is not defined by first-order conditions because the discontinuity in residual

demand implies that a firm's payoff function is not differentiable everywhere. Neither firm would charge a price below marginal costs as this would imply making losses. However, unilaterally charging a higher price is not possible, either, because the firm with the higher price looses all demand to the firm with the lower price. Therefore, the existence of just two firms, which are not constrained in capacity, would be enough to cause a perfectly competitive market-outcome. This is called the **Bertrand paradox**. The equilibrium price is independent of the number of firms when there are at least two firms.

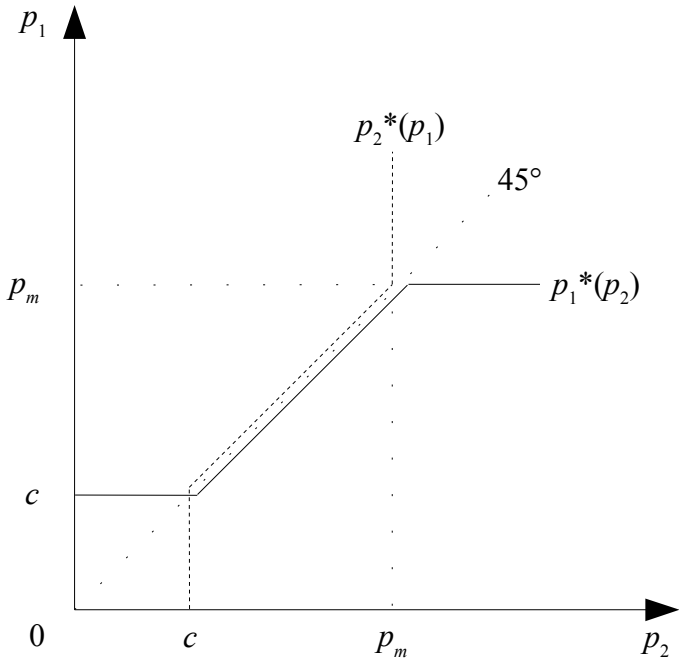


Figure 9: Best Responses in Bertrand Competition

Pepall et al. (2008: 228) provide the following example of Bertrand competition:

> *"Perhaps one of the most dramatic examples of Bertrand competition comes from the market for flat screen TVs. Such screens use one of three basic technologies [(LCD, DLP or plasma). ... O]ver time, the differences between the three types have diminished. The result has been the eruption of a severe price war. From mid-2003 to mid-2005, prices for new TVs based on these technologies fell by an average of 25 percent per year. Fifty-inch plasma TVs that sold for $20,000 in 2000 were selling for $4,000 in 2005. Nor has this pressure let up. In November 2006, Syntax-Brillian cut the price on its 32-inch LCD TV by 40 percent. Sony and other premium brands were forced to follow suit. Prices on all models fell further. Indeed, when Sony was rumored to be thinking of further reducing its 50-inch price to $3,000, James Li, the chief executive of Syntax-Brillian, was quoted as saying, "If they go to $ 3,000, I will go to $ 2,999." Bertrand would have been proud."*

Another example for intense price competition among capacity-unconstrained firms is the US-American industry for solar panels:[14]

> *In November 2011 the US-American commerce department opened an investigation into the market for solar panels because American producers accuse Chinese producers of being subsidized[15] and dumping solar panels into the US-market at prices even below production costs. Despite demand for solar panels in USA has been growing since 2008 at a rate of 70% per year, Chinese producers have grown faster to export about 95% of their production built up a US market share of more than 50%. As a consequence, prices of solar panels per watt of capacity have been falling from USD 3.30 in 2008 to USD 1.00-1.20 in November 2011.*

Solving the Bertrand-Paradox

The result of prices equaling marginal costs is not necessarily realistic because in most real-world oligopolies firms may be assumed to make more than zero profits. This Bertrand-paradox is caused by the strong assumptions of the Bertrand-model (Cabral 2000: p. 105).

1. The above assumption 2 implies that all firms supply a homogeneous product. However, when firms sell **differentiated products** and consumers possess a love for this variety, firms can charge prices above marginal costs. The idea of this result is that firms specialize on different segments of the market which lowers competition in each of these segments. Therefore, the firms may charge prices above marginal costs. Bertrand-competition with differentiated products is introduced in section G.2 .
2. The above assumption 6 implies that the firms play a one-shot game. This prevents retaliatory actions by the competitors. Consider the case of a **dynamic game** where firms interact over many periods. In this case, firms could set a price above marginal costs. If one firm decided to unilaterally lower its price in order to gain additional demand the other firms could lower their prices in the following even further in order to punish the deviator. In sectionError: Reference source not found, we explore the conditions under which firms can sustain such supracompetitive prices.
3. The above assumption 5 implies that firms are not **capacity-constrained**. Thus, by setting a lower price than its rivals, a firm wins the entire demand. In section B.3 , we show one way how capacity-constraints affect the competitive market-outcome, i.e. we assume that firms

14 http://www.nytimes.com/2011/11/10/business/global/us-and-china-on-brink-of-trade-war-over-solar-power-industry.html?pagewanted=1#

15 The economic analysis of state aid in the European Union is described in section K .

compete in quantities.

A second possibility for considering capacity-constraints is to model them explicitly in the Bertrand-model (see Cabral (2000: p. 105) and Motta (2004: p. 555)). Thus, consider the industry shown in Figure 10. Market demand $D(p)$ is assumed to be downward sloping. Marginal costs c are assumed to be zero. Firm 1 is capacity-constrained and cannot sell more than quantity k_1. Firm 2 is capacity-constrained and cannot sell more than quantity k_2. The capacity-constraints are binding because $k_i<D(p_i=c)$.

Now, consider the profit-maximization problem of firm 2. If firm 2 sets a price $p_2 \leq p_u$ firm 1 will set a price $p_1=p_2-\varepsilon$ and sell as much quantity as possible, i.e. it will sell the quantity k_1. The residual demand of firm 2 $D_2(p_2)$ equals the market demand at price p_2 minus the quantity supplied by firm 1, i.e. $D_2(p_2)=D(p_2)-k_1$. Moreover, we show firm 2's marginal revenue $dR_2(q_2)/dq_2$.

What price should firm 2 optimally choose? For any price $p_2>p_l$ makes a marginal revenue above zero. Consequently, a capacity-unconstrained firm would set an optimal price p_l. However, at this price firm 2 would have to supply a greater quantity than it can produce. Given its capacity-constraint, firm 2 sets an optimal price price p_{opt}. We find that, if total industry capacity is low in relation to market demand, equilibrium prices are greater than marginal cost and every firm sells and output equal to its capacity.

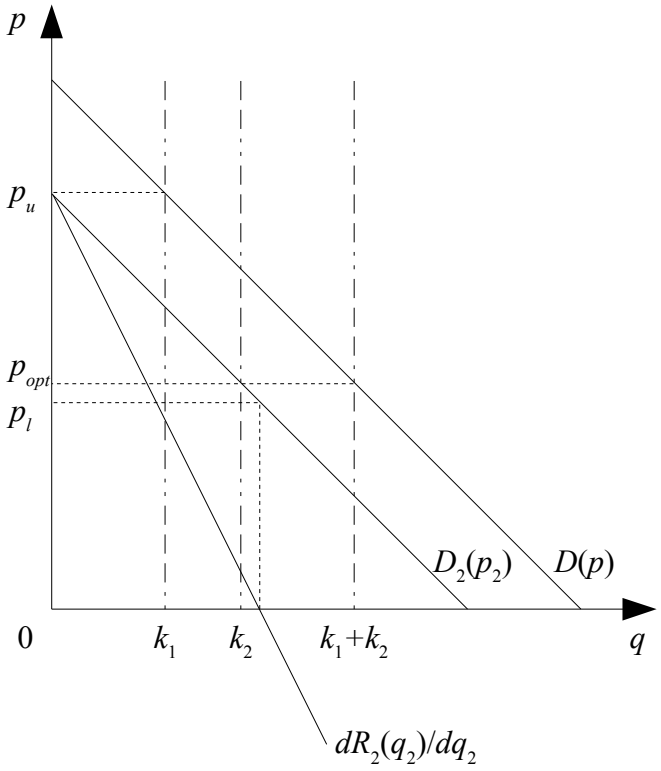


Figure 10: Bertrand Competition with Capacity Constraints

B.3 Pricing in Cournot-Competition with Homogeneous Products

In Bertrand-competition, firms reason what price to choose in order to sell the output produced. Cournot-competition asks what output capacity-constrained firms produce and what price they charge. In a Cournot-model one can, e.g., answer the question: What is the effect of the number of firms on price?

The Basic Cournot-Model

A simple Cournot-model is characterized by the following assumptions.

1. The market consists of ***n* identical firms**. The marginal costs of firm i are constant in output and equal those of firm j ($c_i=c_j \; \forall \; i,j$).
2. The firms supply a **homogeneous product**.
3. Firms' strategic variable is quantity. They **simultaneously set quantities**. The price is set as to clear the market.
4. The firms face a **downward-sloping demand** $D(p)$, which is perfectly observed by all firms.
5. The firms are **capacity-constrained**, i.e. neither firm would be able to supply the entire market.
6. The firms play a **one-shot game**, i.e. they are only interested in the profits of the current period.

The output q supplied by all firms is the sum of the output of all other firms q_{-i} plus the output of firm i, i.e. q_i. Given the inverse demand function (2), the market clearing price at total output q is

$$p = p\left(q_{-i}+q_i\right) \quad . \tag{23}$$

Hence, the profit function of firm i may be denoted as follows.

$$\pi_i = \left(p\left(q_{-i}+q_i\right)-c\right)\cdot q_i \tag{24}$$

Assume for the moment that n=2 applies. Thus, the two firms 1 and 2 choose quantities q_1 and q_2 in order to maximize their profits.

$$\pi_1 = \left(p\left(q_1+q_2\right)-c\right)\cdot q_1$$
$$\pi_2 = \left(p\left(q_1+q_2\right)-c\right)\cdot q_2$$

To illustrate firm 1's decision consider the demand curve $D^{-1}(q)$ as shown in Figure 11. When firm 2

decides to supply quantity q_2, firm 1's **residual demand** curve $D_1^{-1}(q_1,q_2)$ moves to the left by exactly this amount. Firm 1's best response to firm 2's output choice is determined by its first-order condition

$$\begin{array}{rcl} \frac{d\pi_1}{dq_1} & = & p(q_1+q_2)+\frac{dp(q_1+q_2)}{dq_1}\cdot q_1 - c \stackrel{!}{=} 0 \\ & & dR_1(q_1)/dq_1 = c \end{array}, \qquad (25)$$

i.e. firm 1 chooses an optimal quantity $q_1^*(q_2)$ such as to equalize marginal cost and marginal revenue. Re-arranging equation (25) for the optimal price yields

$$p = c - \frac{dp}{dq_1}\cdot q_1 \quad . \qquad (26)$$

Because of $dp/dq<0$, the market **price** in Cournot-competition **is above marginal costs**.

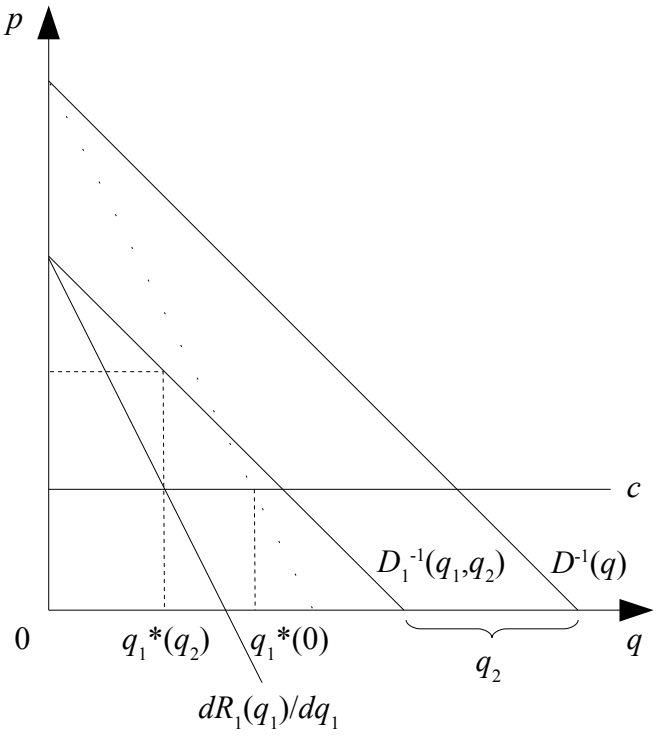


Figure 11: Cournot Optimum

$q_1^*(q_2)$ is firm 1's best-response or **reaction function** given 2's choice of quantity q_2. Suppose firm 2 chose a quantity $q_2=0$ so that $D_1^{-1}(q_1,q_2)=D^{-1}(q_1)$. Using first-order condition (25), it is easy to show that firm 1's best response is setting $q_1^*(0)$. Given that firm 2 sets a quantity $q_{2,c}$ such that $p=c$, firm 1 should set $q_1^*(q_{2,c})=0$. This reaction function of firm 1 is shown in Figure 12. Because firm 2 is assumed to be symmetric to firm 1, the above reasoning applies to firm 2 as well. Therefore, the reaction-function of firm 2 is also shown in Figure 12. The **Nash-equilibrium** of this game is at the point where both reaction functions intersect. To see this, suppose firm 2 sets quantity

$q_{2,A}$. In this case, it would be optimal for firm 1 to set $q_{1,A}$. This would induce firm 2 to set $q_{2,B}$ so that firm 1 optimally chooses $q_{1,B}$ (Cabral 2000: p. 123). The combination of quantities $q_{1,opt}$ and $q_{2,opt}$ is the only set of quantities where neither firm would have an incentive to adjust its output.

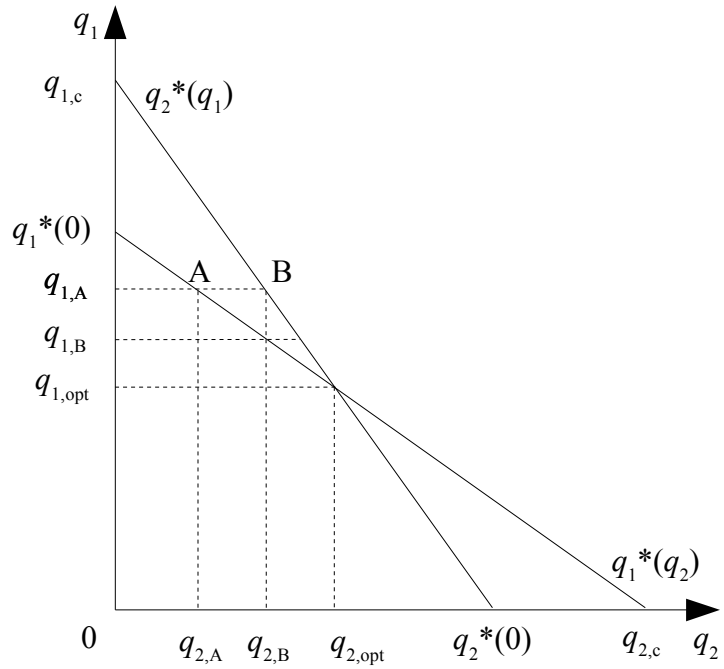


Figure 12: Cournot Reaction-Functions

Comparing the Cournot-Model to other Models

Monopoly is the extreme case of a Cournot-model with n=1 sellers. Consider that $q_1^*(0)=q_2^*(0)$ is the quantity that a profit-maximizing monopolist would set. Therefore, the line that connects these points in Figure 13 shows all divisions of the quantities of firm 1 and firm 2 that add up to the monopoly quantity. We find that the aggregate quantity supplied by the Cournot-duopolists exceeds the quantity of a monopolist.

Perfect competition is the extreme case of Cournot-competition with infinitely many sellers. The quantity $q_{1,c}=q_{2,c}$ implies a price p equaling marginal costs c. This is the condition for a short-run equilibrium in perfect competition (see equation (12)). Therefore, the line that connects these quantities in Figure 13 shows all divisions of the quantities of firm 1 and firm 2 that add up to the competitive quantity.

We find that the aggregate quantity supplied by the Cournot-duopolists is below the quantity in a perfectly competitive market. Likewise, the Cournot-price is lower than the monopoly-price and greater than the price in perfect competition.

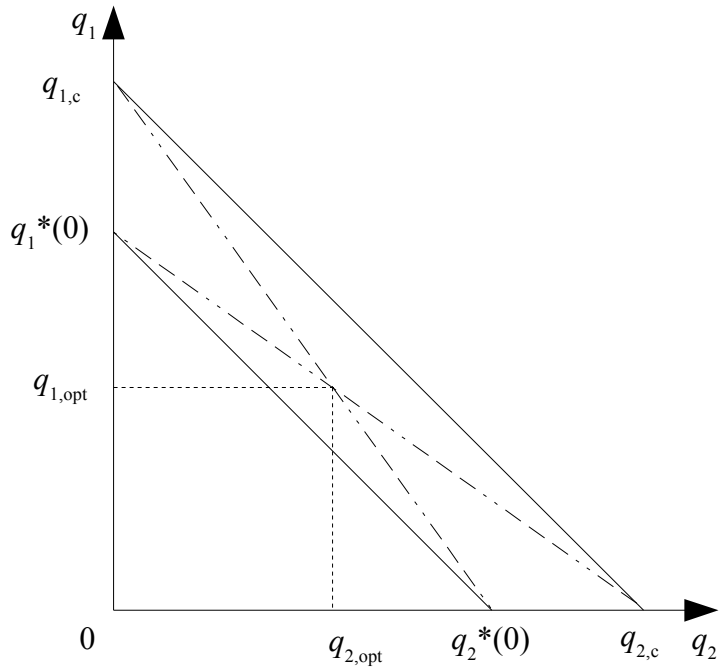


Figure 13: Oligopoly, Monopoly, Competition

In contrast to the Cournot-model, the Bertrand-model predicts that duopoly competition is sufficient to drive prices down to the level of marginal costs. Hence, two firms are enough to achieve the perfectly competitive price level. This decisive difference implies that the two models describe two very different sorts of industries. If capacity and output can be easily adjusted (→ no capacity constraints), the Bertrand model is a better approximation of duopoly competition. Examples include software, insurance, and banking. If output and capacity are difficult to adjust (→ existence of capacity constraints), the Cournot model is a good approximation of duopoly competition. Examples include wheat, cement, steel, cars, and computers (Cabral 2000: 113).

Kreps and Scheinkman (1983: 326) show that the outcomes of a Cournot-model are identical to those of a "two-stage oligopoly game where, first, there is simultaneous production, and, second, after production levels are made public, there is price competition." The first stage can also be interpreted as one where the firms choose a production capacity. The second stage, corresponds to Bertrand-like price competition where production of a homogeneous good is carried out subject to the capacity constraints generated by the first-stage decisions. The size of these capacities is assumed to be common knowledge. The results provided by Kreps and Scheinkman (1983) provide a justification for applying the quantity game to industries in which firms are choosing price.

Lessons Learned

After reading this section you should be able to answer the following questions.

1. What is meant by a Nash-equilibrium?
2. Define the term *dominant strategy*.
3. In what way do the assumptions of the Bertrand model affect its outcome?
4. What is the most important characteristic that distinguishes the Bertrand model from the Cournot model?
5. Calculate the Cournot-equilibrium for a duopoly and an inverse demand function $p(q)=a-bq$ (Cabral 2000: p. 110).
6. Explain why welfare in at a Cournot-equilibrium is lower than welfare in perfect competition.

References

Brandenburger, A.M. and Nalebuff, B.J. (1996). "Co-opetition." Doubleday: New York

Cabral, L.M. (2000). "Introduction to Industrial Organization." The MIT Press: Cambridge

Motta, M. (2004). "Competition Policy – Theory and Practice." Cambridge University Press: Cambridge

Kreps, D.M. and Scheinkman, J.A. (1983). "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes." The Bell Journal of Economics. Vol. 14 No. 2, pp. 326-337

Pepall, L. and Richards, D. and Norman, G. (2008). "Industrial Organization – Contemporary Theory and Empirical Applications." 4th edition. Blackwell Publishing: Malden, MA

C DYNAMIC AND PRODUCTIVE EFFICIENCY

Production technologies and, thus, production costs are an important factor for explaining firm-behavior. To see this, recall that the existence of capacity-constraints is *the* important distinguishing feature between Cournot-competition and Bertrand-competition. Moreover, technology and costs are an important determinant of the industry's structure. Therefore, subsection C.1 provides an introduction to relevant cost-concepts as well as technological issues, e.g. economies of scale and/or scope, under consideration of productive efficiency and industry-structures. Subsection C.2 extends this discussion to dynamic efficiency, considering the effect of research and development on a firm's performance.

C.1 Productive Efficiency and Production Technologies

A firm's **production technology** is a production relationship that describes how a given quantity of inputs ι_j is transformed into the firm's output q (Pepall et al. 2008: ch. 4.1). Microeconomic theory describes this technological relationship by the **production function** f of a firm.

$$q = f(\iota_1, \iota_2, \ldots, \iota_k) \tag{27}$$

Firms' desire to maximize profits (see equations (11), (16), and (25)) can also be formulated as a desire for minimizing the costs for some quantity of output $\bar{q}$, with w_i being the prices of the inputs ι_j.

$$\min_{\iota_j} \sum_{j}^{k} w_j \iota_j \tag{28}$$

Productive efficiency is achieved if a firm produces output $\bar{q}$ at these lowest possible costs. The relationship between (minimum-)costs and output is described by the **cost function** $C(q)$ of the firm.

$$C(q) = F + \int_0^q c(x)\,dx \tag{29}$$

F denotes the **fixed costs** of the firm that it must bear irrespective of the quantity produced. Fixed costs have to be distinguished from sunk costs S. While fixed costs are incurred every period, sunk costs are a cost component that is incurred just once – typically as a prerequisite for entry – and cannot be recovered when the firm decides to exit the market. Additionally, the firm incurrs **variable costs** $\int c(x)dx$ with $dC(q)/dq=c(q)$ denoting its **marginal costs**, i.e. the addition to total cost that is incurred in increasing output by one unit. Hence, **average costs** $AC(q)$ and **average variable costs** $AVC(q)$ can be expressed as

$$AC(q)=C(q)/q \quad \text{, and} \tag{30}$$

$$AVC(q)=\frac{\int_0^q c(x)\,dx}{q}=AC(q)-\frac{F}{q} \quad . \tag{31}$$

In the following, we do not analyze the production-functions of firms, i.e. the physical processes of transforming inputs in outputs, as these are highly industry-specific. Instead we concentrate on cost-functions as a more general concept, i.e. we focus on the costs that must be incurred for producing some quantity of output.

Economies of Scale and Market Structure

In section A.2 , we have already relied on these definitions for characterizing the free-entry (or long-run) equilibrium in perfect competition. If entry is free and all firms have access to the same production-technology firms enter the industry until all firms produce at the **minimum efficient scale** and charge a price equaling their average costs at this output. Entry is treated in greater detail in section E.2 .

The existence of **sunk costs** of entry S affects the entry-decision (see section E.2) – not the decision on how much to produce after entry has occurred nor the decision to exit. We find that the higher the sunk cost, the fewer firms are in the market in equilibrium. A high sunk entry cost requires that each firm that enters subsequently earns a fair bit of profit from its operations to repay the initial entry expense. This can only happen if the number of firms that enter is small so that competition is weak and price can rise above marginal (and average) cost.

Figure 5 also illustrates the concept of **economies of scale**. Whenever marginal cost is less than average cost ($q<q_{MES}$), an expansion of output will lead to a reduction in average cost. Conversely, when marginal cost is greater than average cost ($q>q_{MES}$), an expansion of output will lead to an increase in average cost (**diseconomies of scale**). At the minimum of the average cost-function (q_{MES}, minimum efficient scale), average and marginal costs intersect.

$$\begin{aligned} \frac{dAC(q)}{dq} &= \frac{c(q)\cdot q - C(q)}{q^2} \stackrel{!}{=} 0 \\ c(q_{MES}) &= AC(q_{MES}) \end{aligned} \tag{32}$$

The presence of (dis)economies of scale reflects the underlying technology and, thus, affects the market structure. Some factors of production simply cannot be scaled down to small levels of production.

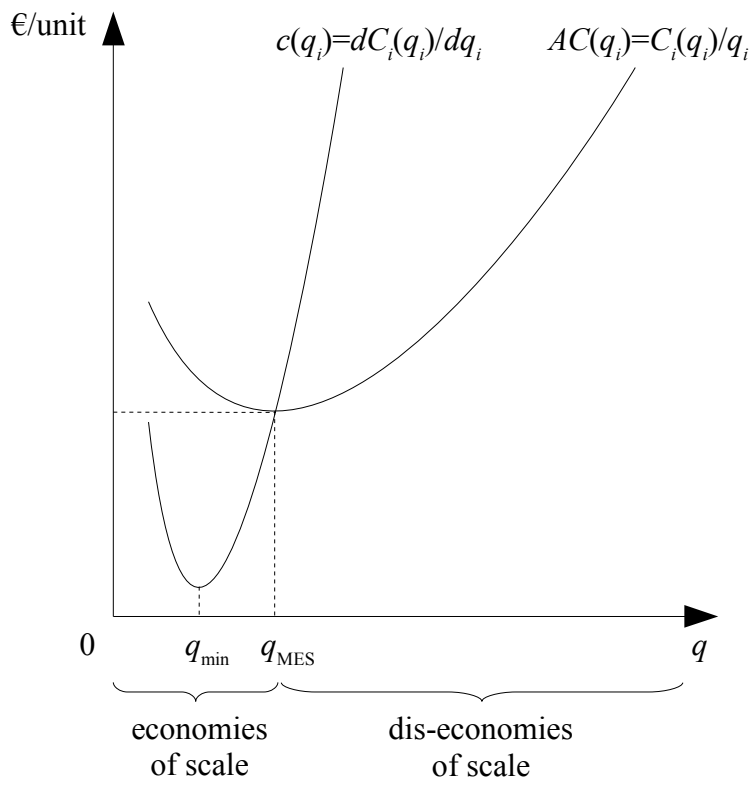


Figure 5: Minimum Efficient Scale

The above discussion suggests that the existence of economies of scale can be inferred from the ratio of average costs and marginal costs. Therefore, let us define a scale economy index *SI* as follows.

$$SI = \frac{AC(q)}{c(q)} \tag{33}$$

In case of economies of scale, *SI* takes values larger than 1. In such a setting, a 1 percent increase in output is associated with a less than 1% increase in costs. In case of diseconomies of scale, *SI* takes values smaller than 1. Increasing output by 1 percent now leads to more than a 1 percent increase in costs. At the minimum efficient scale, *SI* takes a value of 1.

One **source for economies of scale** is the presence of large fixed costs that are distributed on more units of output when the quantity sold rises. Moreover, a greater output also permits, e.g., a greater division of labor, which in turn permits specialization and more efficient production. This effect can be seen for outputs smaller than q_{min} where marginal costs decline with a higher output-quantity.

Now, suppose marginal costs are constant ($c(q)=c$) and production requires fixed costs *F*. It is easy to verify that in this case the scale economy index *SI* is positive for any output q, i.e. economies of scale are global in this market. If scale economies are global then the market is a

natural monopoly. The term "natural" is meant to reflect the implication that monopoly is an almost inevitable outcome for this market because it is cheaper in such cases for a single firm to supply the entire market than for two or more firms to do so, i.e. costs are **subadditive**. For symmetric firms this is shown by equation (34).

$$\begin{array}{rcl} C(q) & < & n \cdot C(q/n) \\ F + c \cdot q & < & n \cdot (F + c \cdot q/n) \end{array} \tag{34}$$

More generally, we can state that the greater is the extent of scale economies – i.e. the larger the minimum efficient scale – the fewer firms can operate efficiently in the market.

Economies of Scope and Multiproduct Firms

Firms often produce more than one product, e.g. several varieties of this product (different types of breakfast cereals) or even more dissimilar products such as operating systems and office-software. Analyzing the relationship between the output of multiproduct firms an the costs of production is rarely easy because the production of such goods can be subject to **economies of scope** (Pepall et al. 2008: ch. 4.3).

Economies of scope are said to be present whenever it is less costly to produce a set of goods in one firm than it is to produce that set in two or more firms. To see this, suppose that the cost of the joint production of two goods is $C(q_1,q_2)$. The costs of producing each good separately are $C(0,q_2)$ and $C(q_1,0)$. The production of these goods is subject to economies of scope when the condition

$$C(0,q_2) + C(q_1,0) > C(q_1,q_2) \tag{35}$$

applies. One example for the existence of economies of scope is the airline industry. About 50% of the freight that is transported via airplane is carried in passenger aircrafts. This is cheaper than using different aircrafts and, thus, allows airlines to charge comparably low prices on passenger-tickets.[16] The degree of economies of scope can be measured by the index *SC*.

$$SC = \frac{C(0,q_2) + C(q_1,0) - C(q_1,q_2)}{C(q_1,q_2)} \tag{36}$$

$SC>0$ means that production involves economies of scope. $SC<0$ would indicate diseconomies of scope.

The concept of scope economies provides the central technological reason for the existence of multiproduct firms. Scope economies may give rise to scale economies where we might not have expected any to exist. Looking at the production of only one product may not indicate any scale

16 http://www.spiegel.de/spiegel/print/d-74948197.html

economy effects. However, if producing more of one product lowers the cost of producing another, then the firm may be able to lower its costs per unit of aggregate output.

Economies of scope can arise for two main reasons. First, particular outputs share **common inputs**. This is the case in the above airline-example where the same planes are used to transport both passengers and freight. Second, economies of scope can also arise because of **cost complementarities**, i.e. producing one good lowers the cost of producing a second good. E.g., consider a refinery for crude oil. It is technically impossible to produce only gasoline or gasoil/diesel from crude oil. Both fuel must be produced together. It is cheaper to produce both instead of producing diesel only and disposing some part of the gasoline. As a consequence, a high demand for diesel increases the costs of producing diesel.[17]

Therefore, the existence of economies of scope may increase concentration in an industry, as a firm can produce a great variety of products (e.g., imperfect substitutes that belong to the same relevant market) more efficiently than each good separately.

The Productive Inefficiency of a Monopoly

Above, we have defined productive efficiency as production of some output-quantity $\bar{q}$ at the lowest possible costs. In this subsection, we show that a dominant firm (here: a monopolist) does not only cause a welfare loss (see Figure 7) because of charging prices above marginal costs but may also cause an additional loss in welfare (see Figure 14) because of productive inefficiency (Motta 2004: ch. 2.3).

Consider the situation from Figure 7. If the monopolist produces efficiently at marginal costs $c(q)$ it charges a price p_m and sells quantity q_m. The resulting welfare-loss in comparison to perfect competition is the sum of the areas C_1, C_2, and C_3. Now, suppose the monopolist produces at inefficiently high marginal costs $c'(q)$, charges a price p'_m, and sells a quantity q'_m. Selling a lower quantity at higher prices causes an additional loss in consumer surplus (D_1+D_2). Moreover, selling a lower quantity that is produced at higher costs reduces the monopolist's producer surplus (E_1+E_2).

Now, the relevant question is: "Why should a monopolist produce inefficiently at marginal costs $c'(q)$ rather than lower marginal costs $c(q)$?" There are three main answers to this question. First, managers of a monopolistic firm have less incentive to make effort (**managerial slack**) while "competitive pressure makes organizations internally more efficient by sharpening incentives to avoid sloth and slack" (Vickers 1995: p. 1). Second, when competition exists, more efficient firms will survive and thrive, whereas less efficient firms will shut down (**Darwinian mechanism**). If a monopoly exists, the market will not operate any selection and an inefficient firm is as likely to

17 http://www.spiegel.de/spiegel/print/d-72462692.html

survive as an efficient one. Third, competition to innovate is a major source of gains in productive efficiency over time (**dynamic efficiency**). In the following, I concentrate on the issues of managerial slack and competitive selection of efficient firms. The issue of dynamic efficiency will be discussed somewhat further below.

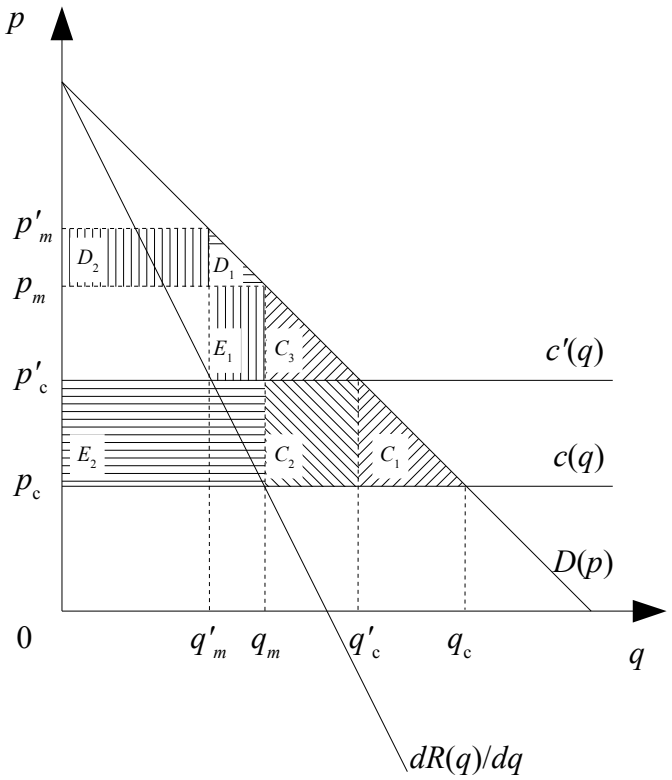


Figure 14: Additional Loss from Productive Inefficiency

Concerning the argument of **managerial slack** Vickers (1995: p. 7) notes that already "Hicks (1935) remarked that 'the best of all monopoly profits is a quiet life'. An important element of the quiet life is slack or 'X-inefficiency' (see Leibenstein 1966) within an organization inefficiently low levels of effort by its members to reduce costs, improve quality, and introduce new ways of doing things, and correspondingly high levels of leisure." Sometimes, even managers hold the opinion that other managers are more interested in prestige and money than in their job's challenges.[18] The lack of competition prevents performance comparisons among firms and thus precludes efficient production.

18 http://www.spiegel.de/wirtschaft/0,1518,632825,00.html

This can be shown based on a **principal-agent-model** (Vickers 1995: p. 9). Assume that x_t denotes the measured performance of a firm's manager (agent A) in period t. This performance is measured by the owner of the firm (principal P) and consists of the manager's effort e_t, his ability a and a measurement error ε_t.

$$x_t = e_t + a + \epsilon_t \tag{37}$$

Principal P offers agent A a wage contract $w_t(x_t)$ that consists of a base component α and a variable (effort-specific) component βx_t.

$$w_t(x_t) = \alpha + \beta x_t \tag{38}$$

The principal is interested in choosing the wage contract such as to maximize his payoff

$$max \;\; x_t - w_t(x_t) \quad . \tag{39}$$

The agent is interested in maximizing his expected utility from receiving wage $w_t(x_t)$ minus the cost of his effort $c_t(e_t)$

$$max \;\; u\left(w_t(x_t) - c_t(e_t)\right) \tag{40}$$

where the cost of his effort is defined as

$$c_t(e_t) = 0.5\, e_t^2 \quad . \tag{41}$$

A difficulty comes in by the imperfect measurability of the agent's performance x_t. The principal cannot decide with certainty whether a good performance-measure is a result of the agent's effort, his good overall quality, or because of a mis-measurement. This causes a **moral hazard**-problem, i.e. for a given payment the agent tries to exploit the measurement error in a way such as to minimize his effort. Under the assumption that the agent is risk-averse, it can be shown that the principal performs best by offering the manager no lump-sum payment (β_t=0) but make his wage depend on the measured performance x_t. Moreover, it can be shown that the moral hazard-problem is the more severe the less accurate x_t can be measured, i.e. the higher is the variance of the measurement error ε_t. The existence of competition may be assumed to reduce the variance of this measurement-error because the performance of a manager in firm i can be compared to the performance of a manager in firm j. This reduction in the variance of the measurement error reduces the moral-hazard problem and, thus, raises the manager's effort (to produce at the lowest cost possible).

The **Darwinian Mechanism**-argument goes as follows: In an industry with more efficient and less efficient firms, competition forces the inefficient firms to exit. Thus, welfare improves because output is produced at a lower cost (Motta 2004: ch. 2.3.2.3). Moreover, it is possible to show that it can be profitable for productively efficient outsiders to enter an industry and drive

inefficient incumbents out of the market (Vickers 1995: 15 – more information on entry is provided in section E.2).

Vickers (1995: p. 2) provides an example from the telecommunications-industry, which shows that issues of productive efficiency are relevant for modern competition policy.

> *"Consider a question that arose in the 1990 review of the UK government's previous policy of giving British Telecom and Mercury a protected duopoly of (non-mobile) public network operations. The issue was how much to open up the industry to competition. The existence of scale economies provided a possible case for limiting competition, but, as the telecommunications regulator put it 'some competition between networks is likely to be desirable because monopoly suppliers do not normally operate at the greatest possible level of efficiency'. Should market forces be allowed to resolve the tradeoff, or should government intervene to strike a balance, say by licensing just one or two additional competitors? Is there a danger that 'too much' competition might be self-defeating, leading eventually to the return of monopoly?"*

C.2 Dynamic Efficiency and Incentives to Innovate

Above, we were concerned with productive efficiency, i.e. producing with the most efficient technology available. In this section, we are concerned with **dynamic efficiency** (Motta 2004: ch. 2.4), which refers to the extent to which a firm introduces new products or processes of production (**process innovations**). We will particularly focus on process innovations.

The relationship between competition and dynamic efficiency is an important one. For example, Schumpeter's concept of creative destruction – because introducing new products or processes inevitably means the destruction of old ones – rests on the proposition that market power is necessary to spur innovations. With regard to welfare, Schumpeter considers competition even more important for creating new products and production technologies than for achieving allocative efficiency (see, e.g., Vickers (1995: p. 16) or Pepall et al. (2008: ch. 22)).

A central result of this section is that some extent of market power is necessary for motivating firms to invest in (process) innovations.[19] The idea underlying this result is the following. If firms compete only moderately, each firm makes enough profits to finance investments in innovations that give them a competitive advantage over their rivals. However, if

19 This interview from 1986 with then IBM-CEO John F. Akers provides a nice illustration for the topic of innovation and market power. Moreover, a comparison of the situation in 1986 to that in 2011 illustrates the impact of innovations on both producers and consumers: http://www.spiegel.de/spiegel/print/d-13521624.html

firms compete fiercely their (expected) profits are possibly too low for financing investments.

However, too much market power reduces the incentive of a firm to invest in innovations. The idea of this result is the following. A firm with substantial market power cannot gain even (much) more market power by innovating and, thus, is less keen to invest in innovations. One proposed example for this effect is Microsoft[20]:

> *In February 2010, Dick Brass – a former vice president of Microsoft – in an article in the New York Times proposed that Microsoft's productive efficiency is not at its peak by saying that "Microsoft has become a clumsy, uncompetitive innovator". Brass states that Microsoft's profits "come almost entirely from Windows and Office programs" where Microsoft still has a position quite close to monopoly.[21] However, in his view Microsoft overly relies on these products and fails in inventing new ones such as tablet-PCs or e-books.*

Combining the above findings leads us to expect that there is an **inverted u-shaped relation between market power and innovation**. There are only small incentives to innovate when market power is either very small or very large. The incentive to innovate is at its maximum for intermediate levels of market power. We will analyze the two parts of this statement in turn after having defined some key terms.

A Taxonomy of Innovations

The research-component in **research and development** (R&D) consists of two parts basic research and applied research. **Basic research** includes studies that will not necessarily lead to specific applications but aim to improve our fundamental knowledge in a manner that may subsequently be helpful in a range of activities, for example the development of lasers at 405 nm wavelength. **Applied research** involves substantial engineering input and is aimed at a more practical and specific usage than basic research. Its goal is the creation of a prototype of a product, for example the development of the Blu Ray-technology. The **development**-component of R&D takes the prototype and develops a product that can be used by consumers and that is ready for (mass) production, for example Blu Ray-players.

We also distinguish between **process innovations**, which is the discovery of cheaper methods for producing existing goods, and **product innovations**, which is the creation of new goods. This section concentrates on process innovations.

20 http://www.nytimes.com/2010/02/04/opinion/04brass.html?_r=1&pagewanted=all

21 Also see http://www.spiegel.de/wirtschaft/0,1518,223335,00.html

Market power is small → Incentive to innovate rises with market power

We will provide evidence for this proposition on basis of a theoretical example. Consider an industry with demand

$$Q = a - p \tag{42}$$

and n ex-ante symmetric firms who produce a homogeneous product at marginal costs c_h and compete à la Cournot. Production does not require fixed costs. It is easy to check that a firm i in this market makes a profit

$$\pi_i = \left(\frac{a - c_h}{n+1} \right)^2 \quad . \tag{43}$$

Now, firm i is offered the opportunity to buy a technology that gives it the sole right to produce at marginal costs c_l (with $c_h > c_l$) while its competitors go on producing at marginal costs c_h. It is straightforward to show that i's profit becomes

$$\pi'_i = \left(\frac{a - nc_l + (n-1)c_h}{n+1} \right)^2 \quad . \tag{44}$$

Therefore, i's willingness to pay for this technology is the gain in profits that it can make by investing into this technology

$$\Delta \pi_i = \pi'_i - \pi_i \quad . \tag{45}$$

Firm i will buy the new technology when $\Delta\pi_i$ is larger than the investment F required to acquire the new technology.

It can be shown that $\Delta\pi_i$ increases as the number of firms decreases, i.e. when concentration increases. Hence, the incentive to innovate rises with the amount of market power. The reason for this is that in the case of fewer firms the innovator gains a larger additional market share than in the case of many firms. Therefore, amortization of the investment is easier when market power is high.[22]

As a further consequence of the innovation, producer surplus (i.e. the sum of all firms' profits) rises. Additionally, consumer surplus rises because the innovation leads to lower prices and, thus, a higher quantity sold. Since firms neglect the positive externality that their R&D-efforts impose on consumers, these efforts could be argued to be too low as compared to the efforts done

22 Pepall et al. (2008: ch. 22.3) provide a model on this issue where firms may decide on the size of investment in R&D. They also find that "the reduction in a firm's output that results from increasing the number of firms also reduces the marginal benefit that R&D spending yields to an individual firm." Hence, R&D-spending will fall as the number of firms rises.

by a social planner who is interested in welfare-maximization. One solution to this problem would be the payment of subsidies for research that are sometimes requested by business-leaders.[23]

23 http://www.spiegel.de/wirtschaft/unternehmen/0,1518,728594,00.html

Market power is large → Incentive to innovate declines with market power

Solving the above Cournot-model

Given demand-function (42), the asymmetric firms' profit-functions can be written as follows.

$$\begin{aligned} \pi_l &= (a-Q_{-l}-q_l-c_l)\cdot q_l \quad \text{with} \quad Q_{-l}=(n-1)\cdot q_h \\ \pi_h &= (a-Q_{-h}-q_h-c_h)\cdot q_h \quad \text{with} \quad Q_{-h}=(n-2)\cdot q_h+q_l \end{aligned}$$

Maximizing with regard to quantities yields the following reaction-functions.

$$\begin{aligned} q_l &= \frac{a-c_l-(n-1)q_h}{2} \\ q_h &= \frac{a-c_h-q_l}{n} \end{aligned}$$

Solving for the Nash-equilibrium of this game yields the following results.

$$\begin{aligned} q_l &= \frac{a-n\cdot c_l+(n-1)c_h}{n+1} \\ q_h &= \frac{a-2\cdot c_h+c_l}{n+1} \\ Q &= \frac{n\cdot a-c_l-(n-1)\cdot c_h}{n+1} \end{aligned}$$

$$P=\frac{a+c_l+(n-1)\cdot c_h}{n+1}$$

$$\begin{aligned} \pi_l &= \left(\frac{a-nc_l+(n-1)c_h}{n+1}\right)^2 \\ \pi_h &= \left(\frac{a-2\cdot c_h+c_l}{n+1}\right)^2 \end{aligned}$$

Equation (43) is found from the above profit-functions by setting $c_l=c_h$.

Exercise: Calculate the above equilibrium and show that (i) the incentive to innovate and (ii) welfare rise in the number of firms. Do this based on a numerical example with a=110, c_l=5, and c_h=10. Also show that a downturn in the business cycle (demand-parameter a shrinks) lowers the incentive to innovate.

However, the incentive to innovate in a monopoly can be lower than in a Cournot-duopoly. To see this, consider that a monopolist's profits in the above case are

$$\pi_i = \left(\frac{a-c}{2}\right)^2 \quad . \tag{46}$$

It can be shown that in some cases the incentive to innovate $\Delta\pi_i$ in a monopoly is lower than in a Cournot-duopoly. This is particularly the case when the innovation reduces a firm's costs a lot. The reason for this is that the innovation "only" lowers the monopolist's unit-production costs and allows for attracting some more customers because of reduced prices. However, a duopolist gains from these two effects plus a **business-stealing effect**, i.e. the firm increases its market share at the cost of its competitor.

In a nutshell, the intuition behind this result is as follows. Competition pushes firms to invest, in order to improve their competitive position relative to their rivals. The absence of competition reduces this incentive to innovate, and this in turn means that a monopolist will be less efficient (less innovative) than firms which operate under competition.

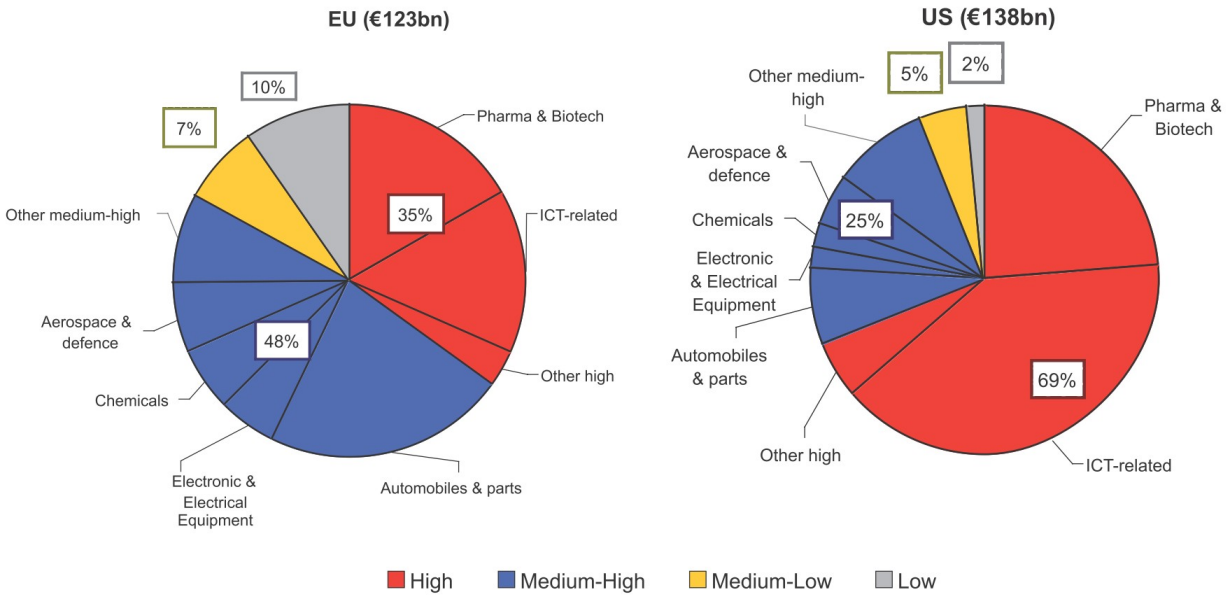
Empirical findings and the inverted U-relationship

Taking together the above findings we conclude that some intermediate levels of competition might be optimal for innovations and productive efficiency. This result is confirmed by empirical studies that do find an inverted U-relationship between competition and innovation (see Pepall et al. (2008: ch. 22.4) for an overview).

These studies often find that R&D-efforts rise with industry-concentration – and, thus, with firm-size – up to moderate levels of concentration. A further increase in concentration tends to decrease R&D-efforts. In this context, one should note that these relationships are highly industry-specific because some industries are more research-intensive than others. One quite research-intensive industry is the pharmaceuticals-sector, which also is a good example for a moderately concentrated industry.[24] The below table gives an overview on research-intensive sectors in the EU as compared to USA (EC 2010: p. 10).

24 http://www.manager-magazin.de/unternehmen/industrie/0,2828,727239,00.html

Figure S4. R&D investment shares by sector group



Sectors are split into four groups according to the R&D intensity of the sector worldwide:

***High R&D intensity** sectors (intensity above 5%) include e.g. Pharmaceuticals & biotechnology; Health care equipment & services; Technology hardware & equipment; Software & computer services.*

***Medium-high R&D intensity** sectors (between 2% and 5%) include e.g. Electronics & electrical equipment; Automobiles & parts; Aerospace & defence; Industrial engineering & machinery; Chemicals; Personal goods; Household goods; General industrials; Support services.*

***Medium-low R&D intensity** sectors (between 1% and 2%) include e.g. Food producers; Beverages; Travel & leisure; Media; Oil equipment; Electricity; Fixed line telecommunications.*

***Low R&D intensity** sectors (less than 1%) include e.g. Oil & gas producers; Industrial metals; Construction & materials; Food & drug retailers; Transportation; Mining; Tobacco; Multi-utilities.*

Source: The 2010 EU Industrial R&D Investment Scoreboard. European Commission, JRC/DG RTD.

Moreover, empirical studies on the link between innovation and concentration show correlations between these variables but have a hard time identifying causal links. Therefore, it is not necessarily clear if research-intensive industries cause a concentrated market-structure, or if a high concentration is a reason for firms doing research.

Bertrand-competition and appropriability

The above results hinge on the assumption that the innovator is the only firm that may use the new technology. Results change when this full **appropriability** is not given and other firms can also use the innovative technology.

To see this, consider a Bertrand-duopoly where the firms produce at marginal costs c_h. In period 1, they set $p=c_h$ and make zero economic profits. At the beginning of period 2, they may decide to invest an amount F in a new technology which enables them to produce at lower marginal costs c_l. Either firm must now decide whether to invest in the new technology or not. This game is shown in the below strategy-matrix. When neither firm invest, both firms make zero economic profits. When both firms invest, they set $p=c_l$ and make zero economic profits but have to incur the costs for acquiring the new technology. When only one firm invests, it steals all business from its competitor and makes a positive economic profit. This game is at equilibrium when only one firm invests.

		firm 2	
		innovation	*no innovation*
firm 1	*innovation*	$\pi_1=-F$ | $\pi_2=-F$	$\underline{\pi_1>0}$ | $\underline{\pi_2=0}$
	no innovation	$\underline{\pi_1=0}$ | $\underline{\pi_2>0}$	$\pi_1=0$ | $\pi_2=0$

The pricing-strategy of the innovating firm in either equilibrium can take two forms as shown in Figure 15. First, the innovator's monopoly price given its lower costs ($p^*(c_l)$) might be below the marginal costs of the non-innovating firm ($p^*(c_l)<c_h$). This is called a **drastic** (or major) **innovation**. The innovator sets its price at the monopoly-level and drives its competitor out of the market. Second, the innovator's monopoly-price might be above the other firm's marginal costs ($p^*(c'_l)>c_h$). This is called a **non-drastic** (or minor) **innovation**. In this case, the innovator sets its price just below its competitor's marginal costs ($p'=c_h-\varepsilon$, with $\varepsilon \rightarrow 0$) and drives its competitor out of the market.

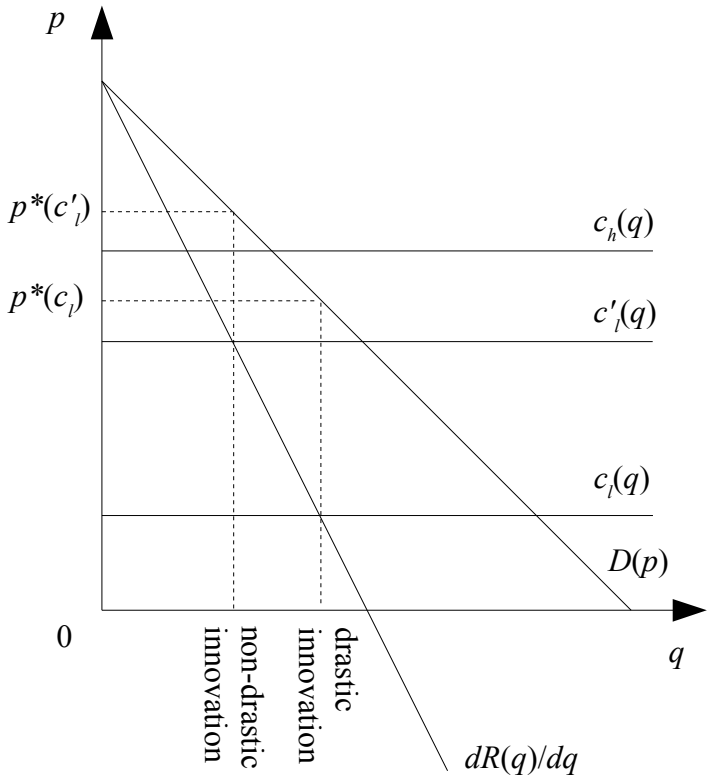


Figure 15: (Non-)drastic innovations

Now, consider the case where appropriability is not given. In this case, the non-innovating firm can appropriate and use the innovative technology without having paid for it. As a result, prices will always be at the level of marginal costs with the firms making zero economic profits. However, the innovator incurs the investment *F* of this innovation. This is shown in the below table. The only equilibrium of this game is that no innovation occurs.

		firm 2	
		innovation	*no innovation*
firm 1	*innovation*	π_1=-*F* | π_2=-*F*	π_1=-*F* | π_2=0
	no innovation	π_1=0 | π_2=-*F*	$\underline{\pi_1}$=0 | $\underline{\pi_2}$=0

This is, of course, an extreme case. However, one may easily imagine examples where one firm's innovation has positive externalities (**spillovers**) on the other firm's production costs. E.g. by poaching engineers of the innovating firm, the non-innovating firm can at least partially benefit from this research. In this case, we expect innovation to be in between the above cases, i.e. R&D occurs but investments in R&D are lower than in the case of full appropriability.

You may now see the value of patents and other forms of intellectual property rights. They ensure (for some time) that an innovator can appropriate the gains from his innovation. This is likely to spur innovations.

Lessons Learned

After reading this section you should be able to answer the following questions.

1. What is meant by the terms *economies of scale* and *economies of scope*? How does the existence of scale and scope economies affect market structure.
2. Why is a monopolist likely to produce at inefficiently high costs?
3. Provide an explanation for the inverted U-shaped relationship between market power and dynamic efficiency.
4. How can patents help to improve dynamic efficiency?

References

European Commission (2010). "The 2010 EU Industrial R&D Investment SCOREBOARD." http://iri.jrc.ec.europa.eu/research/scoreboard_2010.htm

Motta, M. (2004). "Competition Policy – Theory and Practice." Cambridge University Press: Cambridge

Pepall, L. and Richards, D. and Norman, G. (2008). "Industrial Organization – Contemporary Theory and Empirical Applications." 4th edition. Blackwell Publishing: Malden, MA

Vickers, J. (1995). "Concepts of Competition." Oxford Economic Papers. Vol. 47 No. 1, pp. 1-23

D COMPETITION POLICY

In this section, we provide an overview on competition policy.[25] This includes an overview on important competitive concerns that may call for interventions into the working of markets. These (economic) foundations of competition policy are provided in subsection D.1 . The interventions must be based on legal statutes and are performed by competition authorities and courts. In subsection D.4 , we present an introduction to the most relevant competition laws in different jurisdictions. These are (i) USA as a country with a long history of legal assessments of competitive issues, (ii) Germany where this lecture is held, and (iii) Europe whose laws complement and shape national laws in the European Union. For the latter reason, a focus is put on European competition law.

D.1 Foundations of Competition Policy

This subsection shows why many jurisdictions have adopted competition laws and, thus, engage in competition policy. The main reason is that real **markets are not necessarily perfectly competitive** and, thus, not allocatively efficient. In this context, one may think of, e.g., horizontal agreements or other market imperfections (**competitive concerns**) that allow firms to exercise market power and decrease efficiency. In the following, we provide a **definition** of the term competition according to competition laws. Moreover, we show that competition laws should be designed that – given the available budget – both an under- and an over-enforcement are avoided.

The Necessity of Competition Policy and the Ideal of Perfect Competition

"In a world of **perfect competition** [*emphasis added*], life is good. Firms can enter and exit markets instantly and without cost, products are homogeneous, and everyone is perfectly informed. Firms are so numerous that none of them is large enough to influence prices by altering output, and all act independently. Supplier competition for sales thus drives prices for products and services down to the costs of providing them" (Elhauge 2008: p. 1). If demand increased or costs decreased the incumbent-firms would earn supranormal profits. This would induce entry until economic profits have been driven down to zero. A decrease in demand or an increase in costs would induce exit. Such a market would create allocative efficiency.

"In the real world, life is regrettably imperfect. Entry, exit or expansion are costly and take time. Products vary by brand or attributes and information is imperfect. Economies of scale mean many markets cannot sustain a large enough number of firms to leave each without any incentive to

25 A more detailed introduction in German is provided by Schmidt (2012).

consider the effect of its decisions on market prices. But despite such unavoidable realities, typical markets are workably competitive in the sense that they produce results that are fairly close to perfect competition, at least in the long run. In any event, perfect competition provides an aspiration and useful benchmark that helps identify the sort of interferences with market mechanisms that should most concern antitrust law" (Elhauge 2008: p. 1). The ideal of perfect competition as well as some of these market imperfections are presented in Part I of this reader. Additional imperfections as well as the foundations of competition law, i.e. competitive concerns, are presented in Part II.

"Principally, the need for competition law intervention arises when there is a market failure – so long as markets remain competitive, consumers benefit from low prices and innovative products because firms are driven by the desire to maximise profits and sell as many goods as is economically feasible at the lowest price. The market system is perceived to be the ideal mechanism through which the fundamental economic questions are answered: what goods to produce, how many to produce, and how to distribute them" (Monti 2007: p. 55). Putting it broadly, the **aim of competition policy** is to promote the competitiveness of markets and prevent distortions of market outcomes. In Europe, this is specified in Articles 3 and 119 TFEU (Treaty on the Functioning of the European Union): "The Union shall have exclusive competence in [...] the establishing of the competition rules necessary for the **functioning of the internal market** [*emphasis added*]" according to "the **principle of an open market economy with free competition** [*emphasis added*]."

Getting more specific on what competition policy *should* achieve is difficult because the role of competition policy is controversially debated. We follow Monti's (2007: p. 2) position that "it is impossible to identify the 'soul' of competition law; the most that can be done is to show that there are different, equally legitimate opinions of what competition policy should achieve. Moreover, within each country, the purposes of competition law can change over time, even without an amendment to the legislative texts. [...] Understanding competition law thus is not only about dissecting legislative texts [...] but is also about understanding the particular forces that have influenced the direction of competition policy at particular times." Monti (2007: p. 4) considers it "helpful to think about the factors that influence competition law and the decisions that stem from those rules on the basis of the interactions of three components: a *political* decision about the aims of competition law; an *economic* theory about how markets behave, how and when they fail, and how market failures may be remedied; and the *institution* in charge of enforcing competition law." Based on this classification, we pose three questions:[26]

1. **Political question**: Should competition policy (only) be concerned with **economic welfare**, i.e. maximizing productive, allocative, and dynamic efficiency, or should it be used to

26 Monti (2007: ch. 1.2) illustrates the three questions based on an assessment of the proposed *de Havailland*-merger.

pursue a variety of **other goals**, "for example to maximise economic freedom, preserve employment, promote national champions, facilitate restructuring, protect small firms, safeguard cultural values, conserve the environment, and so on" (Monti 2007: p. 4).

2. **Economic questions**: Should competition policy be concerned with creating industry structures (**structural approach**) that make adverse effects to welfare unlikely, e.g. promoting the number of firms in an industry? Alternatively, it could place more weight on the assessment of further effects, e.g. the ease of entry into a market, i.e. stressing the self-healing qualities of a market (**Chicago School approach**). As a second question, it must be answered which **welfare-measure** should be used, i.e. whether competition policy shall be concerned with maximizing total welfare or consumer surplus.
3. **Institutional question**: It should be decided whether the enforcement should be done by **independent (judicative and/or adminstrative) bodies** or by **legislative/governmental bodies**. In the earlier case, one has to decide whether competition authorities may decide cases (in the first instance) or whether the decision must be made by a court. The latter case is most relevant when competition policy is designed to achieve goals beyond maximizing economic welfare.

In this reader, we concentrate on the economic questions and analyze market behavior, market failure, as well as possible remedies. The political and institutional questions are addressed only briefly. Before getting to these points, we present an overview on the main competitive concerns that require intervention by a competition authority.

Competitive Concerns

"The first major concern is that firms might agree to avoid competing with each other, thus elevating prices above cost and increasing their profits to supracompetitive levels" (Elhauge 2008: p. 2). The economic consequences of such collusive agreements are illustrated in sectionError: Reference source not found. Aspects of other (horizontal and vertical) co-operation agreements are presented in H .

"A second concern is that one firm might individually be large enough to raise prices by reducing output [see section A.3]. […] True monopolists are rare. More typical is what economists call a **dominant firm** [*emphasis added*], which is a firm that is much larger than the other firms because it has lower costs or a better product. A dominant firm also has incentives to price above cost but is somewhat constrained by the ability of the other firms to offer the product at their costs. […[27]]

27 The essence of competition policy is summarized by an article of the US-American satirical newspaper *the ONION*

The mere possession of monopoly or dominant power need not, however, be a concern. […] Dominant market power normally reflects the fact that a firm is more efficient because of some cost or quality advantage over its rivals. If a firm has acquired that efficiency advantage through productive investments in innovation, physical capital, or organization, then the additional profits it is able to earn might reasonably be thought to provide the right reward for that investment […]. Typically the antitrust laws are instead focused on anticompetitive conduct that is used to obtain or maintain monopoly or dominant market power at levels that were not earned through productive efforts. A dominant firm has incentives to use anticompetitive conduct to exclude rivals from the market, impair rival efficiency, or impede the sort of rival expansion and entry that would drive down prices toward more competitive levels [see section I]. […]

Firms with market power might likewise have incentives to enter into agreements with suppliers or buyers to try to exclude rivals, diminish their efficiency, or impede their expansion or entry. Because these agreements are up or down the supply chain, they are generally called **"vertical agreements"** [*emphasis added*; see section H.4], in contrast to the **"horizontal" agreements** [*emphasis added*; see section H.3] entered into by rivals at the same level. […]

[A further concern] is that rivals might merge or combine into one firm. **Horizontal mergers** [*emphasis added*; see section G.2] can have anticompetitive effects if the resulting firm has monopoly or dominant market power, or the structure of the rest of the market means the merger will create an oligopoly or exacerbate its ability to coordinate on higher prices [see section B]. The difficulty is determining when this is the effect of a merger and whether the merger is justified by any greater efficiencies it might create. **Vertical mergers** [*emphasis added*] between firms up and down the supply chain [see section] raise issues similar to vertical agreements that might exclude or impair rival competition" (Elhauge 2008: p. 2-4).

Objectives and Design of Competition Policy

The central aim of competition policy is to foster competition and, thus, enhance welfare. This requires a **definition** of the concept **of competition**. In short, competition can be defined as

1. rivalry among firms, which relates to firms' conduct,
2. a competitive market-outcome in terms of welfare (performance), or by
3. the ordoliberal view of competition as economic freedom (Monti 2007: ch. 2.2.1).

In the first definition of competition as **rivalry among firms**, it is regarded as a means (i.e. a particular behavior) to achieve a number of desirable ends (e.g., welfare). Based on this definition it

that in its issue 38-03 titled "Judge Orders God To Break Up Into Smaller Deities". The article can be read online at: http://www.theonion.com/articles/judge-orders-god-to-break-up-into-smaller-deities,404/

is hard to judge whether there is a distortion of competition, as it is difficult to evaluate firms' behavior. Therefore, economists usually prefer to evaluate the ends of competition, i.e. the **market-outcome**. This neoclassical concept uses welfare and, thus, efficiency as clearly defined concepts for measuring competition. In this context, competition contributes to allocative efficiency and can enhance productive efficiency in three ways: "first, competitive pressures give firms greater incentives to reduce costs to being taken over or rendered insolvent by more efficient firms; second, and related, competition is a means of selecting the more efficient firms, which will thrive; third, the competition to innovate is a source of productive efficiency. So competition may also be a means to achieve dynamic efficiency" (Vickers 1995; cited according to Monti 2007: p. 55). Under the **ordoliberal model**, the aim of competition policy is the protection of individual economic freedom of action as a value in itself. It may also be called disciplined pluralism, i.e. all individuals should allow all individuals to participate unhampered by the economic power of the others (pluralism) while the risk of monopolies or cartels necessitates laws to sustain economic freedom (discipline).

"In [the light of ordoliberalism] the discipline of the market is as fundamental as contract law or property rights. [...] liberal discourse is based on the values of personal liberty and equality; in contrast the neo-classical definition of competition is embedded in utilitarian and laissez-faire economic philosophy, where intervention is called for as a second best, when the market fails to deliver economic efficiency [...]. [In contrast], ordoliberalism necessitates rules that safeguard economic freedom in the marketplace by imposing obligations of fair conduct and suppressing economic power" (Monti 2007: p. 23-24). However in many cases, the design practice of competition policy is the same according to the neoclassical and the ordoliberal view of competition.

The above competitive concerns illustrate that competition and economic welfare can be enhanced by means of competition policy. This raises the question how competition laws should be designed in order to achieve this goal in the best way possible. First, the assessment could be based on ***per se* rules**. Such rules are quite strict and prohibit (supposedly) anticompetitive practices without any case-by-case inquiry into their effects. Such strict rules come along with two types of possible errors. On the one hand, a rule may prevent efficient behavior, e.g., it prohibits a merger whose pro-competitive effects would have outweighed its anti-competitive effects. This causes an **over-enforcement** which is also known as type 1 error, α error or false positive. On the other hand, a rule may fail in punishing inefficient behavior, e.g. an anticompetitive agreement remains unprohibited because the firms' combined market share is small and, thus, falls within a safe-harbor rule. This causes an **under-enforcement** which is also known as type 2 error, β error or false negative.

Rules are advantageous, as they are easy to apply and prevent arbitrary decisions by a competition authority. As a disadvantage, rules create false positives and false negatives. Therefore, competition policy should be designed such that a case by case examination of supposedly anticompetitive behavior is possible, i.e. competition authorities should be allowed to apply a **rule of reason**. An assessment based on a rule of reason is required when the legislative body sets a standard rather than a rule. Because more thorough investigations are costly they should only be performed when their benefits outweigh the (expected) costs that come along with false positives and false negatives.

D.2 Economics and Competition Policy

In particular in USA, antitrust laws in the 2nd half of the 20th century were affected by three types of economic thinking. These are the **structure-conduct-performance paradigm** (Harvard School), the **Chicago School**, and the **post-Chicago School**. These schools of economic thought are presented below. The main difference between these schools are their respective definitions of competition (Kling and Thomas 2007: p. 7-17). Thus, they present alternative views of what market power is and how it can be sustained. Therefore, this section should be read in the light of the above section E.2 on market power.

The Harvard School and the Structure-Conduct-Performance Paradigm

"The Structure-Conduct-Performance (hereinafter SCP) paradigm was prominent in the 1950s and 1960s. It suggests that by observing the structure of a market, inferences can be drawn as to how firms conduct themselves, and this allows one to evaluate the market's economic performance. On this basis, certain market structures can be identified as being the cause of anticompetitive conduct, which in turn leads to poor economic performance.

Market structure is a term used to define the characteristics within which firms operate in a given product market. The principal characteristics are the number of firms and their size. Also relevant is the market power of the firms' customers, the ease with which new firms can enter the market and old ones can exit it" (Monti 2007: p. 57). These features are described in economic terms in section E . "**Conduct** describes the way in which the firms behave: what criteria they use to set prices (collusion, independently, or on the basis of consumer demand); how they decide on advertising and research and development expenditure. **Performance** is the yardstick by which the conduct of firms is measured. The standard measure for this is whether the firms enhance economic welfare. [...]

The theory predicts that the more closely the market in question approaches conditions of

monopoly the worse its performance [...], and the more closely the market approaches the conditions of perfect competition the better its performance" (Monti 2007: p. 57-58). These principles are explained in Part I of this reader. "The implications of the SCP paradigm for competition policy is that we can identify which market structures lead to anticompetitive results and devise a competition law that is designed to modify or prevent market structures that are linked with poor economic performance. The policy recommendation that flows from the SCP paradigm is that special attention should be given to those market structures most likely to reduce consumer welfare – monopoly and oligopoly. [... T]he paradigm sees the structure of an industry as the cause of market failure, remedies are designed primarily to alter the structure of the market, and to prevent incumbents from raising entry barriers. It follows that small industry deserves protection from larger firms: large numbers of small firms yield greater economic welfare than a small number of large firms" (Monti 2007: p. 59).

The Chicago School

Chicago scholars launched three central challenges to the SCP approach (Monti 2007: ch. 3.4).

1. They **refuted the connection between industry concentration and anticompetitive effects**. For example, they stressed that coordination is costly so that one cannot presume coordinated anti-competitive behavior in all concentrated industries. Furthermore, the statistical correlation between high concentration and high profit rates was not necessarily caused by anticompetitive behavior, but could be the result of economic efficiency. Finally, economies of scale did justify high levels of concentration.
2. Chicagoans also pointed out that **the relationship between structure and conduct is two-sided**. While the SCP-paradigm implies that market structure affects firms' conduct, the Chicago School stresses that firms' conduct may also affect the structure of the market. Therefore, a firm with a large market share may behave anti-competitively by cutting output. However, this invites new entrants which improves competition.
3. Chicago's criticism of the SCP definition of **entry barriers** is that the latter saw entry barriers whenever entry was more **difficult** (e.g., in the presence of economies of scale, product differentiation, and cost advantages). The question for the Chicago paradigm is whether entry is more **costly** for the new entrant. For Chicagoans the only entry barriers are property rights conferred by the government (e.g., if the incumbent owns the patent to a product necessary for entry).

Re-defining entry barriers in this way means that there can hardly be any concentrated markets. Hence, there would be no need for a competition policy aiming at eliminating market

concentration. These three critiques led to the basic principle of the Chicago approach: **markets normally cure themselves** and competition outcomes are likely without any significant government intervention. In the Chicago view, monopoly behavior is likely to attract new entry, thus the market is the best cure against concentrated markets. On the contrary, legal intervention may lead to inefficient results, and the costs of type 1 errors (= over-enforcement) reduce economic welfare.

Armed with these alternative models of economic behavior, Chicagoans criticized much of the application of competition law as misguided. In their view, the test for antitrust violations is not whether conduct injures competitors or excludes rivals, but whether the practices in question allow firms to reduce output and raise prices. The standard by which to judge antitrust violations shifts from an inquiry into market power to an inquiry about whether the practice in question is efficient. The orthodox Chicago position of the 1960s and 1970s identified **explicit price fixing and large horizontal mergers as the only antitrust problems**.

The Post-Chicago Paradigm

The post-Chicago paradigm relies on complex (quantitative and econometric) tools to determine whether there is a market failure (Monti 2007: ch. 3.5), and the specific features of an industry are crucial to determining market failures. The principal indicator of market failure is the presence of market power, defined by the ability to set prices above marginal cost (see section E). To some extent the potential for market failure is present in all markets, but is not presumed to occur by considering market structure, as in the SCP paradigm. It is rather the way in which firms act that can cause market failures.

Hence, the **industrial organization approach** to market power is concerned with the **strategic means** by which market power is exercised or created. The focus is on firms in imperfectly competitive markets, and how that conduct affects the reactions of other firms in the market. Such behavior is analyzed throughout this reader. The post-Chicago paradigm demands that markets should be studied in more detail before determining whether a practice is pro- or anti-competitive. The likelihood of market failure increases in the post- Chicago paradigm because of their recognition that strategic behavior, which may appear pro-competitive viewed statically (e.g. product improvement), may, if considered strategically, lead to the elimination of competition and to monopoly pricing.

One of the key methodological consequences of this paradigm is the rising importance of **empirical fact-finding**. Armed with fact-specific empirical evidence, the post-Chicago paradigm suggests that in addition to inefficient behavior, firms with market power can cause market failures

by taking steps to exclude entry: by deterring future entrants or by eliminating competitors. Monti (2007: p. 70) presents the Staples-case from the 1990s as an early example of empirical fact-finding:

> *"Increased empirical attention to the specific circumstances in an industry had an impact in the Federal Trade Commission v. Staples merger decision. The nub of the dispute centred on market definition: the parties to the merger (Staples and Office Depot, the two largest office superstore chains in the US) claimed that the relevant market was the sale of consumable office products through all retail outlets, where the firms held a combined market share of 5.5 per cent, so their merger posed no anticompetitive risks. However, the FTC defined the market as one for consumable office supplies sold through office superstores. From a Chicago School approach, this narrow market makes little sense: a pen is a pen wherever it is purchased, and as consumers shop around for the cheapest deal, any attempt by office superstores to raise prices will lead to a loss of sales to other retail outlets. This intuition about the consumer's shopping skills was however denied by the facts: there were three main office superstores in the US and in geographical areas where Staples faced no competition prices were 13 per cent higher than in markets where Staples competed with Office Depot and Office Max; similarly Office Depot's prices were well over 5 per cent higher in areas where it faced no competition. Moreover, the FTC constructed econometric models that demonstrated how little impact other retail outlets have on the pricing decisions of office superstores, and that if all three office superstores were to merge, prices would increase by 8.49 per cent. This econometric evidence led the FTC to conclude that the prices of goods in office superstores are affected primarily by the other office superstores, and that non-superstore competition is not a significant check on prices. Thus, pre-merger the three superstores already enjoyed a degree of market power, which the merger would enhance by eliminating a particularly aggressive competitor. The decision is significant for its use of econometric studies to identify a competition risk which on a cursory analysis, biased by presumptions about consumer reactions to higher prices, appeared unrealistic."*

The Effect of Economics on Competition Law

The above economic doctrines illustrate that decisions in competition cases require an economic model to determine the legality of behavior (Monti 2007: ch. 3.6). A **good model** would have the following three attributes: first, it **distinguishes** in a convincing manner **between pro- and anti-**

competitive behavior; second, the courts are able to **provide a remedy** which they believe will improve consumer welfare; third, the **rule is sufficiently simple** to apply in the courtroom. In other words, a good economic theory can be transformed into legal rules at acceptably low levels of type 1 (over-enforcement) and type 2 (under-enforcement) errors.

This view is rather pragmatical and relies on the observation that it is impossible to provide *the* one and only definition of competition (Kling and Thomas 2007: p. 13-14). In this context, competition laws should not contribute to creating a specific, competitive market *outcome*. They should rather be concerned with providing a level playing field for competitive *behavior*. Hence, the state would act as a referee who ensures that all firms play according to the rules. Economics provides some guidance what rules should be set and how existing rules should be interpreted.

In **USA** since the 2nd half of the 20th century, economics has had an effect on setting competition laws as well as decisions of specific cases. US-competition authorities, i.e. the Federal Trade Commission and the Antitrust Division at the Department of Justice, employ economists, and economic experts provide guidance to courts.

In contrast, **EC competition law** (Monti 2007: ch. 3.7) has not, yet, been affected by economic paradigms to the degree that US antitrust has. There are several explanations for this observation. First, most European Member States have **civil law traditions** that have been traditionally less receptive to the economic analysis of law. Economic analysis requires case-by-case reasoning, considering the potential effects of imposing liability. This type of legal reasoning suits the common law more easily than the civil law. The civil law culture favors reasoning according to legal categories, resolving problems by literal interpretation of the statutes. Second, the European competition authority – the **Directorate General Competition** – used to be **populated by more lawyers** than economists, which creates a bias against the use of economics and in favor of decisions that are legally defensible. Third, the goals of EC competition law were never purely economic. The law's **ordoliberal origins** favored economic freedom over efficiency. In the EC, the concern about economic freedom was motivated by a fear that without such freedom, the market would descend again into the cartelized state which contributed to the rise of totalitarianism in Europe, and the concern over market integration was a corollary to the importance placed by the EC Treaty on creating a single market.

The Role of Economists in Competition Policy

The role of economists in European competition law enforcement (the so-called **more economic approach**) has been described by the former chief economist at DG-Comp, Lars-Hendrik Röller. He (2005: p. 11) makes clear that the “question for effective enforcement is not one of “more” or

"less" economics, but rather what kind of economics and especially *how* economic analysis is used […]." In his opinion, economists can contribute to competition law enforcement in three ways.

1. "The first area is case work […]. **Economic theory is necessary to "frame" a case** [*emphasis added*], which in turn is fundamental to arrive at a particular theory of harm. This typically involves information about the structure of the industry, the firms, the structure of demand and the technology, as well as a preliminary understanding of possible strategies. It will always be the first step in an economic analysis in the context of a competition case (including, in principle, a state aid case). […] The goal of a plausible theoretical framework in the context of a particular case is to come up with testable hypothesis concerning the theory harm. In this sense, competition policy decisions need to be based on empirical evidence. […] An effective economic analysis in the context of a case has to be based on empirical analysis, which in turn needs to be rooted in solid economic principles. The key challenge is to identify a particular theory (or behaviour) from other alternatives. Identification thus involves the uncovering of empirical evidence that is only consistent with the claimed theory, and is inconsistent with other theories" (Röller 2005: ch.3.1).
2. "The second area where economic reasoning is important is guidelines, and similarly block exemptions. [… G]uidelines give general rules that describe the frameworks that will be used under various circumstances. […] The challenge for **economists** in **developing guidelines** [*emphasis added*] is to be able to provide relatively simple rules that are yet economically sound in a large set of circumstances" (Röller 2005: ch. 3.2).
3. "The third area is […] *ex post* and *ex ante* analysis. ***Ex post* analysis** [*emphasis added*] is undertaken in order to understand how antitrust, state aid, and merger decisions have effected markets. A prominent example is the ex post studies that attempt to categorize antitrust and merger decisions in terms of a type I and II error framework. […] The second area where economics is important is ***ex ante* analysis** [*emphasis added*], such as in market monitoring. Market monitoring is the analysis of whether or not markets function well, in principle prior to possible antitrust action. *Ex ante* analysis, such as market monitoring, is important, since whenever an agency relies exclusively on complains, firms' incentives may be negatively effected. The challenge for market monitoring is to identify instances when markets do not function, such as anticompetitive conduct or the existence of entry barriers" (Röller 2005: ch. 3.3).

Consumer Welfare Standard vs. Total Welfare Standard

EC competition policy is concerned about maximizing consumer welfare, while economists often

prefer a standard based on total welfare (Monti 2007: p. 83-86). To see the difference, consider the monopoly in Figure 7. According to a **consumer welfare standard** one would say that consumer surplus (in comparison to a perfectly competitive market-outcome) has decreased by the deadweight loss (area C) *and* a re-distribution from consumers to producers (area B). According to the **total welfare standard** one would merely be concerned about the deadweight loss. By neglecting these re-distributive effects competition policy following a total welfare standard might tolerate practices which the consumer welfare standard would not allow.

Recall that there is an inverted u-shaped relationship between market power and innovations (see section C.2). Therefore, a re-distribution from consumers to producers might not necessarily be intolerable despite its current effect to reduce consumer surplus. In the long run, a firm may invest these higher profits in R&D-efforts resulting in better products or more efficient production. Therefore, a total welfare standard seems more appropriate when **dynamic and productive efficiency** are considered besides allocative efficiency.

D.3 Competition Law in the European Union

This subsection provides an overview on competition policy in the European Union. It starts with presenting the European competition authority at the Commission and presents some basic information on legislation in the European Union. Then, it turns to the relevant competition laws.

The European Competition Authority

In Europe, competition policy is in the area of responsibility of the **European Commission**. The Commission consists of 27 members (one per Member State). The members of the Commission are appointed by the Council, after approval by the European Parliament, for a period of five years. The Commission's departments are divided into **directorates-general**, each of which is responsible for a particular area of activity.[28]

Since February 2010, Joaquín Almunia is the current **competition commissioner**.[29] Between 2004 and 2010 he acted as commissioner for economic and monetary affairs. In his mandate[30], he sees "competition policy as a means of strengthening our social market economy, and enhancing its efficiency and fairness. Together with the euro and the internal market, competition policy gives Europe the means to create more prosperity for its citizens." In a speech, that Mr. Almunia gave on 11 February 2011 he states[31]:

28 http://eur-lex.europa.eu/en/droit_communautaire/droit_communautaire.htm#3.3

29 http://ec.europa.eu/commission_2010-2014/almunia/index_en.htm

30 http://ec.europa.eu/commission_2010-2014/almunia/about/mandate/index_en.htm

31 http://europa.eu/rapid/pressReleasesAction.do?

"I have seen first-hand that the fair and robust enforcement of EU competition law helps business and consumers make the most of the internal market, which is a key asset for Europe. The importance of enforcement is clearest in our fight against cartels, which I regard as the most serious offence in competition law. [...] Cartels are bad for everybody, except for those who participate in them! They increase the price for companies, often SME's, as they often concern intermediate products, and directly or indirectly hurt Europe's consumers.[...] The reason why I always will be tough on cartels is very simple, and has to do with the priorities of the Commission policies. Our priority number one is to help increase our competiveness in the world, our growth potential and the ability of our economies to create jobs – all urgently needed in view of the lasting damages created by the economic and financial crisis. Competition policy will make a substantial contribution to this, by encouraging companies to compete on the merits and innovate, which they are more likely to do if they operate in a sound and undistorted environment. [...]

Apart from cartels, we have also worked hard to protect the single market from the harm done by restrictive agreements and abuse of dominant positions. But above all, over the past year I have seen that our merger-control system is objective, fast, and proportionate. It prefers prevention to the cure, as shown in the recent case involving Intel and McAfee. [...] Just over two weeks ago, I authorised the proposed acquisition of McAfee by Intel accepting Intel's commitment that its hardware will remain open to the security solutions McAfee's competitors will find in the future. This decision shows our ability to intervene before problems actually occur. Complex antitrust investigations and Court proceedings take time and can come too late to restore competition. It also shows our ability to close even complex cases within the tight Phase I deadline, thanks to the cooperation of the parties. [...]

I can see three main areas of development for competition policy. First, as I said before, I intend to bring my full contribution to extend and deepen the internal market. We will follow the conclusions of the latest European Council on energy and innovation and redouble our efforts in the network industries. Also, if we are to ensure that our internal market is truly open to European companies, we need to step up our control of entrenched incumbents and other dominant companies, including in the new Member States. Finally, I intend to increase the protection of consumers and SMEs, which are often the first victims of the restrictions and the higher prices brought about by anticompetitive practices, such as cartels. SMEs deserve special

reference=SPEECH/11/96&format=HTML&aged=0&language=EN&guiLanguage=en

attention, because of their importance for regional development and their role in the social fabric of many parts of Europe. As to State Aid, we are in the process of updating our rescue and restructuring guidelines. In this important process we are identifying the lessons learned in our effort to support the financial and non-financial sectors during the crisis. [...]

Competition is an instrument, not an end in itself. But it is indeed a vital instrument in very many respects. Without fair, robust, and effective competition policy and enforcement, I don't see how we Europeans can overcome the crisis rapidly and shape up to compete with the other, dynamic players that are increasingly present on the world scene. Of course, Competition is not the only tool we should use to pursue this goal. But we need a vibrant and competitive environment in the single market if we are serious about leading in the information age. We need competition to be equal partners with the US, China, and the other leading global players; we need competition to grow; we need competition to preserve our social model for the benefit of our citizens and of the future generations. Considering our demographic trends and the imperative task of building sustainable and green economic and social models, Europe needs all its resources and resourcefulness. The EU competition system is one of the best, if not the best in the world. My commitment is to use it to the full extent of the law, because I am convinced that this is what I must do within my area of responsibility to contribute to a better future for Europe."

The competition rules are enforced by the **Directorate General for Competition** (DG Comp).[32] Currently, DG Comp is headed by the PhD-economist Alexander Italianer as Director General. Economic advice is provided by a team of economists headed by Prof. Kai-Uwe Kühn who succeeds Prof. Damien Neven. The Chief Economist[33] provides independent guidance on methodological issues of economics and econometrics in the application of EU competition rules. He contributes to individual competition cases (in particular ones involving complex economic issues and quantitative analysis), to the development of general policy instruments, as well as assisting with cases pending before the Community Courts.

Legislation in the European Union

This section describes the main types of European laws.[34] The treaties constitute the European

32 http://ec.europa.eu/dgs/competition/index_en.htm

33 http://ec.europa.eu/dgs/competition/economist/role_en.html

34 This section draws on information provided by the European Union.

Union's **primary legislation**, which is comparable to constitutional law at national level. They lay down the fundamental features of the Union. In competition policy we are mainly concerned with the **Treaty on the functioning of the European Union** (TFEU) which according to its Article 1 "organises the functioning of the Union and determines the areas of, delimination of, and arrangements for exercising its competences."

The **secondary legislation** comprises **binding legal instruments** (regulations, directives and decisions) and **non-binding instruments** (resolutions, opinions). A **regulation** is a general measure that is addressed to everyone and binding in all its parts. A regulation is directly applicable, which means that it creates law which takes immediate effect in all the Member States in the same way as a national instrument, without any further action on the part of the national authorities. A **directive** is addressed to the Member States. Its main purpose is to align national legislation. A directive is binding on the Member States as to the result to be achieved but leaves them the choice of the form and method they adopt to realize the Community objectives within the framework of their internal legal order. A **decision** is the instrument by which the Community institutions (e.g., the Commission) give a ruling on a particular matter. By means of a decision, the institutions can require a Member State or a citizen of the Union to take or refrain from taking a particular action, or confer rights or impose obligations on a Member State or a citizen. All the decisions handed down by bodies exercising judicial powers constitute case-law.

Overview on European Competition Laws and their Objectives

In Europe[35], competition law was established in 1957 when the Treaty establishing the European Economic Community (EEC Treaty, or Treaty of Rome) was agreed. This treaty was amended and renamed Treaty establishing the European Community (EC Treaty) in 1993, and again amended and renamed Treaty on the functioning of the European Union (**TFEU**) in 2009. Along with these amendments came a new numbering of these treaties' articles. Therefore, what is now Article 101 TFEU (**agreements between undertakings**) was previously called Article 85 EEC Treaty and later Article 81 EC Treaty.

Article 101 TFEU

1. The following shall be prohibited as incompatible with the internal market: all

http://europa.eu/institutions/decision-making/index_en.htm

http://eur-lex.europa.eu/en/droit_communautaire/droit_communautaire.htm#1

35 A summary of the relevant legislation is provided at:
http://europa.eu/legislation_summaries/competition/index_en.htm
A glossary of the terms used in EU competition policy can be found at:
http://ec.europa.eu/competition/publications/glossary_en.pdf

agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition within the internal market, and in particular those which:

(a) directly or indirectly fix purchase or selling prices or any other trading conditions;

(b) limit or control production, markets, technical development, or investment;

(c) share markets or sources of supply;

(d) apply dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage;

(e) make the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts.

2. Any agreements or decisions prohibited pursuant to this Article shall be automatically void.

3. The provisions of paragraph 1 may, however, be declared inapplicable in the case of:

- any agreement or category of agreements between undertakings,

- any decision or category of decisions by associations of undertakings

- any concerted practice or category of concerted practices,

which contributes to improving the production or distribution of goods or to promoting technical or economic progress, while allowing consumers a fair share of the resulting benefit, and which does not:

(a) impose on the undertakings concerned restrictions which are not indispensable to the attainment of these objectives;

(b) afford such undertakings the possibility of eliminating competition in respect of a substantial part of the products in question.

Article 101 TFEU prohibits both vertical and horizontal agreements among firms. The exemptions according to Article 101 III TFEU are specified in block exemption regulations for, e.g., research and development agreements or specialization agreements, vertical agreements, or concerted practices in the motor vehicle sector. The economic assessment of Article 101 cases is presented in

section H .

Article 102 TFEU (**abuse of a dominant position**) used to be Article 86 EEC Treaty and later Article 82 EC Treaty.

> *Article 102 TFEU*
>
> *Any abuse by one or more undertakings of a dominant position within the internal market or in a substantial part of it shall be prohibited as incompatible with the internal market in so far as it may affect trade between Member States.*
>
> *Such abuse may, in particular, consist in:*
>
> *(a) directly or indirectly imposing unfair purchase or selling prices or other unfair trading conditions;*
>
> *(b) limiting production, markets or technical development to the prejudice of consumers;*
>
> *(c) applying dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage;*
>
> *(d) making the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts.*

Article 102 TFEU prohibits the abuse of a dominant position by **unilateral conduct**, i.e. by a single firm. Therefore, Article 102 TFEU must only be applied after proving a firm's dominant position in the relevant market. According to the Commission's guidelines on the application of Art. 82 EC Treaty, **dominance** "has been defined under Community law as a position of economic strength enjoyed by an undertaking, which enables it to prevent effective competition being maintained on a relevant market, by affording it the power to behave to an appreciable extent independently of its competitors, its customers and ultimately of consumers. [...] The Commission considers that an undertaking which is capable of profitably increasing prices above the competitive level for a significant period of time does not face sufficiently effective competitive constraints and can thus generally be regarded as dominant" (EC 2009). Aspects that need to be taken into account when assessing dominance are, for example, market shares, expansion or entry by other firms, and/or countervailing buyer power. The economic importance of these aspects is shown in section E . The economics of abuse of dominance cases is presented in section I . Following Article 103 TFEU, the broad principles of Articles 101 and 102 TFEU shall be specified by regulations and directives (see above).

Articles 101 and 102 TFEU apply to firms and the relationship among firms. Thus, they shall ensure firms' freedom to compete. Moreover, Article 3 TFEU specifies that the European "Union shall have exclusive competence in [...] the establishing of the competition rules necessary for the functioning of the internal market". In particular, economic policy in the European Union shall be "conducted in accordance with the principle of an open market economy with free competition" (Article 119 TFEU). Therefore, European competition policy is characterized by an ideal of perfect competition (see subsection D.1 and section A.2), where the consumer, not the state, dictates what goods and services are provided. However, markets do not always lead to (almost) perfectly competitive outcomes as has been shown in Part I of this reader. This requires interventions into the market by means of competition policy. Thereby, European "competition law is not designed as a highly interventionist policy to guarantee the welfare of every segment of the economy, nor is it designed to compel or create incentives for firms to behave to promote economic welfare" (Monti 2007: p. 2). The **aim of European competition policy** is more modest: to establish "the competition rules necessary for the functioning of the internal market" (Article 3 TFEU).

"**Economic welfare** [*emphasis added*] is one of the anticipated benefits of membership of the EC, and the Commission noted the contribution of competition policy to economic efficiency early on. In the *First Report on Competition Policy* we find this passage:

> Competition is the best stimulant of economic activity since it guarantees the widest possible freedom of action to all. An active competition policy pursued in accordance with the provisions of the Treaties establishing the Communities makes it easier for the supply and demand structures continually to adjust to technological development. Through the interplay of decentralised decision-making machinery, competition enables enterprises continuously to improve their efficiency, which is the sine qua non for the steady improvement in living standards and employment prospects of the Community" (Monti 2007: p. 44).

Here, the term **efficiency** refers to allocative, productive, and dynamic efficiency as introduced in Part I of this reader (see sections A.1 and C). The last sentence shows that "efficiency is, on the one hand, a result of economic freedom and, on the other, not an end in itself" (Monti 2007: p. 44). For example, economic efficiency is recognized in Article 101 III TFEU as is explained in greater detail in section H.3 . In summary, the shared consensus among economists is that contemporary EC competition policy is concerned about maximizing consumer welfare, while economists prefer a standard based on total welfare (see subsection D.2).

European competition law includes one objective that is not present in, e.g., US-American or German competition law, i.e. the goal of establishing a **single market**. Article 26 TFEU specifies

that the "Union shall adopt measures with the aim of establishing or ensuring the functioning of the internal market [... that comprises] an area without internal frontiers in which the free movement of goods, persons, services and capital is ensured [...]." "Some eminent scholars and the Commission have gone so far as to say that market integration is the 'first principle' of EC competition policy" (Monti 2007: p. 39). In the light of this objective, Article 101 TFEU prohibits anticompetitive agreements between firms which may **affect trade between Member States** while Article 102 prohibits any "abuse [...] of a dominant position [...] in so far as it may affect trade between Member States." The main concern is that firms may create market divisions, e.g., agreements on the allocation of exclusive territories, the prevention of parallel trade between Member States, or price discrimination across Member States.

First, the concentration on conduct that may affect trade between Member States is capable of shaping even the meaning of terms such as 'agreement' in Art. 101 TFEU in order "to cover a wide range of tactics designed to thwart parallel imports. From a commercial perspective, because there remain significant price difference for many goods among Member States, a manufacturer has an incentive to price discriminate to reap higher margins in Member States where prices are high. This strategy would collapse if parallel importers were able to obtain goods sold cheaply in one Member State and resell them to another Member State where the price is higher" (Monti 2007: p. 41). Therefore, agreements with distributors that prevent parallel imports are contrary to Art. 101 I TFEU. Second, the concentration on conduct that may affect trade between Member States facilitates to distinguish more clearly in which cases European competition law is applicable, and when national laws apply. The general rule according to Council Regulation No. 1/2003 Article 3 is that national laws must be consistent with and at least as restrictive as European laws.

Two further principles of EC competition law are the focus on market power and the objective of pluralism (Monti 2007: 86). At the root of the Commission's economic approach is the recognition that competition problems arise more frequently when firms hold **market power**. Hence, stricter standards apply when the firm has a dominant position, and considerably looser standards are applied when the firm is small in a market with many small competitors. This approach has affinities with the SCP paradigm, being premised upon market structure as a measure of presumptive legality/illegality. Finally, a value that underlies EC competition policy is a belief that a **plurality** of market participants guarantees better economic performance. This draws upon the SCP paradigm but is also premised upon the ordoliberal values of competition policy. However, pluralism is no longer favored merely for its own sake (as per the ordoliberal theory) but as a means to increased consumer welfare.

Articles 101 and 102 TFEU apply to firms. In contrast, Article 107 TFEU (**state aid**) refers

to states who may not provide aid to firms which might distort competition and affect trade between Member States.

Article 107 TFEU

1. Save as otherwise provided in the Treaties, any aid granted by a Member State or through State resources in any form whatsoever which distorts or threatens to distort competition by favouring certain undertakings or the production of certain goods shall, in so far as it affects trade between Member States, be incompatible with the internal market.

2. The following shall be compatible with the internal market:

(a) aid having a social character, granted to individual consumers, provided that such aid is granted without discrimination related to the origin of the products concerned;

(b) aid to make good the damage caused by natural disasters or exceptional occurrences;

(c) aid granted to the economy of certain areas of the Federal Republic of Germany affected by the division of Germany, in so far as such aid is required in order to compensate for the economic disadvantages caused by that division. Five years after the entry into force of the Treaty of Lisbon, the Council, acting on a proposal from the Commission, may adopt a decision repealing this point.

3. The following may be considered to be compatible with the internal market:

(a) aid to promote the economic development of areas where the standard of living is abnormally low or where there is serious underemployment, and of the regions referred to in Article 349, in view of their structural, economic and social situation;

(b) aid to promote the execution of an important project of common European interest or to remedy a serious disturbance in the economy of a Member State;

(c) aid to facilitate the development of certain economic activities or of certain economic areas, where such aid does not adversely affect trading conditions to an extent contrary to the common interest;

(d) aid to promote culture and heritage conservation where such aid does not affect trading conditions and competition in the Union to an extent that is contrary to the common interest;

(e) such other categories of aid as may be specified by decision of the Council on a

proposal from the Commission.

D.4 Competition Law in Germany and USA

Germany

In Germany, competition laws (*Gesetz gegen Wettbewerbsbeschränkungen*; GWB) have been passed in 1957, i.e. much later than in USA. As a consequence, during the time of the *Weimarer Republik* the German economy was characterized by up to 4000 cartels. On 15 July 1933 a law was passed (*Zwangskartellgesetz*) that allowed the NS-regime to gain power over important cartels and control business. Therefore, when the GWB was passed in 1957[36] it was also inspired by the idea to prevent politics to gain power over business which had contributed to the cruelties committed during 1933 and 1945. Since 1957, the GWB has been changed by seven amendments. In 2005, the seventh amendment contributed much to the alignment of the GWB to European laws. The eigth amendment is currently under preparation.[37]

By focussing on firms' freedom (to compete) the GWB's **"definition" of competition** was inspired by the ordo-liberal school (*Freiburger Schule*) who emphasized the ideal of perfect competition. However, the GWB neither was nor is too narrowly focussed on this ideal. It is rather pragmatical by taking into account that perfect competition can hardly be achieved in reality. In this context, one might think of markets for differentiated goods or production technologies characterized by economies of scale. Summarizing, one may say that the GWB does not provide a positive definition of competition, because such a definition can hardly be found as was argued in subsection D.2 . The GWB rather provides a **negative definition** of competition by specifying behavior that is inconsistent with competition such as boycotts and refusals to supply (§22 GWB).

The two central norms of German competition laws[38] are laid down in §1 on **agreements between undertakings** and §19 on the **abuse of a dominant position**.

> *§1 GWB – Verbot wettbewerbsbeschränkender Vereinbarungen*
>
> *Vereinbarungen zwischen Unternehmen, Beschlüsse von Unternehmensvereinigungen und aufeinander abgestimmte Verhaltensweisen, die eine Verhinderung, Einschränkung oder Verfälschung des Wettbewerbs bezwecken oder bewirken, sind verboten.*

36 Also see Der Spiegel (1957). "Der siebenjährige Krieg." Vol. 27/1957, pp. 17-25. http://www.spiegel.de/spiegel/print/d-41757813.html

37 This short history of and introduction into the German GWB draws on Kling and Thomas (2007: p. 473-500)

38 http://www.bundeskartellamt.de/wDeutsch/Rechtsgrundlagen/Rechtsgrundlagen.php

§ 19 Missbrauch einer marktbeherrschenden Stellung

(1) Die missbräuchliche Ausnutzung einer marktbeherrschenden Stellung durch ein oder mehrere Unternehmen ist verboten.

(2) Ein Unternehmen ist marktbeherrschend, soweit es als Anbieter oder Nachfrager einer bestimmten Art von Waren oder gewerblichen Leistungen auf dem sachlich und räumlich relevanten Markt

1. ohne Wettbewerber ist oder keinem wesentlichen Wettbewerb ausgesetzt ist oder

2. eine im Verhältnis zu seinen Wettbewerbern überragende Marktstellung hat; hierbei sind insbesondere sein Marktanteil, seine Finanzkraft, sein Zugang zu den Beschaffungs- oder Absatzmärkten, Verflechtungen mit anderen Unternehmen, rechtliche oder tatsächliche Schranken für den Marktzutritt anderer Unternehmen, der tatsächliche oder potentielle Wettbewerb durch innerhalb oder außerhalb des Geltungsbereichs dieses Gesetzes ansässige Unternehmen, die Fähigkeit, sein Angebot oder seine Nachfrage auf andere Waren oder gewerbliche Leistungen umzustellen, sowie die Möglichkeit der Marktgegenseite, auf andere Unternehmen auszuweichen, zu berücksichtigen.

Zwei oder mehr Unternehmen sind marktbeherrschend, soweit zwischen ihnen für eine bestimmte Art von Waren oder gewerblichen Leistungen ein wesentlicher Wettbewerb nicht besteht und soweit sie in ihrer Gesamtheit die Voraussetzungen des Satzes 1 erfüllen. Der räumlich relevante Markt im Sinne dieses Gesetzes kann weiter sein als der Geltungsbereich dieses Gesetzes.

(3) Es wird vermutet, dass ein Unternehmen marktbeherrschend ist, wenn es einen Marktanteil von mindestens einem Drittel hat. Eine Gesamtheit von Unternehmen gilt als marktbeherrschend, wenn sie

1. aus drei oder weniger Unternehmen besteht, die zusammen einen Marktanteil von 50 vom Hundert erreichen, oder

2. aus fünf oder weniger Unternehmen besteht, die zusammen einen Marktanteil von zwei Dritteln erreichen,

es sei denn, die Unternehmen weisen nach, dass die Wettbewerbsbedingungen zwischen ihnen wesentlichen Wettbewerb erwarten lassen oder die Gesamtheit der Unternehmen im Verhältnis zu den übrigen Wettbewerbern keine überragende Marktstellung hat.

> *(4) Ein Missbrauch liegt insbesondere vor, wenn ein marktbeherrschendes Unternehmen als Anbieter oder Nachfrager einer bestimmten Art von Waren oder gewerblichen Leistungen*
>
> *1. die Wettbewerbsmöglichkeiten anderer Unternehmen in einer für den Wettbewerb auf dem Markt erheblichen Weise ohne sachlich gerechtfertigten Grund beeinträchtigt;*
>
> *2. Entgelte oder sonstige Geschäftsbedingungen fordert, die von denjenigen abweichen, die sich bei wirksamem Wettbewerb mit hoher Wahrscheinlichkeit ergeben würden; hierbei sind insbesondere die Verhaltensweisen von Unternehmen auf vergleichbaren Märkten mit wirksamem Wettbewerb zu berücksichtigen;*
>
> *3. ungünstigere Entgelte oder sonstige Geschäftsbedingungen fordert, als sie das marktbeherrschende Unternehmen selbst auf vergleichbaren Märkten von gleichartigen Abnehmern fordert, es sei denn, dass der Unterschied sachlich gerechtfertigt ist;*
>
> *4. sich weigert, einem anderen Unternehmen gegen angemessenes Entgelt Zugang zu den eigenen Netzen oder anderen Infrastruktureinrichtungen zu gewähren, wenn es dem anderen Unternehmen aus rechtlichen oder tatsächlichen Gründen ohne die Mitbenutzung nicht möglich ist, auf dem vor- oder nachgelagerten Markt als Wettbewerber des marktbeherrschenden Unternehmens tätig zu werden; dies gilt nicht, wenn das marktbeherrschende Unternehmen nachweist, dass die Mitbenutzung aus betriebsbedingten oder sonstigen Gründen nicht möglich oder nicht zumutbar ist.*

As it is easily seen, §1 parallels Article 101 TFEU. §2 GWB (*freigestellte Vereinbarungen*) constitutes the German analog to the provisions of Article 101 III TFEU on efficiencies. The main difference between §19 GWB and Article 102 TFEU is §19 III GWB, which bases the presumtion of dominance on firms' market shares. In contrast, on a European level the existence of dominance must be shown on basis of an economic assessment. The relevant German rules on merger control may be found in §§35-43 GWB.

Violations of §§ 1 or 19 GWB can be enforced in three ways (Kling and Thomas 2007: ch. 23). First, the German competition authority (*Bundeskartellamt*, BKartA) may engage in **administrative proceedings** and interdict the wrongful behavior. To do this, §§ 58 and 59 GWB authorize the *Bundeskartellamt* to carry out searchings at the alleged infringers of competition laws and seize relevant files. The decision of the *Bundeskartellamt* may be appealed according to § 63 GWB at the Higher Regional Court (*Oberlandesgericht*) in Düsseldorf. Second, according to § 81

GWB the BKartA may impose **fines** on the infringers of the laws. Third, according to § 33 GWB the affected customers may – in **civil proceedings** at a district court (*Landgericht*) – claim damages for the harm incurred.

In addition to these rules, the German GWB emphasizes the protection of **small and medium-sized enterprises** (SMEs). For example, §3 GWB specifies that agreements among SMEs may be exempted from the prohibition of §1 GWB if they contribute to improving the competitiveness of those firms. Similarly, §20 GWB allows small firms to engage in impediments of other firms that would be prohibited in case of dominant firms. In this context, the definition of a small or medium-sized firm is not clearly defined but must be assessed *relative* to the size of the other firms in this market (Kling and Thomas 2007: p. 494). In practice, one could also define SMEs consistent with §35 GWB (merger control) and presume that a **firm is large** if its yearly revenues are above 500m EUR, and that it is **small or medium-sized** if its yearly revenues are below 10-15m EUR.

The GWB must be distinguished from the German law against unfair competition (*Gesetz gegen unlauteren Wettbewerb*, **UWG**). In short, the GWB concerns the relationship between firms and consumers and prohibits restraints of competition that might harm consumers. The UWG concerns the relationship between competitors and prohibits methods of unfair competition that harm (some of) the firms.

In Germany, competition laws are enforced by the **Federal Cartel Agency** (*Bundeskartellamt*, BKartA) in Bonn as well as competition authorities of the separate federal states.[39] At the Bundeskartellamt, which is headed by Andreas Mundt, work about 320 employees (about 50% with a background in economics and 50% with a legal background).[40] The decisions concerning cartels, the abuse of a dominant position as well as mergers are made by in total twelve decision-making departments (*Beschlussabteilungen*) where each department is responsible for a specific set of industries. Two of those departments are solely concerned with the prosecution of cartels. Assistance to the preparation, execution and assessment of searchings at suspected cartel-firms as well as assistance for leniency-applications is provided by a special commission (*Sonderkommission*) on cartels. Moreover, the *Beschlussabteilungen* receive general assistance by a **department for general economic and legal affairs** (*Grundsatzabteilung*) that includes a department for economic affairs, which is somewhat comparable to the European Chief Economist team. Moreover, the *Bundeskartellamt* is run by several administrative departments.[41] In the first

39 http://www.bundeskartellamt.de/wDeutsch/service/LKB.php

40 http://www.bundeskartellamt.de/wDeutsch/bundeskartellamt/Bundeskartellamt_4.php

41 http://www.bundeskartellamt.de/wDeutsch/bundeskartellamt/Organisation/Organisation.php
Further information is provided in this brochure:

instance, the decisions on cases are made within the BKartA's decision-making departments by the majority vote of the department's chairman and two assessors. In the second instance, decisions may be appealed at the **Higher Regional Court** in Düsseldorf.

According to §§ 44-47 GWB, further information is provided by a **commission on monopolies** (*Monopolkommission*)[42] that in every second year publishes statistics on the concentration in several German industries as well as a report (*Hauptgutachten*) on merger control and further matters of competition policy. The most recent report titled "Mehr Wettbewerb, wenig Ausnahmen" was published in 2010 and concerned the regulation of water-suppliers and pharmacies, the effects of labor-markets on product-markets, and an assessment of competitive issues in health insurance. Additionally, the *Monopolkommission* may publish special reports (*Sondergutachten*) such as the ones on demergers, postal services, telecommunications, railway services, and energy. The Monopolkommission consists of five members and is currently headed by Justus Haucap, professor of economics in Düsseldorf. Its staff consists of ten scientific employees, four administrative employees, and a general secretary.

USA

The US-American Sherman Act from 1890 may be considered the foundation for modern competition laws. At that time the necessity to protect competition by law had arisen because of the developments in many industries. The extension of the American railway-network had contributed to the creation of large national markets rather than local or regional ones. Moreover, trusts gained power causing, e.g., a monopolization of business or restraints to trade. Trusts are business entities where a settlor, e.g. the owner of stocks of a company, entrusts these shares to the trustee or board of trustees in exchange for dividend-paying certificates. This gave the trustees control over otherwise competing companies, whose behavior could now be coordinated to achieve a collusive outcome.

This led to the formulation of the **Sherman Act** (see sections 1-7 of Title 15 on Commerce and Trade in the United States Code (U.S.C.)). The Sherman Act provides the basic laws condemning the the two main restrictions of competition, i.e. **anticompetitive agreements between unterakings** (Sherman Act § 1) and unilateral conduct that monopolizes or attempts to monopolize (Sherman Act § 2). In the terminology of European competition laws the latter refers to the **abuse of a dominant position**. The goal to limit the effect of trusts on market-outcomes provides an explanation for US-American competition laws being called **antitrust laws**.

http://www.bundeskartellamt.de/wDeutsch/download/pdf/Infobroschuere/1009_Infobroschuere_deutsch.pdf

42 http://www.monopolkommission.de/

Sherman Act § 1, 15 U.S.C. § 1

> *Every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is declared to be illegal. Every person who shall make any contract or engage in any combination or conspiracy hereby declared to be illegal shall be deemed guilty of a felony, and, on conviction thereof, shall be punished by fine not exceeding $100,000,000 if a corporation, or, if any other person, $1,000,000, or by imprisonment not exceeding 10 years, or by both said punishments, in the discretion of the court.*

In other words, §1 Sherman Act can be read as: *every agreement in restraint of trade or commerce is declared to be illegal* or *every agreement whose anticompetitive effects on trade outweigh its procompetitive effects is illegal*. "But while this is the general standard, the U.S. Supreme Court has also held that certain agreements are so likely to be anticompetitive, and so unlikely to have procompetitive effects, that they are condemned **"per se,"** [*emphasis added*] which means without any case-by-case inquiry into their effect. The following horizontal agreements have been held to be per se illegal: price-fixing, market divisions, output restraints, and boycotts. [...] If a per se rule does not apply. Then general **"rule of reason"** [*emphasis added*] review applies. Under the rule of reason, courts consider on a case by case basis whether the agreement has a plausible procompetitive justification" (Elhauge 2008: p. 49-50).

Sherman Act § 2, 15 U.S.C. § 2

> *Every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons, to monopolize any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a felony, and, on conviction thereof, shall be punished by fine not exceeding $100,000,000 if a corporation, or, if any other person, $1,000,000, or by imprisonment not exceeding 10 years, or by both said punishments, in the discretion of the court.*

"The statute is generally targeted at unilateral conduct [...]. However, agreements or combinations to form a corporation [...] that exercises monopoly power have long been held to constitute monopolization in violation of § 2" (Elhauge 2008: p. 52). Therefore, in principle it is possible to break up dominant firms according to US-law. A general prohibition of unfair methods of competition is provided in Federal Trade Commission Act § 5:

Federal Trade Commission Act § 5, 15 U.S.C. § 45

Unfair methods of competition in or affecting commerce, and unfair or deceptive acts or practices in or affecting commerce, are hereby declared unlawful.

Lessons Learned

After reading this section you should be able to answer the following questions.

1. Explain why measures of competition policy may increase welfare in an economy.
2. What are the main differences between the *Harvard School* and the *Chicago School*? How is competition defined today?
3. What are the main laws in European competition policy? Which authority enforces competition laws in Europe?
4. What are the main laws in US competition policy? Which authority enforces antitrust laws in USA?
5. What are the main laws in German competition policy? Which authority enforces competition laws in Germany?

References

Elhauge, E. (2008). "United States Antitrust Law and Economics." Thomson Reuters/West Foundation Press: New York

European Commission (2011). "Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements." Official Journal of the European Union. C11/1. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2011:011:0001:0072:EN:PDF

European Commission (2009). "Communication from the Commission – Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings." Official Journal of the European Union. C45/7. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:045:0007:0020:EN:PDF

Kling, M. and Thomas, S. (2007). "Kartellrecht." Verlag Franz Vahlen: München

Monti, G. (2007). "EC Competition Law." Cambridge University Press: Cambridge

Röller, L.-H. (2005). "Economic Analysis and Competition Policy Enforcement in Europe." http://ec.europa.eu/competition/speeches/text/sp2005_011_en.pdf

Schmidt, I. (2005). "Wettbewerbspolitik und Kartellrecht." 8th edition. UTB Lucius & Lucius: Stuttgart

Schmidt, I. (2012). “Wettbewerbspolitik und Kartellrecht.” 9th edition. Oldenbourg Verlag: München

Vickers, J. (1995). “Concepts of Competition.” Oxford Economic Papers. Vol. 47 No. 1, pp. 1-23

E MARKET POWER

E.1 Market Power, Structure, and Performance

Market power can be defined in two ways (Motta 2004: p. 115), i.e. as the ability of a firm to set its price p

(i) substantially above marginal costs c, or

(ii) above the competitive price level.

In perfect competition both definitions are equivalent, because profit-maximizing firms set prices at the level of marginal costs (see equation (14)). Hence, this narrow definition of market power focuses on its effects on prices and, thus, allocative efficiency.

In the European legislation (EU 2004.a: para. 25) market power is defined more broadly as "the ability to maintain prices above competitive levels for a significant period of time or to maintain output in terms of product quantities, product quality and variety or innovation below competitive levels for a significant period of time." This broad definition of market power does not only focus on allocative efficiency but also regards the effects on dynamic efficiency (see section C.2) and product variety.

In the following, we concentrate on the narrow definition of market power, i.e. its effects on prices. We show that the market power of firms, e.g., depends on the structure of the market. To see this, we provide information on **measures of market power** such as the Lerner-index and market shares, and **measures of market structure** such as concentration ratios and the Herfindahl-Hirschmann-index.

These measures fit well in the structure-conduct-performance (SCP) literature. This strand in the literature proposes that the structure of a market (see for example the assumptions of the Cournot-oligopoly in section B.3) determines firms' behavior, which results in a certain market outcome (Perloff et al. 2007: ch. 2). Moreover, the below measures are static in the sense that they measure market structure and market power at a single point in time and neglect intertemporal (i.e. dynamic) aspects. For example, the behavior of firms may have a feedback-effect on the structure of the market. This reverses the causal relationship that is proposed by the SCP-paradigm and, thus, extends this literature.

Measures of Market Power and Market Structure in a Static Market

To see the relationship between market power and market structure, consider the case of the monopoly from section A.3 The first-order condition (16) in conjunction with the assumption of downward-sloping demand ($dp(q)/dq<0$) implies that the monopolist prices above marginal costs, i.e. according to the above definition it possesses market power.

$$p(q)=c(q)-\frac{dp(q)}{dq}\cdot q \tag{47}$$

As was shown above, this equation can be transformed into equation (17).

$$\frac{p(q)-c(q)}{p(q)} = \frac{1}{\eta} \tag{17}$$

Consequently, the degree of market power is inversely related to the demand elasticity faced by the monopolist (Cabral 2000: p. 72). The left-hand side of equation (17) is called **Lerner-index**.

The Lerner-index increases with the mark-up charged by the firm, which should be the most desirable feature of any index of market power (Motta 2004: p. 116). However, in reality the Lerner-index can hardly be determined because marginal costs are rarely known and can be estimated only with great difficulty. Measuring the price-elasticity of demand is even more complicated. Moreover, the Lerner-index may both over- and underestimate market power. The Lerner-index underestimates market power if the existence of market power prevents productive efficiency. In this case, the observed marginal costs may well be above their competitive levels. The Lerner-index may underestimate market power if the production of a good requires significant sunk costs or fixed costs but rather small marginal costs as is, e.g., the case with software or medicine. Therefore, firms set prices above marginal costs in order to cover these fixed or sunk costs. By setting prices at the level of average costs, these firms might behave perfectly in line with the above condition (14) for the perfectly competitive long-run equilibrium. However, the Lerner-index would wrongly indicate the existence of market power.

The Lerner-index can also be calculated for Cournot-oligopolies (Motta 2004: p. 123). To see this, recall that the first-order condition of profit-maximization of some firm j (see equation (25)) is given by

$$\frac{d\pi_j}{dq_j} = p\left(\sum_i q_i\right)+\frac{dp\left(\sum_i q_i\right)}{dq_j}\cdot q_j - c_j \stackrel{!}{=} 0 \quad .$$

After defining $\Sigma_i q_i = q$, this first-order condition can be rewritten as

$$\frac{p(q)-c_j}{p(q)} = -\frac{dp(q)}{dq}\cdot\frac{q_j}{p(q)}\cdot\frac{dq}{dq_j} \quad . \tag{48}$$

The **market-share** m_j of firm j can be written as

$$m_j = \frac{q_j}{q} \quad . \tag{49}$$

Moreover, note that $dq/dq_j=1$ so that equation (48) for the Lerner-index L_j can be written as follows.

$$\begin{aligned} \frac{p(q)-c_j}{p(q)} &= -\frac{dp(q)}{dq}\cdot\frac{q}{p(q)}\cdot\frac{q_j}{q} \\ L_j &= \frac{m_j}{\eta} \end{aligned} \tag{50}$$

Hence, the price-cost margin for each Cournot-oligopolist is determined by its own market share m_j and the overall market demand elasticity η. Hence, equation (17) is a special case of (50) for $m_j=100\%$. Based on the Lerner-index for a particular firm j, the Lerner-index for an entire industry can be written as follows.

$$\begin{aligned} L &= \sum_i m_i L_i \\ &= \frac{p(q)-\sum_i m_i c_i}{p(q)} \end{aligned} \tag{51}$$

Equation (50) shows that the market power of firm j increases in its share of the market m_j. Therefore, competition authorities are often concerned with measuring firms' market shares. One would suspect that a monopolist (with $m_j=100\%$) can exercise most market power while a firm in perfect competition (with $m_j=100\%/n\approx 0\%$) can exercise least market power. Some estimates of Lerner-indices in U.S.-industries are provided in Figure 16.

Industry	*Lerner Index*
Food and kindred products	0.811
Tobacco	0.638
Textile mill products	0.214
Apparel	0.444
Lumber and wood	0.494
Furniture and fixtures	0.731
Paper and allied products	0.930
Printing	0.950
Rubber and plastic	0.337
Leather products	0.524
Stone, clay, and glass	0.606
Primary metals	0.540
Fabricated metals	0.394
Machinery	0.300
Electric equipment	0.676
Instruments	0.284
Miscellaneous mfg	0.777
Communication	0.972
Electric, gas, and sanitary services	0.921
Motor vehicles	0.433
Average	0.57

Figure 16: Estimated Lerner Index for Selected Industries

Source: Pepall et al. (2008: p. 53)

In subsection E.2 , we show that a high market share is not always indicative of market power. Measuring market shares requires a definition of the relevant market, i.e. one must determine the products and the regions that impose competitive constraints on each other. To see this, consider that a firm cannot significantly raise the price of its product in some region if its customers can easily buy the product in another region, if they can substitute the firm's product with that of a rival (demand-side substitution), or if a firm who does not produce the product at current prices would start producing it at higher prices (supply-side substitution). Further remarks on the delineation of markets are provided in section F .

In equation (49) firm *j*'s **market share** is defined as its share in total output. Alternatively, its market share m_j' can be defined as its share in industry-revenue.

$$m_j' = \frac{p_j q_j}{\sum_i p_i q_i} \tag{52}$$

When assessing market shares it might be necessary to consider not only current market shares but also examine the past and (expected) future evolution of market shares. Maybe, a firm with a high market share today is likely to be a rather unimportant player in the future because of, for example, substitute products that will be on the rise in the future. Similarly, fluctuations in demand might drive the current market share above/below its longtime-value. In these cases, one might want to

calculate average market shares over a longer time-span. Furthermore, it is not only the existence of a certain pattern of market shares, but its persistence over time that might give a strong indication of an industry situation.

The market shares as defined in equations (49) and (52) apply to a single firm *j* only but do not provide information on the overall structure of the market. This overall market structure is important because, for example, a firm with a 25% market share can often exercise more market power if it faces 15 rivals each with a 5% market share (industry A) than by facing three rivals with a 25% market share each (industry B). These ideas are reflected by the concept of **concentration ratios** and **concentration curves** (Pepall et al. 2008: p. 44). A concentration ratio CR_x measures the cumulative market share of the *x* largest firms.

$$CR_x = \sum_{i=1}^{x} m_i \tag{53}$$

Hence, the CR_3 in industry A is $CR_{3,A}$=35% and $CR_{3,B}$=75% in industry B. This indicates that industry B is more concentrated than industry A. Concentration curves map these cumulative market shares as shown in Figure 17.

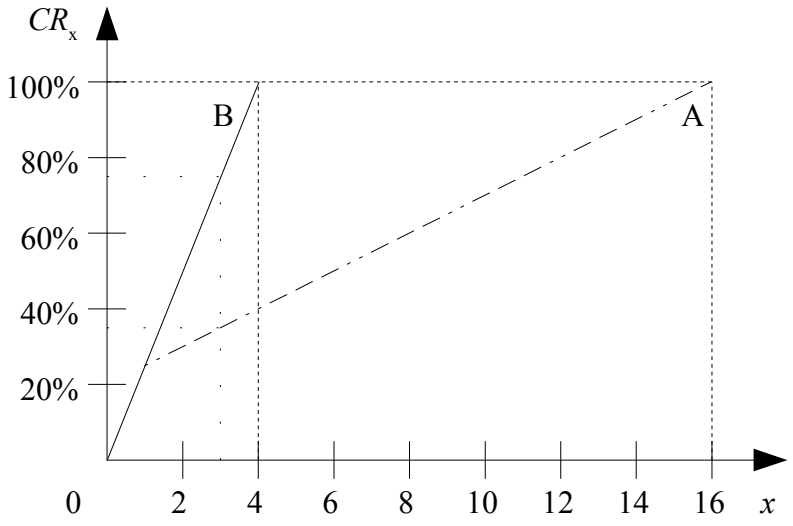


Figure 17: Concentration Curves

The German statistical office (Destatis 2010: p. 378) provides some measures of CR_6 for, e.g., the following industries in 2008. Market shares were calculated according to equation (52) as a firm's share in industry-revenue.

- Tobacco industry 95.5%
- Chemical products 28.8%
- Pharmaceutical products 45.0%
- Machine construction 13.3%
- Car manufacturing 66.4%

Concentration measures can be subject to a variety of problems. First, concentration statistics that only concentrate on domestic firms overstate concentration because they ignore imports. Second, concentration ratios are static measures and neglect dynamic aspects such as entry into the industry. In the short run, an extraordinarily profitable industry may be characterized by few firms and, thus, a high concentration. However, this may be expected to cause entry by additional firms, which intensifies competition (see subsection E.2). Third, concentration measures are biased if the relevant economic market is defined improperly (see section F). Fourth, the effects of seller concentration can be offset by powerful buyers (see subsection E.2).

Given equations (50) and (51), the Lerner-index L for an industry as a whole can be written as

$$\begin{aligned} L &= \sum_i m_i L_i \\ &= \frac{\sum_i m_i^2}{\eta} \\ &= \frac{HHI}{\eta} \end{aligned} \quad , \tag{54}$$

where

$$HHI = \sum_i m_i^2 \tag{55}$$

is the **Herfindahl-Hirschman Index** of concentration. Hence, there is a direct relationship between the degree of industrial concentration and the average degree of market power (in the Cournot-model). However, a particular market structure does not imply a particular market outcome, as market power depends on further factors such as the ones described in subsection E.2 .

The *HHI* is defined on a scale between 0 and 10,000 when market shares are measured in percentage points, or 0 to 1.0 when market shares are measured in decimals. In perfect competition the *HHI* takes a value of (close to) zero. In monopoly the *HHI* takes a value of 10,000. For the two

above industries we calculate HHI_A=1,000 and HHI_B=2,500. The *HHI* is advantageous in comparison to the concentration ratio, because by squaring market shares it puts larger weight on dissimilar distributions of market shares.

Assessment of Market Power (in Competition Policy)

The above measures are used in competition policy for analyzing market outcomes. This is because supra-competitive prices reduce the allocative efficiency of a market and, thus, cause a welfare-loss (see section A.1). Therefore, it is important for a competition authority to have an idea whether, for example, the merger of two firms (one with a 25%-market share, the other with a 5%-market share) in the above industry A is likely to raise prices in this industry much (see section G). The same is true if firms coordinate to affect the market outcome (see section H on Art. 101 TFEU) or if they are accused of abusing a dominant position (see section I on Art. 102 TFEU).

The starting point for the analysis of market power is the position of the parties on the affected markets. This requires a definition of the relevant market (see section F). In a second step, European legislation (EU 2004.b: para. 14) specifies that market "shares and concentration levels provide useful first indications of the market structure and of the competitive importance of both the merging parties and their competitors." The above merger would raise the concentration ratio by 5% to $CR_{3,A}$'=40% and the HHI by 250 to HHI_A'=1,250. The increase in HHI is referred to as **delta** (Δ). Similar statements can be found regarding the application of Art. 101 and 102 TFEU (see, e.g., EC (2011: para. 44)).

In either of these cases, one must specify a critical level of market power. In the case of horizontal mergers, European law (EU 2004.b: paras. 17-18) considers a combined market share of 50% or above evidence of a dominant position by the merged firm while a combined market share at or below 25% does not raise competition concerns. Moreover, markets with a post-merger HHI below 1,000 normally do not require extensive analysis. However, more intensive analysis is required for markets with (i) 1,000<HHI≤2000 and Δ≥250 or (ii) HHI>2,000 and Δ≥150 (EU 2004.b: para. 20).

These critical values are not necessarily uniformly defined even in a single jurisdiction. For example, the degree of market power required for the finding of an infringement under Article 101 (agreements between undertakings) of *the treaty on the functioning of the European Union* (TFEU) is less than the degree of market power required under Article 102 TFEU (abuse of dominance; EU 2004.a: para. 26). This is because the definition of a 'low combined market share' depends on the type of agreement or merger in question. "Depending on the market position of the parties and the concentration in the market, other factors such as the stability of market shares over time, entry

barriers and the likelihood of market entry, and the countervailing power of buyers/suppliers also have to be considered" (EC 2011: paras. 44+45). These points are explained in greater detail in sections E.2 andError: Reference source not found.

Estimating Market Power

Above we proposed a unique relationship between the value of the Lerner-index and market structure, in particular monopoly (equation (17)) and Cournot-oligopoly (equation (50)). Thus, we asked the question, how much market power does a firm possess if its conduct and the structure of the market perfectly match the underlying theory? However, empirical measures of the Lerner-index answer the question, how much market power does a firm exercise? The theoretical prediction and the empirical estimate of the Lerner-index may be different if, for example, firms not perfectly behave as proposed by the theoretic model or if there are countervailing factors (see subsection E.2) that are not perfectly considered in the model. In the following, we provide an overview on the empirical estimation of the above static measures (also see Perloff et al. 2007).

A typical SCP-analysis of market power would regress a performance measure on measures of industry-structure (e.g. concentration ratios or HHI). Typically, one of the following three performance measures is used (Perloff et al. 2007: pp. 14-19). First, one might use rates of return that – broadly speaking – are defined as a firm's operating profits relative to the value of its capital stock. This measure is imperfect as accounting measures of profitability not necessarily reflect economic profits. Second, one might use Tobin's q, i.e. the ratio of a firm's market value relative to the replacement cost of its assets. If a firm is worth more than it would cost to rebuild it, then it is earning an **excess profit**: a profit that is greater than the level necessary to keep the firm in the industry. This measure is imperfect because quantifying a firm's market value as well as its replacement costs is difficult. A third measure is the Lerner-index as defined above.

Quantifying the Lerner-index L_i for some firm i is difficult because a measure of marginal costs is rarely available. Therefore, researchers frequently approximate firm i's marginal costs by its **average variable costs** (AVC). However, this causes a bias in the estimate of L_i. To see this, suppose that marginal cost is constant and is given as

$$c = AVC + (r+\delta)\frac{p_k K}{q} \quad , \tag{56}$$

where r is the competitive rate of return, d is the depreciation rate. And AVC is the (constant) average variable cost of the labor and materials needed to produce one unit of output, q. Equation (56) describes a technology that requires K/q units of capital (at a cost of p_k per unit of capital) to produce one unit of output. By plugging (56) into (17) we find

$$\frac{p(q)-AVC}{p(q)} = \frac{1}{\eta}+(r+\delta)\frac{p_k K}{p\cdot q} \quad . \tag{57}$$

This shows clearly that approximating marginal costs by average variable costs results in a biased estimate of the Lerner-index. As a consequence, when regressing the price-*AVC* margin on concentration measures and other explanatory variables, one would have to include p_kK/pq as regressor in the estimation-equation (Perloff et al. 2007: p. 27).

Hosken et al. (2011) provide a **price-concentration study** that shows, first, how to conduct such a study in a convincing way and, second, that even then the results of a carefully conducted study should be treated with care. They estimate the price effects of two petroleum refinery transactions in the San Francisco Bay Area: Tosco's purchase of Unocal's Rodeo refinery in April 1997, and UDS's purchase of Tosco's Avon refinery in August 2000. The Tosco/Unocal merger reduced the number of Bay Area refiners from 5 to 4 (post-merger *HHI*=2702, Δ=574) and was expected to raise gasoline prices. The UDS/Tosco merger restored the area to 5 refiners (post-merger *HHI*=2167, Δ=-454) and was expected to lower gasoline prices.

In a first step, Hosken et al. (2011) **define the relevant market** (also see section F) as the market for CARB (California Air Resources Board) gasoline in the Bay Area. This narrow market definition is justified by the nature of CARB gasoline as a unique specification for the state of California. Moreover, they argue that there is little competition with refineries both in other US-American states and in other regions in California.

In a second step, the authors specify three **regressions** to identify the price effects of the mergers. First, they regress the difference in rack (or: wholesale) prices in San Francisco ($P_{SF,t}$) and Los Angeles ($P_{LA,t}$) on a constant a and monthly indicator variables M_t^i. Moreover, they include dummy variables for the 1st year ($Year\ 1_t$) and the 2nd year ($Year\ 2_t$) following the merger.

$$P_{SF,t}-P_{LA,t}=a+\sum_{i=1}^{11} b_i M_t^i+c_1\,Year\,1_t+c_2\,Year\,2_t+e_t \tag{58}$$

Equation (58) is estimated separately for each transaction using one year of daily pre-transaction data and two-years of post-transaction both for branded and unbranded wholesale prices. The coefficients c_1 and c_2 can in interpreted as the change in price (cent/gallon) in the first and second year following the transaction. They are hypothesized to be positive for the concentration-increasing Tosco/Unocal merger and negative for the concentration-decreasing UDS/Tosco merger. Second, Hosken et al. (2011) run a panel-regression for daily retail prices P_{it} in one of six cities i.

$$P_{it}=a_i+\sum_{i=1}^{11} b_i\cdot Bay_i\,M_t^i+\gamma_t+c_1\cdot Bay_i\cdot Year\,1_t+c_2\cdot Bay_i\cdot Year\,2_t+e_{it} \tag{59}$$

a_i is a city-specific fixed effect, Bay_i is an indicator equaling one if the city is in the Bay, and the γ_ts are daily indicator variables allowing for general demand or cost shocks. Equation (59) is estimated using one year of pre-transaction data and two years of post-transaction data. Third, a similar specification is estimated for station-specific – rather than city-specific – prices.

In a third step, the authors **interpret their results**. Table 2 presents the estimates of the coefficients c_1 and c_2 for the Tosco/Unocal merger (rows 1 and 2) and the UDS/Tosco merger (rows 3 and 4) for the 4 datasets. One finds that the signs of the estimated coefficients do not match our expectations. Hosken et al. (2011: 49) note that given "the coincident timing of UDS/Tosco with a number of significant shocks to refining capacity in the Bay Area and Southern California, it is difficult to isolate the retail price effects of this transaction." For example, outages of refineries caused high prices prior to the UDS/Tosco merger so that part of the estimated retail price decrease may be attributed to the abnormally high pre-merger price level. One finds that it is difficult to isolate the effects of changes in market structure from other factors, e.g. supply shocks, that also affect pricing. Moreover, it is also possible that the merger decision is endogenous to market conditions, such as demand shocks or previous mergers in the same market. Omission of variables correlated with prices and the decision to merge may result in biased estimates of merger impacts (Hosken et al. 2011: 46).

TABLE 1

	Wholesale price effects		Retail price effects	
Merger	Branded rack	Unbranded rack	City aggregate	Station level
(1) Tosco/Unocal Year 1	−3.59	−0.26	−0.68	−1.23
	(0.48)	(0.26)	(1.17)	(0.28)
(2) Tosco/Unocal Year 2	−1.34	1.39	0.16	0.51
	(0.47)	(0.30)	(1.14)	(0.30)
Observations	918	918	918	824,139
(3) UDS/Tosco Year 1	−2.87	−3.96	−4.05	−4.59
	(0.27)	(0.38)	(1.55)	(0.25)
(4) UDS/Tosco Year 2	0.52	−0.55	−4.43	−7.33
	(0.24)	(0.32)	(1.82)	(0.37)
Observations	903	903	10,675	1,616,317
(5) UDS/Tosco Year 1*	−5.45	−5.71	0.62	3.28
	(0.40)	(0.54)	(1.88)	(0.46)
(6) UDS/Tosco Year 2*	−2.11	−2.44	0.15	0.74
	(0.35)	(0.47)	(1.64)	(0.49)
Observations	903	903	10,675	1,582,344

Notes: Each regression uses one year of premerger and two years of postmerger data. Tosco/Unocal covers 4/96–3/99. UDS/Tosco covers 9/99–8/02. UDS/Tosco* covers 9/98–8/99 and 9/00–8/02. Standard errors for wholesale price regressions are calculated using Newey-West, standard errors for the city aggregate regressions are clustered at the city level, and standard errors for station level regressions are clustered at the zip code level.

Table 2: Results from Hosken et al. (2011: 48)

E.2 Persistence of Dominance

"The creation, maintenance or strengthening of market power can result from superior skill, foresight or innovation" (EC 2011: para. 41). In this context, recall our result from section C.2 that, e.g., in Bertrand-competition innovation may give a firm a cost-advantage (= superior skill), which allows this firm to drive its competitors out of the market and set monopoly-prices and -quantities. In other words, dominance is a result of asymmetry among firms. Therefore, explaining dominance requires to determine the causes of asymmetries among firms.

Industries that are characterized by supra-competitive profits are likely to attract firms that are currently outside the market. **Entry** into a market is likely to increase competition and reduce firms' markups on marginal costs (for example, see equation (26)). Similarly, **barriers to entry** cause the market power of the incumbent-firms to persist. Sometimes the threat to enter a market is sufficient to drive market prices down to competitive levels. This is analyzed by the theory of **contestable markets**. Moreover, the market power of suppliers can be offset by countervailing **buyer power**. These aspects are analyzed in this section E.2 .

Actual Entry into the Market

In section A.3 we propose that a monopolist charges an optimal price above marginal costs (see equation (16)).

$$p(q)=c(q)-\frac{dp(q)}{dq}\cdot q \tag{60}$$

Note that this is not a long-run equilibrium (see equation (14)) as long as the price is above average costs $AC(q)$. This is because the prospect to make economic profits attracts other firms to enter this industry. If entry is costless firms will enter the industry and, by thus, intensify competition. This drives prices down and entry stops when the condition

$$p(q)=AC(q) \tag{14}$$

is satisfied. Condition (14) denotes the **free-entry equilibrium** because (i) no active firm wishes to leave the market, and (ii) no inactive firm wishes to enter the market.

As a consequence, when costless entry is possible a firm's market power cannot persist in the long-run. Average costs are calculated by dividing total costs C through the output q, where total costs consist of fixed costs F and the sum of marginal costs $c(q)$ that are incurred in the production of q.

$$C = F + \int_{x=0}^{q} c(x)\,dx \tag{29}$$

The relevant questions are: What structure should we expect in a particular industry in the long run? Why can market power persist in some industries but not in others.

To illustrate firms' entry dynamics, consider the following Cournot-model. Each firm i has a cost function given by

$$C_i = F + cq_i$$

with constant marginal costs c and fixed costs F that are identical for all firms. The demand function and curve are given by

$$Q = (a - P)M \Leftrightarrow P = a - Q/M \quad ,$$

where M is a measure of market size. Calculating the Cournot-equilibrium yields the following optimal quantities, price, and profits.

$$\begin{aligned} q_i &= M \cdot \frac{a-c}{n+1} \\ Q &= n \cdot M \cdot \frac{a-c}{n+1} \\ P &= c + \frac{a-c}{n+1} \\ \pi_i &= M \cdot \left(\frac{a-c}{n+1}\right)^2 - F \end{aligned}$$

We find that an increase in the number of firms reduces each firm's individual quantity q_i but raises the aggregate quantity Q. This reduces both the price P and each firm's individual profit π_i.

Note that the free-entry equilibrium is determined by the number of firms n_{fe} that eliminates economic profits for all firms.

$$n_{fe} = (a-c)\sqrt{\frac{M}{F}} - 1$$

Thus, the number of firms in an industry in the free-entry equilibrium is an (i) increasing function of the market-size M and (ii) a decreasing function of both marginal and fixed costs (Cabral 2000: p. 243). However, the equilibrium-number of firms rises under-proportionally with the size of the market. This is because a higher number of firms causes an increase in competition, which leads to lower price-marginal cost margins. Moreover, given the assumption of symmetric firms, each firm possesses a market share of $m_i = 1/n_{fe}$. Therefore, the concentration ratio CR_4 is given by

$$CR_4 = \frac{4}{(a-c)\sqrt{M/F} - 1} \quad , \tag{61}$$

and concentration declines as market size grows (Pepall et al. 2008: ch. 4.4.1).

The free-entry number of firms need not necessarily be socially optimal, i.e. welfare-maximizing. This is the case when production requires fixed costs (as in the above example) or entry requires sunk costs. It can be shown that the welfare-maximizing number of firms is smaller than n_{fe}. This is because of a business stealing effect as long as $n<n_{fe}$ a firm that enters the market may make positive economic profits by stealing quantity from its rivals and, thus, lowering their profits. This raises the consumer surplus *CS* that is increasing in the number of firms.

$$CS=\frac{M}{2}\left(\frac{n\cdot(a-c)}{n+1}\right)^2$$

However, it does not necessarily raise the sum of profits, as additional firms come along with additional fixed costs that are welfare-detrimental. In the long-run equilibrium this effect is present so that total welfare is not at its maximum (Cabral 2000: ch. 14.3).

"[T]he negative externality that an additional entrant imposes on existing firms by taking business from them outweighs the positive externality to consumers in terms of lower price. In this very simplified setting, then, competition is good for allocative efficiency but bad for productive efficiency, and the net effect is a tendency for there to be too much entry" (Vickers 1995: p. 14).

Bresnahan and Reis (1991) provide empirical results on entry and competitive conduct. They assume that firms incur fixed costs and that marginal costs rise in output – rather than being constant as in the above case. This makes the average total costs of the firms u-shaped. They analyze the effect of entry on competitive conduct in 202 isolated local markets in USA in the following five industries: doctors, dentists, druggists, plumbers, and tire dealers. Among others, they (*ibid.* p. 988) observe that the population in each market imperfectly predicts the number of dentists. They assume that market size and, thus, entry decisions also depend on variables such as expected future population growth, market demographics, changing economic conditions, consumer incomes, and factor prices. This underlines the importance of a sound market delineation both for academic studies as well as practical competition policy. They find that a market must not be too small for firms to enter. Their estimates suggest "that a monopoly tire dealer or druggist requires about 500 people in town to set up business. A monopoly doctor or dentist needs between 700 and 900 people. Monopoly plumbers require at least twice what monopoly doctors or dentists need to break even" (*ibid.* p. 995). Their results show that markets become more competitive when the number of firms rises. However, once there are three firms in a market additional firms do not make the market more competitive any more. Most "of the increase in competition comes with the entry of the second and third firms. These results initially surprised [the authors. They] expected to find entry threshold ratios that declined more gradually. It instead appears that the competitive effect of

entry occurs rapidly [...]. Whether this pattern appears in other industries remains an open question. " (*ibid.* p. 1007).

Barriers to Entry

At least some industries are characterized by firms that can exercise market power, which is inconsistent with the prevalence of a free-entry equilibrium. Therefore, we may conclude that in these cases firms face barriers to entry. In the following, entry barriers are analyzed in greater detail.

One such entry barrier is the existence of **economies of scale**. We have shown above that firms in the free-entry equilibrium produce at their minimum efficient scale. At this output firms cannot lower their average costs any further by producing a higher quantity. This situation is characterized by the absence of economies of scale. However, consider the case where a firm incurs a fixed cost F and produces with constant marginal costs ($c(q)=c \ \forall \ q$). This cost function exhibits economies of scale for every output q. As a result, productive efficiency is at its optimum when the entire output is produced by a single firm. This is the case of a **natural monopoly**.

Moreover, **first-mover advantages** can create barriers to entry. Suppose, firm i enters a market earlier than some firm j. This enables firm i to achieve learning effects – i.e. it moves down the learning curve – and produce at lower costs than firm j. Thus, first-mover advantages can contribute to the creation of **technological differences**, i.e. firms produce with asymmetric production technologies. This also indicates that a market's structure may to some extent be explained by the **history of the industry**. A further element in this context are changes in demand. For example, consider an industry with high demand in the past that has decreased over time. Therefore, a number of firms that was adequate for the past may be inefficiently high (especially with regard to productive efficiency) in the present situation.

If the **sunk costs** of entry into a large market exceed the entry costs into a small market (**endogenous entry costs**) the number of firms in the free-entry equilibrium is less responsive to market size. For example, this is the case when entry requires advertising or R&D. Advertising is certainly more expensive in large markets than in small markets.

In **summary**, we may conclude that concentration is greater the higher the barriers to entry. The above definition of entry barriers combines the views of the Chicago School and the Structure-Conduct-Performance (SCP) School. From the viewpoint of the SCP paradigm, e.g., economies of scale, product differentiation, and cost advantages constitute barriers to entry. From the viewpoint of the Chicago school, entry barriers only exist when it is more costly for an entrant to enter the market than for the incumbent (Monti 2007: p. 64; see also section D.2). In addition to the existence of entry barriers, incumbent-firms may actively engage in strategic behavior with the aim

to keep entrants outside the market. Cabral (2000: ch. 15) provides an overview on such behavior. Exclusionary practices are discussed in section I . The consideration of entry withing merger-cases is decribed in section G.2 on p. 166.

Potential Entry into the Market (Contestable Markets)

Above, we have shown that market power cannot persist when costless entry is possible *and* rivals of a dominant firm actually enter the industry. However, under some conditions it is not even necessary that rivals enter the industry. In these cases, the mere threat of rivals entering the industry suffices to drive prices down to competitive levels. Markets where potential competition is sufficient for ensuring a competitive market outcome are called **contestable markets** (Motta 2004: p. 73). To see this, suppose the following assumptions.

1. The market is served by an incumbent monopolist.
2. This monopolist cannot adjust its price p after announcing it at the beginning of the period.
3. The monopolist is not capacity-constrained and can serve the entire market.
4. If a competitor enters the market it sets a price slightly below the price of the monopolist and steals the entire market (price-competition).
5. Entry into and exit from the market are completely costless.

If the incumbent monopolist sets a price above average costs ($p_i>AC$) it makes positive profits $\pi_i>0$. This attracts the entrant who sets a price $p_e=p_i-\varepsilon$ and wins the entire market ($\pi_e>0$), while the monopolist incurs a loss equaling its fixed costs $\pi_i=-F$. If the monopolist sets a price equaling average costs ($p_i'=AC$) it makes zero economic profits $\pi_i'=0$. Therefore, it is not profitable for the entrant to enter the market.

		entrant	
		"in" & $p_e=p_i-\varepsilon$	"out"
incumbent	$p_i>AC$	$\pi_i=-F-\pi_e>0$	$\pi_i>0-\pi_e=0$
	$p_i=AC$	$\pi_i=0-\pi_e<0$	$\underline{\pi_i=0}-\underline{\pi_e=0}$

This result hinges crucially on assumption 5 that the entry into the market is completely costless. However, if entry requires a **sunk cost** S the incumbent would set a price $p_i''=AC+S/q$ such that the entrant would make zero profits. This prevents the entrant from entering the market but allows the monopolist to make positive economic profits $\pi_i''>0$. Therefore, the existence of sunk

costs of entry makes hit-and-run-entry by the potential competitor more difficult and, thus, ensures the incumbent positive economic profits. Sunk costs also occur if the entry into the market requires an investment in (physical) assets that cannot be sold at the purchasing price any more when leaving the industry. The more specific the physical assets, the greater the extent to which the investment in the assets is sunk (Martin 2000: p. 17). Sunk costs can also occur in the form of intangible assets such as advertising or R&D expenditures. Such investment is specific to the industry and cannot be resold.

Additionally, the result of the monopolist setting its price such that the entrant would make zero economic profits depends on assumption 4 that the firms would engage in **price-competition**. However, if the firms would engage in, e.g., Cournot-competition they could both make positive economic profits. Therefore, accepting an entrant in the market but making positive economic profits maybe expected to be the lesser evil for the incumbent than keeping the entrant outside the market and making zero economic profits. Moreover, one crucial assumption of the Bertrand-model that leads to the result of prices equaling marginal costs is the homogeneity of products (see section B.2). If the buyers have a preference for differentiated products and firms offer differentiated products they can charge a premium for this service. As a consequence, an entrant that offered a modified product would be able to profitably enter the industry and co-exist with the incumbent. In this case, actual competition rather than potential competition could raise welfare.

A further crucial assumption is assumption 2 that the incumbent cannot alter its price after announcing it at the beginning of the period. If the incumbent was able to make a **strategic response** and change its price immediately after the entry by the incumbent the monopolist would be able to sustain the profit-maximizing price p^* from equation (60). To see this, suppose that the monopolist starts with charging p^*. If the entrant entered the industry the monopolist would instantaneously lower its price to p_i'' and, thus, make entry unprofitable. As this behavior is anticipated by the entrant, no entry will occur and the monopolist charges p^*.

However, the contestability-result can occur in such a market-environment where the entrant has the ability to commit to long-term contracts with its potential buyers before entry. To see this, suppose the incumbent would set the monopoly-price $p_i=p^*$. The entrant may now engage in a long-term contract with the buyers to supply the good at $p_e=p_i-\varepsilon$. It would win the entire market without having to be afraid of retaliation by the incumbent. The only way for the incumbent to prevent entry is to set prices such that the entrant would make zero economic profits.

In summary, we can say that in some cases the threat of hit-and-run-entry is sufficient to drive prices in concentrated markets down to competitive levels. However, this is not the case when entry into the market is costly and when the incumbent can respond to entry quickly by adjusting its

prices. Additionally, product-differentiation allows for profitable entry into the market, which creates actual competition and improves welfare. One piece of evidence for the price-effects of potential competition is provided by Kwoka and Shumilkina (2010):

> *"This study examines the 1987 merger between USAir (now US Airways) and Piedmont Airlines. Our key question is:* ***What is the effect on price when an incumbent firm merges with a potential entrant into the same market?*** *[emphasis added]*
>
> *A* ***potential competitor*** *[emphasis added] on a route is conventionally defined as a carrier that, while not serving the route, operates at either or both endpoint cities. Such a carrier is viewed as positioned for relatively quick and easy entry since it has (a) feed or connecting traffic for the route in question, (b) ground infrastructure, such as gates, terminal space, and baggage handling, and (c) first-hand information about and perhaps marketing investment in the route endpoint*
>
> ***Prior to the merger*** *[emphasis added], both USAir and Piedmont were large regionally-based carriers. While USAir's operations concentrated in the Northeastern U.S. and Piedmont's in the Southeast, each had extensive route networks with considerable overlap. In January, 1987, USAir and Piedmont announced their intention to merge. Financial aspects of the merger were completed in November of that year, but both regulatory and labor issues delayed integration of their operations until August, 1989.*
>
> *Our analysis indicates that where one of the carriers was a potential entrant into a route served by the other merger partner, the USAir-Piedmont merger allowed the incumbent to raise price by an amount between 5.0 and 6.0 per cent. This increase is statistically significant in all model specifications and establishes that* ***the elimination of a potential competitor can have a very considerable impact on market price*** *[emphasis added]. This result should be compared to the price change on routes where both carriers had previously operated.* ***There the merger eliminated an actual competitor, and prices increased between 9.0 and 10.2 percent, nearly twice the increase from eliminating a potential competitor*** *[emphasis added]."*

Buyer Power (Monopsony)

Buyer power may have a countervailing effect on the market power of sellers. For example, leading retail firms with a significant market share may possibly obtain more favorable terms, i.e. discounts, from suppliers than those available to other buyers. Moreover, firms who agree on purchasing

jointly may exercise buyer power. The treatment of such purchasing agreements in European competition policy is described in section H.3 . One example of a purchasing agreement is the German supermarket corporation EDEKA (*Einkaufsgenossenschaft der Kolonialwarenhändler*). The EDEKA-retailers belong to regional purchasing cooperatives which own the EDEKA head office. This head office controls the joint purchasing of goods.[43] Other manifestations of buyer power are the following (Dobson et al. 2002: p. 248-249):

- **Listing charges** – The buyer requires the payment of a fee before goods are purchased from the listed supplier.
- **Slotting allowances** – The buyer charges a fee for allocating shelf-space to the supplier.[44]
- **Retroactive discounts** may be awarded on goods already sold.
- **Most favored nation** (MFN) **clauses** – The supplier is obliged not to sell to another retailer at a lower price.
- **Exclusive supply**

The price-effects of buyer power can nicely be illustrated using the model of monopsony. This model makes clear that one needs to take into account two markets, an upstream- (or purchasing-) and a downstream- (or selling-)market, when analyzing the effects of buyer power (EC 2011: para. 197).[45] To see this, consider the case of a retailer i that buys a quantity q of some product (**upstream**-level) and sells this quantity to its customers (**downstream**-level). The suppliers in the upstream market sell the good at price p_u and the supply-curve is upward sloping.

$$p_u = S_u^{-1}(q) \text{ with } \frac{dS_u^{-1}(q)}{dq} > 0 \tag{62}$$

For the moment, we assume that the retailers are perfectly competitive. This means, they sell the good at prices equaling their marginal costs (see equation (12)).

$$S_u^{-1}(q) = c(q) \text{ so that } \frac{dS_u^{-1}(q)}{dq} = \frac{dc(q)}{dq} \tag{63}$$

The retailer sells the good at price p_d in the downstream market and the demand-curve is downward sloping.

$$p_d = D_d^{-1}(q) \text{ with } \frac{dD_d^{-1}(q)}{dq} < 0 \tag{64}$$

43 http://www.edeka.de/EDEKA/Content/Unternehmen/Profil/Unternehmensstruktur/index.jsp

44 In this context, also see paras. 203-208 in the EU's Guidelines on Vertical Restraints (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52010SC0411:EN:NOT).

45 A further model of vertical structures is provided in section G.3 .

The profit of retailer *i* is

$$\pi_i = p_d \cdot q - p_u \cdot q \quad . \tag{65}$$

Maximizing this profit with respect to quantity *q* yield the following first-order condition.

$$\begin{array}{ccc} D_d^{-1}(q) + \dfrac{dD_d^{-1}(q)}{dq} \cdot q & = & S_u^{-1}(q) + \dfrac{dS_u^{-1}(q)}{dq} \cdot q \\ MR_d(q) & = & MC_d(q) \end{array} \tag{66}$$

The retailer maximizes its profit at the quantity that equalizes its marginal revenue (*MR*) to its marginal costs (*MC*). We may now analyze market outcomes in the following three cases. In all cases, we maintain the assumption that the upstream sellers behave perfectly competitive.

(A) If the upstream-market and the downstream-market are **perfectly competitive**, retailer *i* can neither affect its purchasing-price p_u (i.e. $dS_u^{-1}(q)/dq=0$) nor its selling-price p_d (i.e. $dD_d^{-1}(q)/dq=0$). The market clears according to the following condition.

$$D_d^{-1}(q) = S_u^{-1}(q) \tag{67}$$

This is shown by point A in Figure 18.

(B) If retailer *i* is a **monopsonist** – i.e. its suppliers deliver to *i* only – and the downstream-market is perfectly competitive – i.e. other retailers offer substitute products from other suppliers – *i* cannot affect the downstream price p_d (i.e. $dD_d^{-1}(q)/dq=0$). However, its purchasing-decision has an effect on the purchasing price p_u. If the retailer buys a higher quantity the marginal costs of the supplier rise. This raises the price that the upstream supplier charges from the downstream buyer (i.e. $dS_u^{-1}(q)/dq>0$). The market clears according to the following condition.

$$D_d^{-1}(q) = MC_d(q) \tag{68}$$

This is shown by point B in Figure 18. This market outcome is characterized by higher selling-prices ($p_{d,B}$) in the downstream market and lower purchasing-prices ($p_{u,B}$) because of a reduced quantity (q_B). The light-gray triangle shows the welfare-loss ΔW_B that arises because of the existence of buyer power. This welfare-loss is the larger the more inelastic is supply. Hence, in the case of a horizontal or even downward-sloping supply-curve (e.g. as a result of economies of scale) buyer power does not cause a loss in welfare.

(C) If retailer *i* is both a **monopsonist** in the upstream-market *and* a **monopolist** in the downstream-market, its output decision affects both the purchasing- and the selling-price. The market clears according to condition (66). This is shown by point C in Figure 18. The light-gray triangle and the dark-gray trapezoid show the welfare-loss ΔW_C that arises

because of the existence of buyer power *and* seller power.

If retailer *i* is not the single but one of few buyers, its is called an **oligopsonist**. In analogy to oligopoly-theory, the market outcome in oligopsony is somewhere in between the monopsonistic and the perfectly competitive market outcome. For example, this is the case when some firms are part of a joint purchasing agreement while others are not. In this case, an additional problem arises, i.e. "[b]uying power of the parties to the joint purchasing arrangement could be used to foreclose competing purchasers by limiting their access to efficient suppliers" (EC 2011: para. 203).

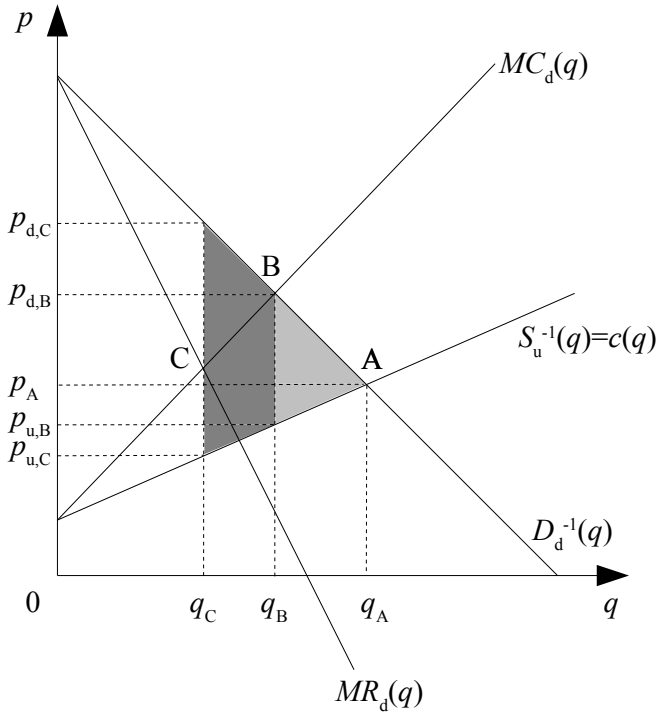


Figure 18: Monopsony-buyer with competitive seller

So far we have been concerned with a perfectly competitive upstream market. In this case, the existence of buyer power downstream led to higher prices, a lower quantity, and thus a lower consumer surplus. Let us consider the case with **market power upstream** and **buyer power downstream**. Retailer *i* is a monopolist downstream that faces a downward-sloping demand $D_d^{-1}(q)$ according to equation (64). Its profit-function looks like equation (65) so that its profit-maximum is determined by condition (66).

$$\begin{aligned} D_d^{-1}(q)+\frac{dD_d^{-1}(q)}{dq}\cdot q &= S_u^{-1}(q)+\frac{dS_u^{-1}(q)}{dq}\cdot q \\ MR_d(q) &= MC_d(q) \end{aligned} \tag{66}$$

In the perfect competition in the upstream market, the suppliers upstream offer the good at a price

equaling marginal costs. This means that the supply curve and the marginal-cost curve are identical. This is not the case when the upstream-seller is a monopolist itself.

To see this, note that the upstream seller faces a demand-curve $D_u^{-1}(q)$, which can be derived from condition (66) as follows.

$$D_u^{-1}(q) = MR_d(q) - \frac{dS_u^{-1}(q)}{dq} \cdot q \tag{69}$$

We arrive at equation (69) by noting that the upstream monopolist supplies a quantity q at price $p_u = S_u^{-1}(q)$ (see equation (62)). At this price p_u the downstream retailer demands quantity $q = D_u(p_u)$, which is equivalent to $D_u^{-1}(q) = p_u$. It is easy to show that the upstream monopolist attains its profit maximum

$$\pi_u = D_u^{-1}(q) \cdot q - \int_0^q c(q) \tag{70}$$

when the following standard condition is satisfied.

$$MR_u(q) = c(q) \tag{71}$$

We analyze the following cases.

(D) If the retailer is a monopolist but **no monopsonist** its purchasing decision does not have an effect on its purchasing-price p_u (because $dS_u^{-1}(q)/dq = 0$). Equation (69) simplifies to

$$D_{u,D}^{-1}(q) = MR_d(q) \quad . \tag{72}$$

Let $MR_{u,D}(q)$ denote the marginal revenue that is associated with demand-curve $D_{u,D}^{-1}$. In this case, the upstream monopolist attains its profit maximum at point D in Figure 19. The welfare-loss ΔW_D of having monopolists upstream and downstream in comparison to perfect competition (point A) is the sum of the dark-gray trapezoid and the light-gray triangle in Figure 19.[46]

(E) If the retailer is a monopolist and a **monopsonist** we assume they agree on setting a quantity level that maximizes their joint profits

$$max\ \pi = \pi_l + \pi_u \tag{73}$$

and then divide the spoils through bargaining. The joint profit-maximum is achieved at point E in Figure 19 where the condition

$$MR_d(q) = c(q) \tag{74}$$

46 Issues related to a monopoly both upstream and downstream are also described in the context of vertical mergers in section G.3 .

applies. This is the profit-maximum of a vertically integrated monopolist. To achieve this solution, the upstream-monopolist might set a two-part tariff t with a fixed component T and a variable component $p_{u,E}$.

$$t = q_E \cdot p_{u,E} + T \tag{75}$$

The two firms would have to bargain the size of $T \in [0;(p_{d\text{-}E}\text{-}p_{u,E})q_E)]$. The welfare-loss ΔW_E of this bargaining-solution in comparison to perfect competition is the light-gray triangle in Figure 19.

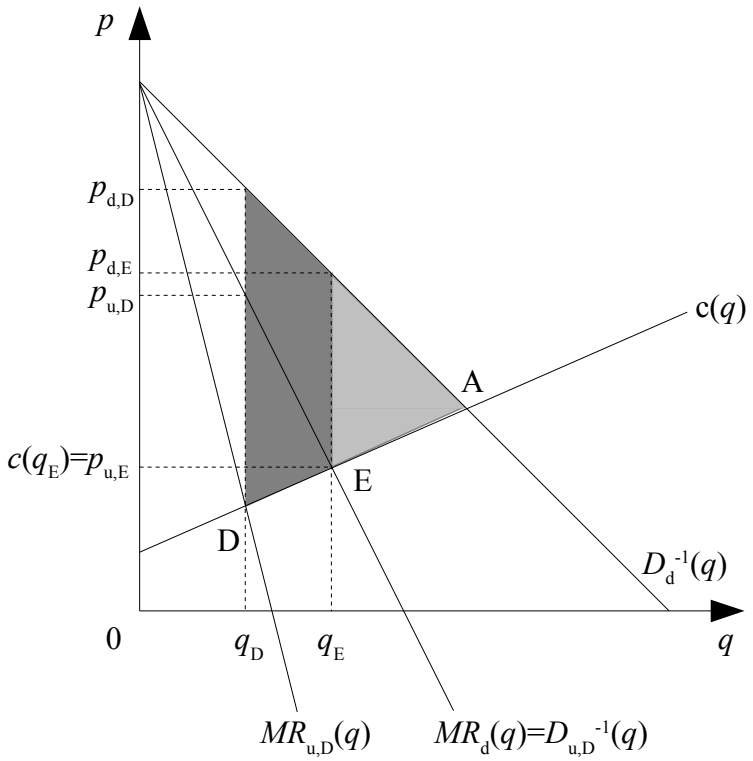


Figure 19: Monopsony-buyer with monopolistic seller

What have we learned from this analysis of buyer power? First, suppose the producer of some good is a competitive firm (in the upstream market) and a retailer (in the downstream market) possesses buyer power (situation B). In this case, the retailer buys the good from the producer at a lower purchasing price and sells it at a higher retail-price as in a situation of perfect competition. Welfare is lower than in perfect competition and the producer makes zero economic profits. Second, if the retailer possesses market power in the downstream market these effects are even severe (situation C). In this context the European Commission (2011: para. 201) notes: "If the parties have a significant degree of market power [...] on the selling market [...], the lower purchase prices achieved by the joint purchasing arrangement are likely not to be passed on to consumers." Moreover, "joint purchasing arrangements are less likely to give rise to competition concenrns

when the parties do not have market power on the selling market" (EC 2011: para. 204).

Third, suppose the producer of the good is a monopolist in the upstream market and the retailer is a monopolist in the downstream market but does not possess buyer power (situation D). In this case, prices are higher and welfare is much lower than under perfect competition. The producer makes economic profits. Fourth, if in this situation the downstream monopolist possesses buyer power the firms can jointly negotiate an outcome that increases consumer surplus and both firms' profits. Summarizing we can argue that buyer power is likely to be detrimental to welfare when markets would otherwise be competitive. However, the existence of buyer power can raise welfare when it is used to offset the market power of upstream firms.

One example for the supposed existence of buyer power is the ***German market for dairy products****. On the one hand, grocery retail-chains may possess buyer power with respect to (competitively behaving) dairies. This would cause the retail-prices of milk to be high while the price that is paid to the dairies is low (situation B in the above model). On the other hand, (competitively behaving) farmers argue that they often face (local) monopsonist respectively oligopsonist dairies resulting in low prices for raw milk as well as the necessity to commit to exclusive supply-contracts (BKartA 2009: p. 12).*

This complicated market structure is suitable to illustrate a central concept of competition law (in Germany). The Bundeskartellamt (2012: 14) makes clear that competition law is not designed to protect certain groups of market participants (for example, producers or consumers) but to protect and ensure the ***process of competition*** *as such.*

The Bundeskartellamt notes concerning the ***relationship between farmers and dairies*** *that concentration has increased over the last years because of dairy-mergers and the acquisition of German dairies by foreign firms (BKartA 2012: 19). For example, the Danish dairy Arla attempts to establish a position as one of the three largest milk processing firms in Germany (BKartA 2012: 20). In its assessment of these mergers the Bundeskartellamt identifies two distinct markets: the market for raw milk, i.e. the buying (or: upstream) market of the dairies, and markets for different types of processed dairy products, i.e. the selling (or: downstream) markets for dairies (BKartA 2012: 21). The Bundeskartellamt (2012: 42) notes that none of the dairies possesses a dominant position in the (regional) upstream markets. Moreover, an assessment of collective dominance by more than one dairy can only be made in specific cases but not on an abstract level.*

*In this context, consider that the market or buyer power of dairies may result from **cooperation** among them. The firms cooperate in the form of purchasing, production, or research and development agreements (BKartA 2012: 30). While these agreements can have positive efficiency effects especially with regard to productive and dynamic efficiency, they can also have negative effects when the dairies coordinate their conduct towards farmers and/or retailers (see section H).*

*The detection of a possible **abuse of a dominant position** by one or more dairies (see section I) is difficult for two reasons (BKartA 2012: 44). First, the abuse of a dominant position typically refers to a dominant upstream firm that charges excessive prices from its downstream buyers. Therefore, there is little experience in addressing cases where a supposedly dominant buyer pays lower than competitive prices to its upstream sellers. Second, establishing the abuse of a dominant position requires determining the price level in competition absent the abuse. Such counterfactual prices can be determined by comparing the prices in a particular market to prices in a clearly competitive market. Accounting for the differences in these markets makes price comparisons a difficult task.*

*The Bundeskartellamt (2012: 60) argues that so called **reference price models** may cause a distortion of competition. According to these models, dairies guarantee farmers a minimum price that equals the average price paid by other dairies. Reference price models distort competition if they cause a coordination of dairies' pricing behavior that leading to lower prices. One of their effects is to prevent the negative consequences that accrue to dairies when farmers exploit differences in milk prices and switch dairies. By creating more similar prices, reference price models have the potential to prevent such switching and harm farmers.*

*The analysis of buyer power in the milk market is also difficult because farmers often own the dairies via cooperatives and receive a dividend from these dairies. Therefore, it is not clear why the dairies would pursue a strategy that harms their owners (BKartA 2009: p. 78). In these cooperatives, the farmers supply milk to the dairy and receive a payment with a lag of one month. This **"Milchgeld"** (milk money) is calculated as a farmer's share in the profit of the dairy.*

*The **relationship between dairies and retailers** is complex because the suppliers of dairy products face oligopsonistic buyers. Following our above model of buyer power the effects of this market structure depend on the market power of*

the dairies. If the dairies compete intensely in the output market the buyer power of the retailers is likely to harm both dairies and final consumers. However, if the dairies possess market power themselves the true situation might be more like the bargaining-situation E above.

The Bundeskartellamt (2012: 87) cannot find conclusive evidence of the (ab)use of ***buyer power by grocery chains****. For example, dairies only sell 40% of their products to German grocers while 44% are exported (BKartA 2012: 87). However, it does neither preclude that there is no evidence of buyer power because German dairies make 42% of their revenues by selling to German grocers. 30% of revenues are generated by exports and 28% by means of sales to other German buyers (BKartA 2012: 89).*

It is interesting to note that there is no particular difference between the producers of ***branded and unbranded dairy products*** *regarding their ability to sell to different groups of buyers. The producers of branded products often cannot profitably produce a greater share of unbranded products because they cannot evade the fixed (or sunk) costs for advertising its branded product. Moreover, producing a greater share of unbranded products would be likely to harm the image of the branded product. As a consequence, the dairies have a limited ability to pass on cost-increases to grocery chains by means of higher prices (BKartA 2012: 92).*

Dairies cannot necessarily ***switch buyers*** *in order to achieve better prices. This is because especially non-cooperative ("nicht-genossenschaftlich") organized dairies and producers of branded products already sell to all relevant grocery chains. The export of dairy products is only an imperfect alternative to selling in Germany because of, e.g., differences in tastes, transport costs, or a short shelf-life of products.*

There is some (faint) evidence that the dairies have to accept ***long payment periods****, i.e. they are typically paid at a time later than the sale of the product to final consumers by the grocer. The length of this delay varies between a few days and several weeks (BKartA 2012: 103). In other words, the dairies (have to) grant the grocery chains a credit free of interest.*

Switching Costs

Another situation where market power does not necessarily decrease under free entry arises when there exist consumer **switching costs** (Motta 2004: ch. 2.6.3.2). Switching costs are costs that a

buyer must bear in consequence of switching from the consumption of good A to good B. Thus, a consumer's current purchasing-decision depends on a previous investment that he has made by purchasing good A in the past. Switching costs encompass the following types of costs:

- **Transaction costs** – They arise when a customer, for example, closes his bank account and opens a new account at another bank.
- **Learning costs** – This customer must learn how to operate the new online banking software.
- **Need for compatibility with existing equipment** – If the new online banking software is not compatible with the customer's existing operating system (OS) the customer must purchase a new OS.
- **Artificial/contractual switching costs** – Some firms use bonus-programs to create some customer loyalty. For example, *Volksbank Mittelhessen* offers its customers a *Mehrhaber*-bonus[47] that includes a ticket-buying service for cultural events, a free locksmith service, cinema-discounts and many more. Customers would have to forsake these services by switching to another bank.
- **Uncertainty about the quality of untested brands** – Additional (opportunity) costs occur when the new bank has not been operating in this market until previously. In this case the quality (and hidden costs) of its services can hardly be anticipated. One might also think of such goods as **experience goods** whose quality can only fully be appraised after purchasing them.
- **Psychological costs of switching** - "Social psychologists cite evidence that people change their own preferences in favour of products that they have previously chosen or been given in order to reduce "cognitive dissonance" (Brehm (1956))" (Klemperer 1995: p. 518). Therefore, people tend to prefer the products that they know to the products that they do not know.

As a consequence of the existence of switching costs, products that are *ex ante* identical will *ex post* become differentiated. When a customer opens his first bank account he will find that most banks offer pretty much the same service, i.e. they offer a relatively homogeneous product. However, after having decided for a bank the account holder will to some extent prefer this bank's products to that of other banks because switching would impose the above types of costs on him. This indicates that switching costs are not only relevant for the purchase of a certain good or service, e.g. a bank account. They rather determine the choice for a variety of related products because a customer might be tempted to hold a savings account at the same bank where he already

47 http://www.mehrhaber.de

holds a checking account. Therefore, switching costs may be seen in analogy to **brand loyalty** and in the context of **follow-on goods** in **aftermarkets**.

Farrell and Klemperer (2007: ch. 2.2) present some empirical evidence on switching costs. For example, the switching costs for breakfast cereals are estimated to be $3.43, which is even above the price for these cereals. With regard to the above banking-example, the cost of switching banks (in Finland) is found to vary between 0 and 11% of the average balance on the account.

In a market with switching costs, a firm's current market share is an important determinant of its future profitability (Klemperer 1995: 515). Therefore, when a new market is established (period 1), firms will compete for market shares where $m_{i,1}$ denotes the market share of firm i in period 1. In period 2, the existence of switching costs allows firm i to raise its price and "harvest" the customers that were attracted in period 1. To see this in a 2-period model, consider that firm i aims to maximize its present value of profits

$$V_i = \pi_{i,1} + \delta \pi_{i,2}\left(m_{i,1}\right) \quad , \tag{76}$$

where $\pi_{i,1}$ and $\pi_{i,2}$ denote firm i's profits in periods 1 and 2 with δ being the firm's discount factor. We assume that the firm's profits in period 2 depend positively on its market share in period 1, i.e.

$$\frac{\partial \pi_{i,2}}{\partial m_{i,1}} > 0 \quad , \tag{77}$$

while this market share is decreasing in firm i's first-period price

$$\frac{\partial m_{i,1}}{\partial p_{i,1}} < 0 \quad . \tag{78}$$

Then, maximizing the present value of profits with respect to the first-period price yields the following first-order condition.

$$\frac{\partial V_i}{\partial p_{i,1}} = \frac{\partial \pi_{i,1}}{\partial p_{i,1}} + \delta \frac{\partial \pi_{i,2}}{\partial m_{i,1}} \frac{\partial m_{i,1}}{\partial p_{i,1}} \stackrel{!}{=} 0 \tag{79}$$

Given (77) and (78), the second summand in (79) is negative. This implies

$$\frac{\partial \pi_{i,1}}{\partial p_{i,1}} > \frac{\partial \pi_{i,1}}{\partial p^*_{i,1}} = 0 \quad , \tag{80}$$

where $p^*_{i,1}$ is the price that firm 1 would set in period 1 if the profits in the second period would not matter. This suggests that second-period switching costs lower first-period prices. A similar result can be obtained for models with more than two periods (Klemperer 1995: p. 525).

Despite switching costs having a dampening effect on first-period prices, second-period prices are usually higher than in a case without switching costs. This is because the existence of

switching costs and, thus, lock-in effects makes demand more inelastic in period 2.

When firms anticipate that attracting new customers is more difficult once a market has been divided among some incumbent firms, they might compete more fiercely for the market. Therefore the presence of switching costs can explain price wars when (i) new markets open, (ii) a new group of customers enters the market, or (iii) new firms enter the market. With regard to the latter point, firms entering an industry have to offer lower prices than the incumbent firms in order to attract some of the incumbents' customers. Additionally, incumbent firms face a trade-off between maintaining their market share by lowering its price or harvesting profits by charging high prices that, however, run down the firm's existing stock of market share.

We have now established the finding that the existence of switching costs can have pro-competitive effects when, e.g., a new market opens, and anti-competitive effects in the periods thereon. What can we say about the aggregate effects of switching costs? In most cases a market is established just once (intense competition) and then exists for many periods (reduced competition because of switching costs). Therefore, the anti-competitive effects of switching costs usually dominate the pro-competitive effects. Klemperer (1995: p. 536) concludes that switching costs generally raise prices and create substantial welfare losses. They may also discourage new entry and so further reduce the market's competitiveness. Switching costs reduce the product variety available to consumers by reducing firms' incentives to differentiate their products. Finally, because switching costs tend to reduce competition, firms may dissipate more social surplus in costly activities to create them. These conclusions suggest that public policy should discourage activities that increase consumer switching costs and encourage activities that reduce them.

Farrell and Klemperer (2007: fn. 89) provide an example where switching costs were an important element in blocking a merger of two banks:

> *"[T]he UK Competition Commission in July 2001 blocked the proposed merger of two banks, Lloyds TSB and Abbey National, even though Abbey National accounted for only 5 per cent of the market for personal banking. An important part of the Commission's reasoning was that consumer switching costs, combined with some scale economies, make new entry very hard, and that existing firms with low market shares tend to compete more aggressively than larger firms in markets with switching costs, so smaller firms are particularly valuable competitors to retain."*

Network Effects

Switching costs arise when a consumer wants his current purchases to be compatible with his past purchases. **Network effects** (Motta 2004: ch. 2.6.3.3) arise when a consumer wants his purchases to

be compatible with the purchases of other users. This creates economies of scope between different users' purchases. We say that there are network effects if one agent's adoption of a good (i) benefits other adopters of the good (total effect) and (ii) increases others' incentives to adopt it (marginal effect). Moreover, we distinguish two types of network effects.

- **Direct network effects** – "A good exhibits direct network effects if adoption by different users is complementary, so that each user's adoption payoff, and his incentive to adopt, increases as more others adopt. Thus users of a communications network or speakers of a language gain directly when others adopt it, because they have more opportunities for (beneficial) interactions with peers" (Farrell and Klemperer 2007: p. 1974). Such direct network effects refer to physical (e.g. telecommunications) networks.
- **Indirect network effects** – Sometimes, markets are more efficient when the number of buyers is large. To see this, consider that an increase in buyers (i) might attract more sellers. This intensifies competition among sellers and makes the market allocatively more efficient. Alternatively, (ii) when firms produce with economies of scale the additional quantity produced might lower average production costs. However, this (iii) need not always be beneficial for the buyers, as the increase in their number shifts out demand. This effect may dampen or even overcompensate the effect of additional sellers or economies of scale on prices. Such indirect network effects refer to virtual (i.e. non-physical) networks.

Farrell and Klemperer (2007: ch. 3.2) provide some cases in which network effects are important:

- ***Telecommunications*** *– In this industry, interconnection is an important topic, i.e. a call originated on one network can be completed in another network. In some cases, dominant firms in this industry refused to interconnect with non-dominant firms in order to strengthening their dominant position. To some extent, such problems can be solved by regulatory agencies by issuing standards for telecommunication systems which ensure interconnection.*
- ***Microsoft*** *– "Because they have many users, Microsoft's operating system platforms attract a lot of applications programming. An indirect network effect arises because application software writers make it their first priority to work well with the dominant platform [...]. The U.S. antitrust case against Microsoft relied on this network effect or "applications barrier to entry", but did not claim that Windows is "the wrong" platform. Rather, Microsoft was convicted of illegal acts meant to preserve the network barrier against potential weakening through the Netscape browser and independent "middleware" such as Java" (Farrell and Klemperer 2007: p. 2010).*

- ***Betamax vs. VHS*** *– After video recording systems had been developed in the 1970ies three incompatible systems were competing for this market: JVC's VHS (Video Home System), Sony's Betamax and (in Europe) Grundig/Philips' Video2000. Although Betamax was believed to be the better format from a technical point of view, VHS won the format-race. The reason for this is network-effects. First, JVC only charged a moderate fee from its license takers (in contrast to Sony). Second, VHS was more popular for distributing porn-movies.[48] Therefore, video stores started to mainly rent out videos in VHS-format. This again made home-users buy VHS-recorders instead of Betamax-recorders.*

- ***HD-DVD vs. Blu Ray*** *– A similar, more recent story can be told for HD-DVDs (capacity up to 30GB, promoted by Matsushita, Pioneer, Philips, Sony, Thomson, LG Electronis, Hitachi, Sharp, and Samsung) and Blu-Ray-DVDs (capacity up to 54GB, promoted by, e.g., NEC, Toshiba, Intel, and Warner). Both formats are incompatible and talks about a common standard already failed in the development-stage.[49] Although, Sony was reluctant to have porn-movies on Blu-Ray disks[50] it successfully negotiated with video strore chains and, thus, heavily promoted its format.[51] However, this case study shows that winning a format-war is not sufficient for winning a market. This is because Blu-Ray also competes with different video-solutions such as the standard DVD[52] or IPTV.*

The example of Blu Ray vs. DVD shall be used to illustrate the **economic analysis of network-effects**. Suppose there is a group A with n_a DVD-consumers and a group B with n_B DVD-consumers (with $n=n_A+n_B$). Now, both groups may decide whether to adopt Blu Ray-DVDs or stay with the standard DVD. When both groups stay with the standard DVD, the groups have a payoffs $u_A(0)$ and $u_B(0)$. When only one group adopts Blu Ray-disks it receives a payoff $u_A(n_A)$ and $u_B(n_B)$ respectively. The payoff of the non-adopting group does not change. If both groups adopt Blu Ray-disks, they receive payoffs $u_A(n)$ and $u_B(n)$. Adoption requires a cost K, e.g., for buying a DVD-player. These payoffs are shown in the following table.

		group B	
		Blu Ray	*DVD*
group A	*Blu Ray*	$u_A(n)$-K – $u_B(n)$-K	$u_A(n_A)$-K – $u_B(0)$
	DVD	$u_A(0)$ – $u_B(n_B)$-K	$u_A(0)$ – $u_B(0)$

48 http://www.spiegel.de/netzwelt/tech/0,1518,336456,00.html

49 http://www.spiegel.de/netzwelt/tech/0,1518,362020,00.html

50 http://www.spiegel.de/netzwelt/tech/0,1518,459652,00.html

51 http://www.spiegel.de/wirtschaft/0,1518,527648,00.html

52 http://www.spiegel.de/netzwelt/tech/0,1518,558763,00.html

Network-effects arise if $u_i(x)$ is increasing in x. Hence, one group's adoption is to the benefit other adopters. This is called the **total effect**. The adoption incentive is $u_i(x)$-K-$u_i(0)$. Thus, given the total effect, one group's adoption raises the other group's incentive to adopt. This is called the **marginal effect** (Farrell and Klemperer 2007: p. 2019). Depending on the values of the payoffs and the adoption incentive, this game can have various equilibria.

1. Suppose the adoption incentive is negative for both groups when only one group adopts Blu Ray ($u_i(n_i)$-K-$u_i(0)<0$) and positive when both groups adopt Blu Ray ($u_i(n)$-K-$u_i(0)>0$). In this case, neither group has an incentive to adopt Blu Ray although adoption by both groups would be welfare-enhancing. This is called **under-adoption** until the prices for Blu Ray-players are lowered such that at least one group has a positive incentive to adopt.

2. Suppose the individual adoption incentive of group A is positive ($u_A(n_A)$-K-$u_A(0)>0$) but negative for group B ($u_B(n_B)$-K-$u_B(0)<0$). In this case, A will adopt Blu ray while group B stays with DVDs.

    a. Observing that A adopts the Blu Ray-technology alters group B's adoption incentive, which may now be positive ($u_B(n)$-K-$u_B(0)>0$). Therefore, B will adopt Bluy Ray in the future, which is a consequence of the marginal effect (see above). This pattern is called **sequential adoption**.

    b. However, if B's adoption incentive is still negative ($u_B(n)$-K-$u_B(0)<0$) the price of players must be lowered for motivating group B to adopt Blu Ray.

3. Suppose the individual adoption incentive of both groups is positive ($u_i(n_i)$-K-$u_i(0)>0$), then both groups will adopt Blu Ray right away.

The competitive evaluation of network-effects is ambiguous. On the one hand, larger networks are to the benefit of consumers and raise welfare. On the other hand, a large network that is operated by a monopolist allows this firm to exercise market power by charging supra-competitive prices. In this context, one should also note that networks-effects constitute switching costs. With regard to the above discussion of switching costs this also suggests intense competition for the market (**market tipping**) but reduced competition within the market. Moreover, a strong incumbent in a network-industry may more easily deter entry into this market or force other firms in the value-chain into exclusivity agreements (see section I).

Essential Facilities

Above, we have seen that the establishment of a (physical) network can be an important element for serving the market. Any input which is deemed necessary for all industry participants to operate in a given industry and which is not easily duplicated might be seen as an **essential facility** (Motta 2004: ch. 2.5.2). An incumbent's control over an essential facility potentially prevents other firms from entering the market and reducing the market power of the incumbent. Moreover, it is often more efficient if two competing firms supply their products via the same network rather than establishing a second network. In this context, think of natural monopolies such as electricity or rail-transport. We distinguish between natural monopolies whose behavior is controlled by regulatory agencies and markets that are under scrutiny by competition authorities. Here, we concentrate on the latter markets.

It is of crucial interest to determine conditions under which a competition authority should force an incumbent to open its essential facility to an entrant. In its Bronner-decision (EU 1998), the European Court established steps for evaluating the request by an entrant to receive access to a dominant incumbent's essential facility.

1. The **relevant product market** must be defined such that it comprises “all the products or services which in view of their characteristics are particularly suited to satisfy constant needs and are only to a limited extent interchangeable with other products or services” (EU 1998: 33; also see section F).
2. The incumbent must be a dominant firm according to Article 102 TFEU (EU 1998: 4, 23; also see section I). Otherwise, a refusal to grant access to its essential facility would not constitute an **abuse of a dominant position**.
3. First, entering the market by means of access to an existing essential facility must be **feasible**, i.e. there are no technical, legal or legal obstacles that prevent the joint usage of the facility by both firms (EU 1998: 44). Second, there must be no alternative ways (**supply-side substitutes**) of entry into this market (EU 1998: 4) such as a duplication of the essential facility. In this context, the opportunities for entry into the market must not be assessed for the actual (potentially weak) entrant but for an efficient firm (EU 1998: 47).

Government Policy

A further reason for the existence of dominant firms are legal rules that prevent competition. One example for such policies is the German medicine-market. German law stipulates that a pharmacy has to be run by a pharmacist, who must not own more than four pharmacies.

On the one hand, in 2006 this law prevented the entry of the Dutch chain DocMorris into the German drug-market although experts and health-insurers believed that the increased competition would have driven drug-prices down without a loss in service-quality.[53] Therefore, legal rules may have prevented competition. DocMorris later decided to issue franchise-contracts for opening pharmacies in Germany, which complies with German laws.

On the other hand, in 2007 DocMorris was acquired by Europes largest drug-trader Celesio. This would have strengthened the competitive position of DocMorris relative to other pharmacies in case of a relaxation of the German laws on pharmacies (see section I on the abuse of a dominant position). In consequence, some competing pharmacies reduced purchases from the Celesio-subsidiary Gehe in order to punish Celesio (see sectionError: Reference source not found on coordinated behavior). As a further consequence of the creation of pharmacy-chains in Germany, these chains would have gained buyer power with regard to drug-producers.[54] In 2009, the European Court of Justice confirmed the lawfulness of the German laws on pharmacies.[55]

53 http://www.spiegel.de/wirtschaft/0,1518,430718,00.html
http://www.spiegel.de/wirtschaft/0,1518,431857,00.html

54 http://www.manager-magazin.de/unternehmen/it/0,2828,479563,00.html
http://www.spiegel.de/spiegel/print/d-52345019.html

55 http://www.spiegel.de/wirtschaft/0,1518,527348,00.html
http://www.manager-magazin.de/unternehmen/artikel/0,2828,625685,00.html

Lessons Learned

After reading this section you should be able to answer the following questions.

1. What is meant by the term *market power*?
2. How can market power be measured?
3. Name and describe two measures of a market's structure. How are these measures used in the European Union for pre-assessing horizontal mergers?
4. Describe the advantages and disadvantages of the above measures of market structure and market power.
5. Explain why free entry not necessarily leads to a socially optimal market structure.
6. Name the assumptions of contestable markets. Explain their meaning for the market outcome in this model.
7. Does buyer power always cause a welfare-loss? Describe, in what way the welfare-effects of buyer power depend on the structure of the upstream and the downstream market.
8. Name and explain some categories of switching costs. Does the existence of switching costs always lower welfare?
9. Describe some problems that might arise when network-effects are present in a market.
10. What is meant by the term *essential facility*?

References

Bresnahan, T.F. and Reiss, P.C. (1991). "Entry and Competition in Concentrated Markets." Journal of Political Economy. Vol. 99 No. 5, pp. 977-1009

Bundeskartellamt (2009). "Sektoruntersuchung Milch – Zwischenbericht Dezember 2009." B2-19/08. http://www.bundeskartellamt.de/wDeutsch/download/pdf/Stellungnahmen/1001_Sektoruntersuchung_Milch_Zwischenbericht_2009.pdf

Bundeskartellamt (2012). "Sektoruntersuchung Milch – Endbericht Januar 2012." B2-19/08. http://www.bundeskartellamt.de/wDeutsch/download/pdf/Stellungnahmen/2012_01_Sektoruntersuchung_Milch_Endbericht_final.pdf

Cabral, L.M. (2000). "Introduction to Industrial Organization." The MIT Press: Cambridge

Destatis (2010). "Statistisches Jahrbuch 2010 für die Bundesrepublik Deutschland mit "Internationalen Übersichten"." http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/SharedContent/Oeffentlich/B3/Publikation/Jahrbuch/StatistischesJahrbuch,property=file.pdf

Dobson, P.W., Clarke, R., Davies, S., Waterson, M. (2002). "Buyer Power and its Impact on Competition in the Food retail Distribution Sector of the European Union." Journal of Industry, Competition and Trade. Vol. 1 No. 3 , pp. 247-281

European Commission (2011). "Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements." Official Journal of the European Union. C11/1. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2011:011:0001:0072:EN:PDF

European Union (1998). "Judgement of the Court (Sixth Chamber) of 26 November 1998." Case C-7/97. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:61997J0007:EN:HTML

European Union (2004.a). "Guidelines on the application of Article 81(3) of the Treaty." Official Journal of the European Union. C 101/97. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52004XC0427%2807%29:EN:NOT

European Union (2004.b). "Guidelines on the assessment of horizontal mergers under the Council Regulation on the control of concentrations between undertakings." Official Journal of the European Union. C 31/5. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?

uri=CELEX:52004XC0205%2802%29:EN:NOT

European Union (2010). "Commission proposes new measures to improve future stability in the dairy sector." Press Release IP/10/1691. http://europa.eu/rapid/pressReleasesAction.do?reference=IP/10/1691&format=HTML&aged=0&language=EN&guiLanguage=fr

Farrell, J. and Klemperer, P. (2007). "Coordination and Lock-In: Competition with Switching Costs and Network Effects." In: Handbook of Industrial Organization, Vol. 3. Armstrong, M. and Porter, R. (ed.). North-Holland: Amsterdam

Hosken, D. and Silvia, L. and Taylor, C. (2011). "Does Concentration Matter? Measurement of Petroleum Merger Price Effects." American Economic Review: Papers & Proceedings. Vol. 101 No. 3, pp. 45-50

Klemperer, P. (1995). "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade." The Review of Economic Studies. Vol. 62 No. 4, pp. 515-539

Kwoka, J. and Shumilkina, E. (2010). "The Price Effect of Eliminating Potential Competition: Evidence from an Airline Merger." The Journal of Industrial Economics. Vol. 58 No. 4, pp. 767-793

Martin, S. (2000). "The Theory of Contestable Markets." http://www.krannert.purdue.edu/faculty/smartin/aie2/contestbk.pdf

Monti, G. (2007). "EC Competition Law." Cambridge University Press: Cambridge

Motta, M. (2004). "Competition Policy – Theory and Practice." Cambridge University Press: Cambridge

Perloff, J.M., Karp, L.S., Golan, A. (2007). "Estimating Market Power and Strategies." Cambridge University Press: Cambridge

Pepall, L. and Richards, D. and Norman, G. (2008). "Industrial Organization – Contemporary Theory and Empirical Applications." 4th edition. Blackwell Publishing: Malden, MA

Vickers, J. (1995). "Concepts of Competition." Oxford Economic Papers. Vol. 47 No. 1, pp. 1-23

F MARKET DELINEATION

The purpose of market delineation is to infer information on the market power of firms from their market shares. Louis Kaplow (2010: 439) defines the stages of market definition as follows: “The first step under the market definition / market share paradigm is to define a so-called relevant market [as is shown in this section F]. This market definition process involves choosing from among candidate markets that which most accurately depicts the extent of market power. The method can usefully be described as one that starts from the homogeneous goods market that includes the firm’s or firms’ product and then considers whether to redefine the market by broadening it to include substitute products (or regions, for geographic market definition) [see section F.2]. Next, one measures the firm’s market share in that market or the collective share of a group of firms, such as when evaluating a horizontal merger, joint venture, or trade association activity. Finally, one infers from this share the degree of market power and assesses it against the pertinent legal standard.” This is done in sections G to I .

F.1 Market Power and the Need to Delineate Markets

Measuring Market Power

In section E , we have defined **market power** from an **economic point of view** as the ability of a firm to set current prices profitably above either marginal costs or the competitive level of prices. A somewhat wider definition provides that a firm has market power when it is able to devise strategies that can harm rivals and so give it, in the future, the power to raise prices and reduce output (Monti 2007: ch. 5). Market power can also be defined from a **legal point of view** in the context of the meaning of dominance. From this point of view, the definition of market power can vary across jurisdictions or even across the articles or paragraphs of the competition laws within a single jurisdiction.

In the European *Michelin*-case, “the Court held that Article [102] 'prohibits any abuse of a position of economic strength by an undertaking which enables it to hinder the maintenance of effective competition on the relevant market by allowing it to behave to an appreciable extent independently of its competitors and customers and ultimately of consumers'. This passage represents the standard test for dominance. The Court's reference to the ability to behave independently 'to an appreciable extent' is relevant for two reasons: first, because it means that Article [102] does not apply to the sort of market power that most firms have due to markets not being perfectly competitive; second, because total control of the market is unnecessary to identify dominance” (Monti 2007: 127).

No matter whether market power is defined more widely or more narrowly, competition authorities would want to **measure market power**. Ideally, the competition authority would be able to determine a firm's **Lerner-index**, which is defined in equation (17) as the difference between a firm's price and its marginal costs relative to its price.

$$L = \frac{p(q) - c(q)}{p(q)} \tag{17}$$

The implicit assumption underlying the measurement of market power using the Lerner-index is that the relevant counterfactual situation is perfect competition where the firms charge prices at the level of marginal costs.

This direct approach for measuring market power "has little practical value as marginal costs are difficult to calculate. Moreover, firms with market power may have high costs (as they face no competition, they may have few incentives to minimise production costs) and their prices will be just slightly above their inefficiently high costs, so the index underestimates their power [see p. 86]. Instead, an indirect method is used to measure market power, based upon a calculation of the firm's market shares and of barriers to entry [see section E.2]. According to this approach, if a firm has very high market shares and entry for new competitors is blocked […], then it holds market power because it is free to raise prices without fearing that its position may be undermined by new entrants. A simple example will illustrate the operation and the potential controversies of measuring market power by this method:

> *assume an ice cream seller in Hyde Park is the sole seller of strawberry flavoured ice creams. Does he have market power? It may be argued that he has a 100 per cent share of the market in strawberry ‚flavoured ice cream in Hyde Park, holding an undisputed dominant position. The ice cream seller may retort that he competes against other ice cream sellers who supply other flavours, so that the relevant market is that for all flavours of ice cream. In this wider market his market share is likely to be much less than 100 per cent. He might go further, and argue that consumers are looking for refreshment, thus chilled drinks would also be substitutes for ice cream, reducing his market share even further. Moreover, he might argue that Hyde Park is surrounded by a number of streets with numerous shops, many selling ice cream of all flavours, thus consumers are free to leave the park momentarily (there being no entry charges) and find cheaper ice cream. All these observations serve the same purpose: widening the definition of the relevant market so as to diminish the defendant's market share. Moreover, he can also argue that there is nothing stopping a new business entering the park and selling strawberry ice cream. These*

observations suggest that entry into the market is easy, so that he has no market power, because a price increase on his part will invite other competitors and bring prices down again [*emphasis added*]" (Monti 2007: ch. 5.3).

The above discussion illustrates three points (Kaplow 2010: 446). First, the direct determinantion of market power by comparing prices and costs does not necessarily require information about the market shares of the firms. However, it is frequently difficult because of problems associated with measuring marginal costs. Second, market definition respectively the calculation of market shares can be helpful in determining market power indirectly. Third, the market definition / market share paradigm is often the best alternative among all the other possible techniques or at a minimum is a helpful complement to assessments of market power.

The Purpose and Meaningfulness of Market Definition

This example demonstrates that the identification of market power is intimately connected with how we define the market. Therefore, defining the relevant market should be considered **an intermediate step in the competitive assessment of a market**. In this context, market definition has two significant benefits. "First, it provides **a focus for the competitive assessment** [*emphasis added*]. By defining the relevant market so as to encompass all those products or services which are considered to be effective substitutes for the products or services at the centre of the investigation, the relevant market focuses the attention of both the Commission and interested affected parties on the main competitive constraints which exist between products and between regions. Excluding such competitive constraints from the analysis will likely result in a flawed competitive assessment.

Secondly, the definition of the relevant market provides **an initial screen for the competitive assessment**. Defining the relevant market enables market shares and market concentration to be calculated. But market shares can provide meaningful information regarding market power for the purpose of the competitive assessment only if the relevant market is defined in a systematic way that captures the competitive constraints [see section E.2] that the firms under investigation face, and hence identifies their effective competitors. In consequence, market shares provide a useful first filter to determine whether a more detailed competitive assessment is required" (Bishop and Walker 2010: p. 109).

Consistent with the above definition of market power the relevant market can be defined as follows: "a relevant market is something worth monopolising. A market is worth monopolising if monopolisation permits prices to be profitably increased. This will be the case if the collection of products contained in this "market" are not subject to significant competitive constraints by products outside the market" (Bishop and Walker 2010: p. 111). This definition is consistent with

the below SSNIP-test for market definition. More generally, one might define market power as follows (Kaplow 2010: 466): First, the best market is taken to be that which most accurately measures market power. Second, the relevant market is that which leads to the right legal answer.

Note that **market shares do not provide unique information about market power**. For example, potential competition (see section E.2) can theoretically eliminate the market power of a monopolist completely while a monopolist in a market with entry barriers can fully exercise its market power (also see Kaplow 2010: 461-62). Therefore, the competitive assessment of markets cannot rely solely on the analysis of market shares but also considers effects related to supply, demand and potential competition as is shown below. Accordingly, Kaplow (2010: 467) asks why one would engage in a market definition exercise in order to undertake further efforts to refining this measure of market power. Therefore, he suggests to abandon the intermediate step of market definition altogether and directly engage in the analysis of market power.

However, Kaplow (2010: 505) also notes that it can be helpful to think "in terms of market metaphors [...]. Yet another, stronger use of market definition is as a quick screen, whether, again, by enforcement agencies allocating scarce resources, or by adjudicators, such as in dismissing cases that almost surely are baseless." Gregory Werden (2012: 1) replies to Louis Kaplow that even "if market shares are not used to infer market power, **market delineation is needed to examine the issue of entry and the durability of market power** [*emphasis added*]. And after market power is established, the relevant market serves an analytic purpose by separating the active forces of competition central to the case from those passive forces in the background." He (*ibid.* p. 8) argues that market delineation facilitates the assessment of firms' market power when direct measures of market power, e.g. the Lerner index, cannot be determined easily.

Werden (2012: 14-15) also argues that the " presentation of an antitrust case is a narrative about actual or likely competitive effects consisting of the actors, the scene, and the action – i.e., competition and the challenged conduct. Decades of experience suggests that market delineation often adds clarity and power to the narrative. The device of a relevant market allows the essence of the narrative to be packed into a single sentence. For example, the essence of the narrative in a merger challenge is that the proposed merger would substantially lessen competition in a particular relevant market, and one sentence can communicate the lead actors, the scene, and the basic action. Such a sentence also facilitates the addition of details, such as the remaining actors, their significance, and who would be injured if the merger were allowed to proceed." In the following, we present methods for the market delineation exercise which in current practice still is an important step in the analysis of market power.

The Relevant Market Notice: Competitive Constraints

Official information is provided by the *Commission Notice on the definition of the relevant market for the purposes of Community competition law* (EC 1997). In §2 it is stated that the "main purpose of market definition is to identify in a systematic way the competitive constraints that the undertakings involved face. The objective of defining a market in both its product and geographic dimension is to identify those actual competitors of the undertakings involved that are capable of constraining those undertakings' behaviour and of preventing them from behaving independently of effective competitive pressure. It is from this perspective that the market definition makes it possible inter alia to calculate market shares that would convey meaningful information regarding market power for the purposes of assessing dominance or for the purposes of applying Article [102]."

According to the *relevant market notice* (§§ 7-9) a "**relevant product market** [*emphasis added*] comprises all those products and/or services which are regarded as interchangeable or substitutable by the consumer, by reason of the products' characteristics, their prices and their intended use [...]. The **relevant geographic market** [*emphasis added*] comprises the area in which the undertakings concerned are involved in the supply and demand of products or services, in which the conditions of competition are sufficiently homogeneous and which can be distinguished from neighbouring areas because the conditions of competition are appreciably different in those area. [...] The relevant market within which to assess a given competition issue is therefore established by the combination of the product and geographic markets."

The *relevant market notice* (§ 13) specifies that firms "are subject to three main sources or competitive constraints: demand substitutability, supply substitutability and potential competition. From an economic point of view, for the definition of the relevant market, **demand substitution** [*emphasis added*] constitutes the most immediate and effective disciplinary force on the suppliers of a given product, in particular in relation to their pricing decisions. A firm or a group of firms cannot have a significant impact on the prevailing conditions of sale, such as prices, if its customers are in a position to switch easily to available substitute products or to suppliers located elsewhere." In order for a single supplier to raise price profitably, the reduction in the level of demand must not be too large. Whether or not this is the case will depend upon how responsive demand is to an increase in price.

In other words, for an increase in prices to be profitable, the **price-elasticity of (residual) demand** must not be too large. In this context, recall that (in a Cournot-model with homogeneous products) the markup on marginal costs, i.e. the Lerner-index L_j, of a firm j with market share m_j is the larger the lower the price-elasticity of demand η.

$$L_j = \frac{m_j}{\eta} \quad (50)$$

It should be considered that this formula only applies to markets with homogeneous products while the relevant market often contains imperfect substitutes as is shown throughout this section. Therefore, the Lerner-index as defined in (50) as the ratio of a firm's market share and the price elasticity of demand need not always provide perfect quantitative inferences on the market power of a firm. However, it is appropriate (in a qualitative sense) to guide one's thoughts in the process of defining the relevant market and assessing market power. Equation (50) illustrates that a price increase by firm *j* is likely to result in a loss of market share as some of the firm's customers switch to firms who supply an identical product. Moreover, some potential buyers would stop buying the product at all at the elevated price. An effect, that cannot be inferred from (50), is that consumers might also be switching to products that are imperfect substitutes to the one sold by firm *j*.

Note that, in response to an increase in prices, it is not necessary for all or even most customers to switch, or for those customers that do switch to switch all of their purchases to render the attempted price increase unprofitable. It is sufficient simply that enough switching takes place so that the attempted increase in price is not profitable. This implies that what matters is not the behavior of "average" consumers, but the behavior of **"marginal" consumers** (i.e. those most likely to switch in response to relative price changes) (Bishop and Walker 2010: p. 119).

Supply substitution "means that suppliers are able to switch production to the relevant products and market them in the short term without incurring significant additional costs or risks in response to small and permanent changes in relative prices" (EC 1997: §20). Two products are supply-side substitutes if the supplier of one of the products already owns all of the important assets needed to produce the other product and has the commercial incentives and capabilities to commence such production. This includes not only the resources to physically produce a product but also marketing and distribution assets. Moreover, for the products of a firm to be regarded as supply-side substitutes it is not only necessary for production of the relevant products to be possible without the need for significant new investments, it must also be possible within a relatively short period of time. This is often taken as a period of up to one year (Bishop and Walker 2010: p. 119-120). "When these conditions are met, the additional production that is put on the market will have a disciplinary effect on the competitive behaviour of the companies involved. Such an impact in terms of effectiveness and immediacy is equivalent to the demand substitution effect" (EC 1997: §20).

One possibility for determining the patterns of demand and/or supply substitution and delineate the relevant market is the Hypothetical Monopolist Test (HMT) or SSNIP-test (Small but Significant Non-transitory Increase in Prices) that is presented in subsection F.2 . There is

considerable debate over whether supply-side substitution should be considered when defining the relevant market, or whether it should be taken into account after the market has been defined. The US approach to market definition is to define markets only on the basis of demand-side substitutability, but then to take account of supply-side substitutability when performing the market power assessment. Conversely, the Relevant Market Notice places primacy on demand-side substitution but states that supply-side substitution will be taken into account when defining the market "when its effects are equivalent to those of demand substitution in terms of effectiveness and immediacy." In one sense, the issue of whether to take supply-side considerations into account at the market definition stage or in the second stage of the competitive assessment does not matter. Provided the competitive constraint posed by supply-side substitutability is taken into account at some point of the competitive analysis, the same conclusions should be reached. However, excluding supply-side considerations at the market definition stage may lead to higher market shares which will overstate the degree of market power possessed by firms (Bishop and Walker 2010: p. 120-121).

"The third source of competitive constraint, **potential competition** [*emphasis added*], is not taken into account when defining markets, since the conditions under which potential competition will actually represent an effective competitive constraint depend on the analysis of specific factors and circumstances related to the conditions of entry. If required, this analysis is only carried out at a subsequent stage, in general once the position of the companies involved in the relevant market has already been ascertained, and when such position gives rise to concerns from a competition point of view."

F.2 Methods for Market Delineation: The SSNIP- or HMT-test

The SSNIP-Test Methodology

"The assessment of demand substitution entails a determination of the range of products which are viewed as substitutes by the consumer. One way of making this determination can be viewed as a speculative experiment, postulating a hypothetical small, lasting change in relative prices and evaluating the likely reactions of customers to that increase" (EC 1997: §15). This **Hypothetical Monopolist Test** (HMT) may be considered the standard approach to defining the relevant market in Europe and USA. The 1992 US-Horizontal Merger Guidelines state that: "A market is defined as a product or group of products and a geographic area in which it is produced or sold such that a hypothetical profit-maximising firm, not subject to price regulation, that was the only present and future producer or seller of those products in that area likely would impose at least 'a small but significant and non-transitory' increase in price [SSNIP], assuming the terms of sale of all other

products are held constant."[56] For defining the relevant product market the **SSNIP-test** is performed as follows:

1. One starts with the type of products that "the undertakings involved sell and the area in which they sell them" (EC 1997: § 16).
2. One engages in a thought-experiment and determines the profit-maximizing price-level of a hypothetical monopolist. In this context, one determines the closest substitutes to which the consumers switch as a result of the price-rise.
3. One determines if the price-level in the hypothetical monopoly is at least 5-10% above a benchmark price-level, assuming that the prices of all other products remained constant.[57] For merger cases, this benchmark price-level is the current price. In case of an alleged abuse of a dominant position the benchmark price-level may also be the competitive level of prices (Bishop and Walker 2010: p. 143).
    a. If *no*: This implies that suppliers of other products also provide important competitive constraints. One adds these closest substitutes to the set of products and returns to step 2.
    b. If *yes*: The current set of products defines a relevant market and competition between suppliers of those products provide the main sources of competitive constraints. The SSNIP-test defines the relevant market as the smallest set of products worth monopolizing.

For defining the relevant geographic market one may proceed as above but iteratively adds regions instead of products (EC 1997: §16). In practice, the product market tends to be defined first and then the extent of the geographic market for those products is defined.

The concept of the Hypothetical Monopolist test is "to identify those products and regions that provide the most important competitive constraints on the firms under investigation. The extent to which firms are able to increase prices above the price level appropriate for the particular inquiry depends on, inter alia, the availability of substitute products (**demand-side substitution** [*emphasis added*]) and the ability of other firms to begin supplying those products (**supply-side substitution** [*emphasis added*]). The fewer good substitute products are available and/or the more difficult it is for other firms to begin to supply those products, the less elastic the demand curve facing the hypothetical monopolist and so the more it can raise prices" (Bishop and Walker 2010: p. 113).

The SSNIP-test is advantageous, as it provides a conceptual framework for defining relevant

56 Cited according to Bishop and Walker (2010: p. 111)

57 This is the US-American version of the SSNIP-test. In Europe, one would ask if it is profitable for a hypothetical monopolist to raise its prices by 5-10% no matter what its profit-maximizing price-level is. This may cause markets to be defined wider than in the US-version of the test (Bishop and Walker 2010: p. 113).

markets. In doing so, it takes into account matters of supply-side and demand-side substitution. “It cannot be stressed enough that defining relevant markets on a basis that is not consistent with the principles of the Hypothetical Monopolist Test will, almost by definition, fail to take properly into account demand-side and supply-side substitution possibilities. In consequence, any market shares calculated from such market definitions will not provide, except purely by chance, a good proxy of market power. Although the Hypothetical Monopolist Test is often proposed as one possible way of defining relevant markets, no alternative that is consistent with the principles of assessing demand-side and supply-side substitutability has been proposed” (Bishop and Walker 2010: p. 123). The SSNIP-test *can* be implemented as a quantitative, econometric test but may also be performed as a analytical assessment.

Fallacies and Further Relevant Issues in Market Definition

When delineating markets one needs to take into account a number of possible fallacies and additional issues that complicate market definition. The probably best-known is the **cellophane fallacy** “after the celebrated *Du Pont* case. In that case, Du Pont argued that cellophane was not a separate relevant market since empirical evidence showed that it competed directly and closely with flexible packaging materials such as aluminium foil, wax paper and polyethylene. But, as many commentators have since noted, Du Pont's argument was not sound. Du Pont was the sole supplier of cellophane” (Bishop and Walker 2010: p. 125). Therefore, it is likely that Du Pont had raised its prices to the profit-maximizing level, so that a further price increase would have been unprofitable. Applying the SSNIP-test on basis of this elevated, non-competitive price level means that one would have answered the question in step 3 of the above scheme with “no” and widened the set of products (as described) by other flexible packaging materials although these do not belong to the relevant market. Hence, if the prices of the products under examination are above competitive price-levels the SSNIP-test provides a relevant market that is incorrectly wide. This leads to firms' market-shares that are incorrectly low, which understates the firms' market power.

How can we **resolve the cellophane fallacy**? First, since the problem arises from starting the market definition exercise from an above-competitive price-level, one might try to determine the competitive price-level and start market definition from there. This approach is advocated by the Relevant Market Notice (§ 19): “Generally, and in particular for the analysis of merger cases, the price to take into account will be the prevailing market price. This may not be the case where the prevailing price has been determined in the absence of sufficient competition. In particular for the investigation of abuses of dominant positions, the fact that the prevailing price might already have been substantially increased will be taken into account.” However, this approach is rather impractical because for determining such counterfactual competitive prices one would need to know

the relevant competitive constraints and model the industry in a competitive situation. However, market definition shall help to identify just these constraints.

Second, in some cases (of e.g. exclusionary behavior) the correct benchmark price level for implementing the test is the current, elevated price level. Here, on would ask whether firms can exclude a rival and then raise their prices above the current level. Third, one might try to circumvent this problem by deriving the market definition, e.g., from a geographically different market in the same products whose price is not elevated. Fourth, "even when we cannot avoid the problem raised by the *Cellophane* fallacy entirely, we can still use market definition, and the Hypothetical Monopolist Test, to structure our thinking in an economically coherent fashion" (Bishop and Walker 2010: p. 129).

These points relate to **Kaplow**'s (2010: 440) **main critique** of the market definition process. He argues that "there does not exist any coherent way to choose a relevant market without first formulating one's best assessment of market power, whereas the entire rationale for the market definition process is to enable an inference about market power." Given that "the best market is that which yields the most accurate inference about market power" (*ibid.* p. 442) it is hard and sometimes impossible to decide whether some market definition A is better than B. This is because the errors caused by defining the market as either A or B can only be assessed relative to the most appropriate market definition, which is unknown. "Without this assessment, there is no meaningful way in which we could say that one measurement error is greater or less than the other" (Kaplow 2010: 471) Hence, Kaplow suggests that choosing a relevant market is a useless exercise.

Werden (2012: 11) does not agree to this fundamental critique of the market delineation exercise. He rather sees difficulties of market delineation in the availability of data that is required to identify demand patterns, i.e. own- and cross-price elasticities. Moreover, the tools that are used to delineate markets "are premised on a particular model of competition and cannot be considered reliable when that model does not fit the industry reasonably well."

Another important fallacy is called **toothless fallacy** and rests on the concept of **captive consumers**. In the *United Brands* decision, "the Commission argued that bananas defined a separate relevant market because the very young and the very old (i.e. those without teeth) did not consider other fruit a suitable substitute for bananas. However, the fact that there are no substitute products available is not enough to define the relevant market" (Bishop and Walker 2010: p. 134). The relevant question is not if there are any captive consumers but if enough **marginal consumers** switch to other fruit in response to a rise in the price of bananas to make that price rise unprofitable. In other words, a set of products is not worth monopolizing if the price-elasticity of demand is sufficiently large. Hence, it is not the behavior of the average consumer but that of the marginal

consumer that matters. “The existence of even a large group of customers who would not switch in response to a relative price increase is not by itself sufficient to conclude that the relevant market should be defined narrowly” (Bishop and Walker 2010: p. 134).

The above example shows that **differences in the physical characteristics of products** do not preclude products from being part of the same relevant market. For example, apples and bananas can well be in the same market although they differ in taste and consistency. What matters are the **substitution effects** between apples and bananas as measured by the **cross-price elasticity of demand**.

Moreover, **differences in end uses** do not preclude products from forming part of the same relevant market. For example, apples sold to consumers and apples sold to fruit pressing plants can well be in the same market although apples and apple juice are not necessarily substitutes from customers' point of view. However, apple-farmers may decide whether they sell their harvest to end consumers or to fruit pressing plants. Hence, there exists supply-side substitutability.

The relevant market notice does not specify a **time** for supply-side and demand-side substitution to occur. However, the period for which the assessment of substitution-effects is done can have a considerable effect on the definition of the relevant market. For example, think of customers who are bound by one-year supply-contracts. Within that year price-changes would not affect the customer-base much. One would assume a low price-elasticity of demand and, thus, infer a narrow market. However, if one looked at a period of more than one year the price-elasticity of demand would be found larger, implying a wider relevant market.

Differences in price levels do not imply separate relevant markets. Differences in prices can be due to (perceived) quality differences. For example, a bottle of Perrier water may physically be the same as some no-name water but is priced higher because of some brand-image. This does not imply that both products are in separate markets because a branded water quenches one's thirst as well as any other mineral water. Therefore, the price of the no-name water can still be a constraint to the price of the branded water. If this constraint is strong enough to render a price-increase of the branded water's supplier unprofitable is an empirical question. Below, we argue that it is not the level but the **correlation of prices** that matters for market definition.

One should note that the relevant geographic market is not bound to **nation-borders**. Particularly in Europe it is easy to cross borders and buy a product at a lower price abroad. For example, drivers may cross borders to buy cheaper (e.g., because of lower taxes) gas abroad.[58] Moreover, the **absence of imports** at current prices does not imply separate relevant geographic markets. Here, what matters is whether imports would occur if prices in one area were raised by 5-

58 http://www.manager-magazin.de/finanzen/artikel/0,2828,474530,00.html

10%. To see this, consider a situation where a good is sold at a competitive price of 100 € in areas A and B. Exporting the good from one to the other area would cost 3 € so that exports are unprofitable. If the good's price in area B was raised to 105 € or 110 €, exports would become profitable. Therefore, one would conclude that area A and B constitute the same relevant market although there is no trade among these regions at current prices (Bishop and Walker 2010: p. 142).

In its §57 the *relevant market notice* states that "[i]n certain cases, the existence of **chains of substitution** [*emphasis added*] might lead to the definition of a relevant market where products or areas at the extreme of the market are not directly substitutable. An example might be provided by the geographic dimension of a product with significant transport costs. In such cases, deliveries from a given plant [A] are limited to a certain area around each plant by the impact of transport costs. In principle, such an area could constitute the relevant geographic market. However, if the distribution of plants [say A, B, and C] is such that there are considerable overlaps between the areas around different plants, it is possible that the pricing of those products will be constrained by a chain substitution effect, and lead to the definition of a broader geographic market." Hence, the relevant market in Figure 20 would not solely comprise area A but also areas B and C.

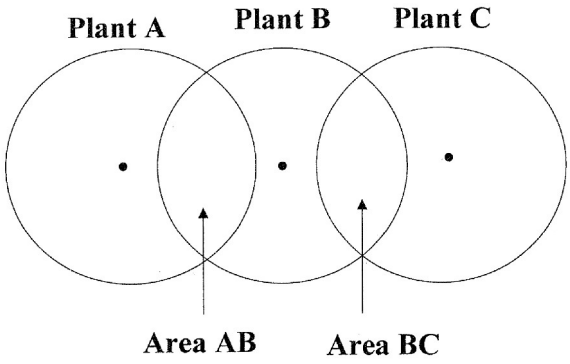


Figure 20: A continuous chain of substitution
Source: Bishop and Walker (2010: p. 146)

> *"Pilkington-Techint/SIV provides an example of this reasoning. This case concerned the acquisition of Società Italiana Vetro Spa, an Italian state-owned glass manufacturer, by a joint venture involving Pilkington Glass. The Commission noted that although glass is a bulky, heavy product, a significant volume of glass is transported across national borders. The Commission concluded that, while the "natural area of supply" for a given producer could be represented by a concentric cycle with the radius determined by relative transport costs, these circles of supply overlapped one another "so that effects can be transmitted from one circle to another". So while a producer in Spain, say, might not compete directly with a*

> *producer in Northern Germany due to transport costs, there might still be a competitive linkage between these two. The Commission therefore defined the relevant geographic market as the whole Community"* (Bishop and Walker 2010: p. 145).

"The same reasoning may apply if product B is a demand substitute for products A and C. Even if products A and C are not direct demand substitutes, they might be found to be in the same relevant product market since their respective pricing might be constrained by substitution to B" (EC 1997: § 57). For example, in the case of cars a "small hatchback does not compete directly with a luxury car, but it may compete directly with a mid-sized car, which may compete directly with an estate, which may compete directly with a luxury car, thus potentially putting all cars into a single market" (Bishop and Walker 2010: p. 146).

"From a practical perspective, the concept of chains of substitution has to be corroborated by actual evidence, for instance related to price interdependence at the extremes of the chains of substitution, in order to lead to an extension of the relevant market in an individual case. Price levels at the extremes of the chains would have to be of the same magnitude as well" (EC 1997: § 58). Hence, the Commission notes that a break in a chain of substitution can hardly be detected by theoretical reasoning. One rather needs to determine empirically what products or regions constrain each other. Some methods for implementing the SSNIP-test empirically are provided in subsection F.3 .

A further relevant question is how competition authorities should deal with **aftermarkets**. These relate to products where the purchase consists of a durable, primary product (for example a printer) and a complementary, secondary product (for example ink cartridges). The buyers of, say, a Canon-printer are to some extent locked into a certain market for ink cartridges (i.e. the aftermarket), as these need to be compatible with its printer. Hence, the suppliers of ink might possibly exercise some kind of market power. It must now be assessed whether the primary and the secondary market constitute a unified market or separate markets.

One relevant aspect in this context is the **frequency of purchases** of the secondary product. On the one hand, business customers print a lot and, thus, need to buy ink cartridges frequently. They would likely consider the entire costs (whole life costing) of printing including the price of the printer and that of the cartridges when buying the printer. Hence, printers and cartridges might be considered being in a unified **system market**. On the other hand, private customers print less often and, thus, need to buy ink less frequently. They concentrate on the price of the printer when buying one and put less emphasis on printing costs. In this case, printers and cartridges would be in separate markets. Again, which of the above effects prevails is an empirical question.

In the latter case of separate markets, one needs to determine if particularly the aftermarket is even more fragmented. It is likely that there is just one market for the primary product comprising, e.g., Canon printers and Epson printers. However, there could be a single or multiple secondary markets. In this context, **physical product characteristics** can play an important role. There would be a single secondary market if Epson cartridges could be used in Canon printers. However, when cartridges are incompatible one would define multiple secondary markets.

Alternatives for the Assessment of Market Power

Kaplow (2010: 475) argues that the market definition exercise is pointless and should be replaced by a direct assessment of market power. Once "a best estimate of market power is formulated – however good or bad such an estimate might be – it should be treated as the market power conclusion rather than used as an input into a market definition process that, as ordinarily conducted, loses information and can produce inferior ultimate outcomes" (*ibid.* p. 480).

For example, one might attempt to directly measure the Lerner-index, which in homogeneous goods markets is directly linked to the price elasticity of demand. In industries with differentiated products this relationship is more difficult. In standard market delineation exercises cross-price elasticities are determined in order to measure substitution effects between products. Kaplow (2010: 485) argues that it is primarily the elasticity of demand that determines a firm's market power. Looking at cross-elasticities would only make sense in order to obtain "a (lower bound) estimate of the market elasticity of demand, from which a direct market power inference would then be made" rather than using cross-price elasticities to define the relevant market.

However, if not in market definition Kaplow (2010: 491) admits that "there is an important use for cross-elasticities when evaluating the prospect of unilateral price increases in horizontal mergers in differentiated product industries. In such settings, the question is not what the level of market power is, either before or after the merger, but how much it will change (this increment, of course, can be calculated by subtracting the before-level from the after-level). Often, the best way to estimate this change (or to predict the level of market power after the merger) is to determine the extent to which the merger relaxes the pricing constraints on the merging firms" (see section G). This requires determining cros-price elasticities and/or the critical loss (see subsection F.3 below).

F.3 Quantitative Techniques for Market Delineation

The SSNIP-test as defined above cannot necessarily be implemented in the form of a quantitative or even an econometric test. However, quantitative evidence often complements qualitative economic reasoning. For example, consider the decision whether caviar and smoked salmon belong to the

same relevant market (Davies and Garcés 2010: ch. 4.1.3). Caviar is potentially a functional substitute for smoked salmon in that it could be served as part of a salad, however it is much more expensive than salmon. The question is whether a sufficient number of consumers would substitute salmon by caviar if the price of salmon was increased, respectively if a sufficient number of consumers would substitute caviar by salmon if the price of caviar was increased. This question cannot be answered by qualitative economic reasoning alone but requires additional quantitative evidence of demand substitutability. This section presents some quantitative tests that are applied in market delineation.[59]

Implementing the SSNIP-Test: Critical Loss and Diversion Ratios

"**Critical loss analysis** [*emphasis added*] does not so much provide an answer to the Hypothetical Monopolist test as rephrase the question in a more user friendly manner. [...] Critical loss analysis asks how many sales a hypothetical monopolist could afford to lose in response to a price rise before the price rise becomes unprofitable. [...] Note that this approach does not answer the Hypothetical Monopolist test. All it does is tell us what is the necessary condition for the Hypothetical Monopolist test to be passed or failed" (Bishop and Walker 2010: p. 552).

In order to answer the SSNIP-test we compare the critical loss (CL) to the **actual loss** (AL) that is likely to occur in response to the price rise. The price rise of the hypothetical monopolist is unprofitable if the actual loss is larger than the critical loss (AL>CL). In this case, the candidate market is still defined too narrowly, and one must add the closest substitutes. Then, one performs the critical loss analysis for the wider market. The relevant market is found when the actual loss is no larger than the critical loss (AL≤CL), i.e. when it is profitable for the hypothetical monopolist to raise prices by 5-10%.

A formula for the critical loss can be found on basis of the idea that profits prior to the price increase (π_0) must not exceed profits after the price increase (π_1; **break-even-condition**).

$$\begin{array}{ccc} \pi_0 & \leq & \pi_1 \\ (p_0-c)\cdot q(p_0) & \leq & (p_0+\Delta p-c)\cdot\left(q(p_0)-\Delta q\right) \\ \Delta q/q(p_0) & \leq & t/(m+t) \\ AL & \leq & CL \end{array} \tag{81}$$

In equation (81), $t=\Delta p/p_0$ denotes the percentage price increase and $m=(p_0-c)/p_0$ the price-cost margin. As stated above, the relevant market is found when the actual loss is no larger than the critical loss, i.e. when inequality (81) applies.

We can also express inequality (81) in terms of a **critical price-elasticity of demand**. To see

59 Much of this subsection is based on chapter 4 in Davies and Garcés (2010).

this, divide both sides by the percentage price increase t.

$$\begin{aligned} \frac{\Delta q/q(p_0)}{\Delta p/p_0} &\leq 1/(m+t) \\ \eta &\leq \eta_{crit} \end{aligned} \tag{82}$$

Condition (82) can be interpreted as follows: A price increase by a percentage t is not unprofitable when the price elasticity of demand η is lower than some critical price elasticity η_{crit}. In this definition of the critical price elasticity one merely assumes a percentage price increase by t. The new price $p_1=p_0+\Delta p$ needs not necessarily be the monopoly price.

The critical price elasticity can also be derived under the assumption that prices are raised by a percentage t such that the new price is the profit-maximizing price of a monopolist, i.e. $p_M=(1+t)p_0$. To see this, consider condition (17) which shows that a profit-maximizing monopolist chooses a profit-maximizing price p_M such that the price cost margin equals the inverse of the price-elasticity of demand η. The question is what value the price elasticity must take such that the profit-maximizing price of the monopolist is a percentage t above the current price p_0.

$$\begin{aligned} \frac{p_M-c}{p_M} &= \frac{1}{\eta} \\ \frac{(1+t)p_0-c}{(1+t)p_0} &= \frac{1}{\eta} \\ \eta_{crit} &= \frac{1+t}{m+t} \end{aligned} \tag{83}$$

This formula for the critical elasticity applies for a constant elasticity demand curve. In case of a linear demand curve, it can be shown that the critical elasticity formula becomes $1/(m+2t)$ (Bishop and Walker 2010: Box 11.1).

If the actual elasticity η is larger than the critical elasticity $(1+t)/(m+t)$, the hypothetical monopolist will raise prices by less than t, because consumers react to an increase of prices by t with a reduction of quantity that is strong enough to render the price increase unprofitable. In this case, the hypothetical monopolist is constrained by competition in substitute products and the candidate market must be widened. One may see that the critical elasticity is the smaller the higher the price-cost margin m, implying a wide market definition. This is reasonable because a small price-cost margin is one sign of intense competition so that all these competitors should be included in the relevant market. In case of linear demand curves the price-elasticity of demand rises when the price of the good rises. If this is the case, assuming a constant elasticity of demand would underestimate the reduction in quantity as a result of the price rise so that the relevant market would be defined too narrowly.

Regarding the use of elasticities, one should note that it is often difficult to measure demand elasticities. However, measuring actual and critical losses needs not necessarily be easier. Moreover, one should determine a **reasonable time period for assessing elasticities**. In some markets, short-run elasticities are larger than long-run elasticities. This is the case when consumers can stock up products, i.e. they buy much when prices are low, for example due to promotional activities. In other markets, long-run elasticities are larger than short-run elasticities. For example, in the short run the owner of a pick-up truck can respond to an increase in the price of gas by driving somewhat less. However in the long-run, she can buy a more fuel-efficient car and reduce her gas-consumption more drastically. As a consequence, the scope of the market may depend on the time period on which the assessment is based.

"The diversion ratio is a concept that is frequently used to measure the closeness of competition between two products in a differentiated products industry" (Bishop and Walker 2010: p. 564). "A **diversion ratio** [*emphasis added*] tries to answer the following question: if the price of good 1 increases, what fraction of lost sales goes to good 2" (Davies and Garcés 2010: p. 191). To answer this question, consider the following system of demand curves for differentiated products.

$$\begin{aligned} q_1(p_1, p_2) &= a_1 - b_{11} p_1 + b_{12} p_2 \\ q_2(p_1, p_2) &= a_2 - b_{22} p_2 + b_{21} p_1 \end{aligned} \tag{84}$$

"The coefficient b_{11} represents the loss of sales of good 1 that will be caused by an increase in p_1 by one unit, say one euro. The coefficient b_{21} represents the increase in sales of good 2 caused by the same price increase. The diversion ratio is then

$$D_{12} = \frac{b_{21}}{b_{11}} = \frac{\partial D_2 / \partial p_1}{-\partial D_1 / \partial p_1} = \frac{\partial D_2 / \partial \ln p_1}{-\partial D_1 / \partial \ln p_1} \quad . \tag{85}$$

The last equality only indicates that the question can be also asked in the terms of the effect of a percentage increase in prices. [...] Estimating the diversion ratio requires knowledge of how consumers would react to a change in the terms of the goods on offer. [...] There are two ways to find out about consumer preferences. One is to observe their choices and try to explain them given the customers' characteristics and the set of possible choices they had available. In such a case, we are using information about consumers' "revealed" preferences. A second method consists of asking consumers about what they would do if they were to face a specific set of choices. In that case, we would be using information based on stated preferences" (Davies and Garcés 2010: ch. 4.4).

In the analysis of **revealed preferences**, one requires data about product characteristics, consumer characteristics, prices, market shares, and consumers' choices. This would allow to estimate a products' demand function and retrieve own- and cross-price elasticities of the products.

To obtain **stated preference** data, one asks the consumers question that allow the researcher

to calculate or at least get an idea of own- or cross-price elasticities as well as patterns of supply-substitution. “Examples of questions would be:

- I notice you have bought brand A. Suppose it cost 50 cents more, would you switch and buy brand B or brand C instead?
- Would you travel to the next big town if tomatos cost 10 cents per kilo less than here? [...]
- How high would the price of yellow paint have to be in order to induce you to switch your red paint machines to start producing yellow paint” (Davies and Garcés 2010: p. 194).

The first two questions directly relate to the assessment that is done in the SSNIP-test to define the relevant product market respectively the relevant geographic market. However, the difficulty with stated preference data is that one cannot be sure whether the interviewees would in reality act as indicated in the survey.

Price Level Differences and Price Correlations

“Examining price differences and correlations is perhaps the most common empirical method used to establish the set of products to be included in a product market. Because correlations require only a small amount of data and are very simple to calculate, they are very commonly presented as empirical evidence in market definition exercises. Correlation analysis rests on the very intuitive assumption that the prices of goods that are substitutes should move together” (Davies and Garcés 2010: p. 169).

The intuition of price correlation analysis is based on **the law of one price**, which states that active sellers of homogeneous goods must sell those goods at the same price. This can be seen in the Bertrand duopoly-model as introduced in section B.2 . Equation (20) implies that the homogeneous product is only supplied by both firms if these firms set identical prices.

$$D_1(p_1, p_2) = \begin{cases} D(p_1), & \text{if } p_1 < p_2 \\ D(p_2)/2, & \text{if } p_1 = p_2 \\ 0, & \text{if } p_1 > p_2 \end{cases} \tag{20}$$

If buying the product requires transaction or transportation costs or the like, the firms will set prices such that the sum of prices and these additional costs is the same whether the consumer buys the product from firm 1 or firm 2. In this model, the law of one price implies that the movement of the two firms' prices is the same. The same applies for the Cournot-model with homogeneous products as shown on p. 48.

“Of course, most goods are not perfect substitutes but may nonetheless be close enough substitutes to ensure that demand schedules and hence prices are closely interrelated. The intuition

from the law of one price is that **similarities in the levels of prices** [*emphasis added*] can indicate that goods are close substitutes. Taking this idea one step further, price correlation analysis is based on the idea that prices of close substitutes will move together. [... I]ntuitively [this] means that we expect prices of substitute goods to move together across time or across regions. Thus, both similarity in the level of prices and also **co-movement of prices** [*emphasis added*] may be helpful when attempting to understand the extent of substitutability between goods" (Davies and Garcés 2010: p. 171). Davies and Garcés (2010: ch. 4.2.2.1) present an example for the analysis of price correlations:

> *"In the Nestle-Perrier merger, a key question became whether the relevant market was the market for still water, the market for water, or the market for nonalcoholic drinks. Price correlations were calculated between brands in the different categories and produced the results shown in [Table 3]. The brands are labeled from A to I. The table reports correlations between prices of goods of individual brands of still water (A-C), sparkling water (D-F), and soft drinks (G-I). From the results, it appears fairly clear that this evidence suggests that the relevant market is the market for water, including both still and sparkling waters but excluding soft drinks. The price correlation between brands of still water and sparkling water is of similar magnitude as the correlation of brands within the group of still waters, at around 0.9. This is clearly a rather high number and is sufficiently close to 1 so as to appear not to leave a great deal of doubt as to its interpretation. In contrast, the positive correlations between the prices of water and soft drinks is low, between 0 and 0.3. That said, the table produces negative price correlations between soft drinks and water, which might suggest that if the price of water rises, the prices for soft drinks decrease and vice versa. This is a rather odd result and it would be interesting to dig a little deeper to understand the causes of such correlation. Although there are a variety of possible causes, one potential explanation is that soft drinks and water are complementary products. The very low correlation within the group of soft drinks is also worth noting. It might be arguable from these data that branded soft drinks present a market of their own.*
>
> *Even with a very high price correlation, other evidence could potentially outweigh the correlation analysis. For example, we might also find survey evidence from consumers suggesting that they are clearly segmented by either having a strong preference for either still or sparkling water. Intuitively, supply substitutability seems likely in tills case but supposing there was evidence from company documents or testimony that the machines for each type of water were impossible to move across to*

produce the other and we also found evidence that company pricing policies were such that they induced a high correlation in prices for some other reason, perhaps simply that the same person currently prices the two goods. The fact that prices are currently correlated may not reassure us that if it were in fact profitable to raise prices for say sparkling water, then prices would indeed be increased."

	A	B	C	D	E	F	G	H	I
A	1								
B	0.93	1							
C	0.91	0.94	1						
D	0.91	0.85	0.86	1					
E	0.94	0.97	0.95	0.92	1				
F	0.93	0.99	0.96	0.88	0.99	1			
G	0.11	0.05	−0.01	0.33	−0.02	0.01	1		
H	−0.57	−0.55	0.25	0.16	0.24	0.27	0.17	1	
I	−0.77	−0.75	−0.81	−0.86	−0.86	−0.79	0.33	−0.11	1

Table 3: Correlations between prices of brands of still water (A-C), sparkling water (D-F), and soft drinks (G-I)
Source: Davies and Garcés (2010: p. 172)

There is a number of **limitations of the price correlation analysis**.

1. There is **no critical level of the correlation coefficient** that would undisputably imply that two firms are in the same relevant market. In the above example, choosing a level of 0.87 would imply that brand D is in the same market as brand A but not with brands B and C. One possibility to define a critical level is to resort to comparators. For example, the lowest correlation among the undisputed substitutes F and D is 0.88, which is of about the same order of magnitude of the correlations among D and B respectively C. Therefore, D could be put in the same market as B and C.

2. Sometimes one finds a positive correlation of prices although the products are not in the same market (**false positives**). This is the case when "two products use the same input and its price varies [...]. [This] will generate a positive correlation in the costs of producing the two products. For instance, both airline travel and rubber are intensive in fuel-based inputs. As the price of oil varies, the costs or producing both airline travel and rubber will covary" (Davies and Garcés 2010: p. 175). As the output-prices vary with marginal costs, they will also be correlated. "A related cause of false positives in a price correlation exercise is the

occurrence of common demand shocks [...]. To see why, consider any two normal goods, say cars and holidays. When the economy is good we will tend to see high demand, and hence high prices, for both cars and holidays and yet, of course, we would not want to define those two goods as being in the same market. [...] Unsurprisingly, in many cases there will be room for substantial debate about the implications of a positive correlation" (Davies and Garcés 2010, p. 176).

3. "**Spurious correlation** [*emphasis added*] occurs when two series appear to be correlated but are in fact only correlated because each of them has a trend. The correlation in this case is a "coincidence" and is not the product of a genuine interrelation between the two products. [...] The basic lesson is that one needs to be very careful when dealing with correlations when variables trend. Seemingly highly robust correlations can be completely spurious and the two variables may be in fact completely unrelated. A formal way to approach this problem is to assess whether a series is "**stationary**" [*emphasis added*]" (Davies and Garcés 2010: p. 177). Alternatively, one may determine whether two time-series are **co-integrated**. "In general, we will want to substantiate claims about stationarity and correlations by checking what happened to the costs of, and demand for, the products during the period of interest. If such shocks exist they may cause a false negative if only one product is affected and substitution is less than perfect. If the shocks are common to both products, they may cause a false positive and the products appear to be more related than they really are" (Davies and Garcés 2010: p. 181).

4. "[T]here are other circumstances when we will underestimate the degree of substitutability [**false negatives**] of two products by just looking at how their prices move together. In particular, if the **signal-to-noise ratio** [*emphasis added*] is low, we will find little correlation between the prices but this result will be driven by random short-lived shocks to the prices of the product and the apparent lack of correlation will not reflect the underlying structural relationship between the products. For instance, suppose the inputs are really different for the two goods and input prices move around a lot. Then the observed correlation in prices will be small due to the variance in the price series caused by shocks to input prices even though the two series may exhibit some limited co-movement. Also, if the data are noisy due to poor quality or measurement problems and the actual prices do not vary much in the period observe, the correlation coefficient will appear small since it will only pick up the noise in the series" (Davies and Garcés 2010: p. 183).

Natural Experiments (Shock Analyses)

"Rather than evaluating the correlation and then checking explanations for its source, shock analysis looks at the reaction of the price(s) of other goods following an exogenous shock on the price of one good, the one at the center of the investigation. Shock analysis is the simplest way of getting a feel for the magnitude of own- and cross-price elasticities of demand without getting involved in a more complex econometric analysis. [...]

To see the **logic of natural experiments** [*emphasis added*], assume a sudden unanticipated exogenous decrease in the price of good A, P^A [...]. Such a change may occur, for example by design, perhaps if a firm conducts a marketing experiment in an attempt to learn about the sensitivity of demand to its price. An exogenous change in the price of good A may feed into (1) the price of good B, (2) the quantity of good B, and (3) the quantity of good A. [...] If the reaction to a decrease in P^A is a sharp decrease of Q^A and a sharp decrease of Q^B, then we can confidently assert that A and B are demand substitutes. [...] A key factor for the success of the methodology is the fact that the original shock on prices is exogenous and not related to the demand of either product A or B, nor related to the cost of inputs for B" (Davies and Garcés 2010: ch. 4.3).

Using Shipment Data for Geographic Market Definition

The Elzinga-Hogarty test for geographic market definition is based on the idea that the geographic market is defined as the area where both imports from and exports to other areas are low. A low level of imports and exports implies that there is no or only insignificant competition among these regions which would impede the exercise of market power. The test consists of two stages, these are known as "little out from inside" (LOFI) and "little in from outside" (LIFO).

The LOFI-measure is defined as

$$LOFI = \frac{Exports}{Production\ in\ candidate\ area} = 1 - \frac{Production\ minus\ exports}{Production\ in\ candidate\ area} \quad . \tag{86}$$

For the LOFI-test to be satisfied the LOFI-measure must take a value below a threshold of 10-25%. In this case, exports are assumed to be sufficiently small so that firms outside the candidate area do not very effectively constrain the competitive behavior of firms inside the candidate area. In order to find the candidate area that satisfies this criterion, one proceeds almost as in the case of the SSNIP-test. One starts with a narrow candidate area and performs the LOFI-test. For LOFI>0.25 one successively adds the areas where most of the candiate area's shipments go. This is done until LOFI<0.25 is satisfied.

The LIFO-measure is defined as

$$LIFO = \frac{Imports}{Consumption\ in\ candidate\ area} = 1 - \frac{Production\ minus\ exports}{Consumption\ in\ candidate\ area} \quad . \qquad (87)$$

Again, for the LIFO-test to be satisfied the LIFO-measure must take a value below a threshold of 10-25%. In this case, imports are assumed to be sufficiently small so that firms outside the candidate area do not very effectively constrain the competitive behavior of firms inside the candidate area. In order to find the candidate area, one starts with a narrow candidate area and performs the LIFO-test. For LIFO>0.25 one successively adds the areas where most of the candiate area's imports come from. This is done until LIFO<0.25 is satisfied. Bishop and Walker (2010: p. 678) provide an example of the Elzinga-Hogarty-test.

> *"In Nestlé/Perrier, the Commission "concluded that the relevant geographic market within which the power of the new entity has to be assessed is France". The Commission found that "Trade flows in the Community are of minor significance". Although the Commission did not formally carry out an Elzinga-Hogarty test, paras 25-27 of the decision provide the [below] statistics.*
>
> *The imports figures imply that the LIFO test is passed for France. However, the export figure implies that the LOFI test is failed. However, [...] the Commission did not take this to mean that the market should be widened. Doubtless the lack of exports from the major contiguous countries to France played an important role in this decision."*

	Imports as percentage of consumption*	**Exports as percentage of production**
France	1–2%	>10%[28]
Italy	< 1%	<1%
Spain	< 1%	<1%
Germany	5%	<1%

* 1990 figures

Table 4: Shipment statistics from the Nestlé/Perrier decision

Source: Bishop and Walker (2010: p. 679)

Lessons Learned

After reading this section you should be able to answer the following questions.

1. Why is it necessary to define a relevant market?
2. What is meant by the concepts of *demand-side substitution*, *supply-side substitution*, and *potential competition*?
3. Describe the basic idea of the SSNIP-test.
4. What problems can arise with regard to market definition when there are primary and secondary markets, i.e. when it is necessary to buy a secondary product in order to use the primary product?
5. How can the SSNIP-test be implemented on basis of the analysis of a critical loss and/or a critical price-elasticity of demand?
6. How can price-correlation analyses be used to define the relevant market?

References

Bishop, S. and Walker, M. (2010). "The Economics of EC Competition Law." 3rd edition. Sweet and Maxwell: London

Davies, P.J. and Garcés, E. (2010). "Quantitative Techniques for Competition and Antitrust Analysis." Princeton University Press: Princeton and Oxford

European Commission (1997). "Commission Notice on the definition of the relevant market for the purposes of Community competition law." Official Journal C 372

Kaplow, L. (2010). "Why (Ever) Define Markets?" Harvard Law Review. Vol. 124, pp. 437-517

Monti, G. (2007). "EC Competition Law." Cambridge University Press: Cambridge

Werden, G. (2012). "Why (Ever) Define Markets? An Answer to Professor Kaplow." U.S. Department of Justice - Antitrust Division. Working Paper (February 13, 2012). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2004655

G MERGER CONTROL

In this section, we provide an overview on the economics of merger control. From the legal side, we concentrate on the European and German legal rules governing merger decisions as outlined in subsection G.1 . “A merger involves the acquisition of one company by another. As a result of that purchase, the acquiring firm gets the physical capital-buildings, equipment, and land – and perhaps certain intangible assets, such as reputation or brand name, that formerly belonged to the acquired company. The ultimate question raised by any merger is, what does the change in ownership permit the merged firm to do that could not be done before?” Therefore, subsection G.2 is concerned with the economic analysis of horizontal mergers. In this context, we provide rationales why competing firms merge and how the effects of such mergers can be assessed. Subsection G.3 is concerned with answering these questions for non-horizontal, i.e. conglomerate or vertical, mergers.

G.1 Overview of Merger Control Regimes

EC Merger Control

The **EC Merger Regulation** first came into force in 1990.[60] The current version[61] (EC council regulation No 139/2004) was passed in 2004. Article 2(1) of the Merger Regulation states that:

> "Concentrations within the scope of this Regulation shall be appraised in accordance with the objectives of this Regulation and the following provisions with a view to establishing whether or not they are compatible with the common market. In making this appraisal, the Commission shall take into account:
>
> > (a) the need to preserve and develop effective competition within the common market in view of, among other things, the structure of all the markets concerned and the actual or potential competition from undertakings located either within or without the Community;
> >
> > (b) the market position of the undertakings concerned and their economic and financial power, the opportunities available to suppliers and users, their access to supplies or markets, any legal or other barriers to entry, supply and demand trends for the intermediate and ultimate consumers, and the development of technical and economic progress provided that it is to

60 This section is based on Bishop and Walker (2010: ch. 7). Quotations that do not constitute original research of the authors are not always stressed.

61 http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32004R0139:EN:NOT

consumers' advantage and does not form an obstacle to competition."

At its introduction, mergers with a Community dimension were assessed with respect to whether the merger created or strengthened a dominant position (see the information on the assessment of mergers in Germany, below). However, in January 2004, the substantive test of the Merger Regulation was changed to whether a merger gives rise to a **significant impediment to effective competition** (hereafter, SIEC). The SIEC test is analogous to the "substantial lessening of competition" test (SLC test) that is used to assess mergers in some other jurisdictions, such as the United States, the United Kingdom, Ireland and Australia. The implications of this change in the substantive merger test were heavily debated at the time and the debate is still on-going.[62] Article 2(3) of the Merger Regulation states that:

> "A concentration which would significantly impede effective competition in the common market or in a substantial part of it, in particular as a result of the creation or strengthening of a dominant position, shall be declared incompatible with the common market."

The Merger Regulation only covers those mergers and joint ventures which have a significant **Community dimension**. A merger (or concentration in the parlance of the legislation) is said to have a Community dimension where:

> (a) the combined aggregate worldwide turnover of all the undertakings concerned is more than EUR 5 000 million; and
>
> (b) the aggregate Community-wide turnover of each of at least two of the undertakings concerned is more than EUR 250 million, unless each of the undertakings concerned achieves more than two-thirds of its aggregate Community-wide turnover within one and the same Member State.

A merger that does not meet the above thresholds is also said to have a Community dimension where:

> (a) the combined aggregate worldwide turnover of all the undertakings concerned is more than EUR 2 500 million;
>
> (b) in each of at least three Member States, the combined aggregate turnover of all the undertakings concerned is more than EUR 100 million;
>
> (c) in each of at least three Member States included for the purpose of point (b), the aggregate turnover of each of at least two of the undertakings concerned is more than EUR 25 million; and

62 See Bishop and Walker (2010: pp. 356-360) for a more detailed overview on this issue.

(d) the aggregate Community-wide turnover of each of at least two of the undertakings concerned is more than EUR 100 million, unless each of the undertakings concerned achieves more than two-thirds of its aggregate Community-wide turnover within one and the same Member State.

Mergers that do not meet these criteria and, thus, do not have a Community dimension are assessed by the competition authorities of one or more Member States.

Subject to the condition of having a Community dimension, all **categories of mergers** fall within the scope of the Commission's jurisdiction, i.e. it covers horizontal mergers (see subsection G.2), vertical mergers and conglomerate mergers (see subsection G.3). The Commission's approach to assessing mergers is set out in two guidelines: Horizontal Merger Guidelines[63] and Non-horizontal Merger Guidelines.[64] **Horizontal mergers** involve companies that operate at the same level of the supply chain, producing substitute goods. Two products are termed substitutes if an increase in the price of one good induces an increase in demand for the other. **Vertical mergers**, in contrast, involve companies that operate at different levels of the supply chain. A common example is a merger between a wholesaler and its retailer, or a manufacturer and its input supplier. **Conglomerate concentrations** involve companies that, although they operate in different markets, produce complementary goods. Examples of conglomerate mergers are Proctor & Gamble/Gillette, which involved different non-overlapping oral products sold to the same retailers, and GE/Amersham which involved medical scanning hardware equipment and diagnostic pharmaceuticals which either enhance or enable images to be produced by such scanning hardware. These different categories of mergers are analyzed in below subsections G.2 and G.3 .

Competition concerns raised by mergers can be categorized as unilateral effects or coordinated effects. **Unilateral effects** (or non-coordinated effects) are said to arise when the merged entity has the ability to increase prices or reduce quality to the detriment of consumers despite the responses of the remaining competitors. As para. 8 of the Horizontal Merger Guidelines makes clear, "increased market power" means the ability of one or more firms to profitably increase prices, reduce output, choice or quality of goods and services, diminish innovation, or otherwise influence parameters of competition. The expression "increased prices" is therefore used as shorthand for these various ways in which a merger may result in competitive harm. In marked contrast, the adverse effects associated with **coordinated effects** depend on one or more competitors to the merged entity choosing to compete less vigorously post-merger, i.e. they collude tacitly. In principle, all categories of mergers can give rise to both concerns, but in practice coordinated effect concerns arise only rarely in non-horizontal mergers. In the following discussion,

63 http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52004XC0205%2802%29:EN:NOT

64 http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52008XC1018%2803%29:EN:NOT

we concentrate on the analysis of unilateral effects with coordinated effects (tacit collusion) being analyzed in sectionError: Reference source not found.

The Merger Regulation also provides for a **fixed timetable** in which decisions are made. Formally, the Commission's assessment of mergers takes the form of a short initial assessment, usually termed **Phase I**. Phase I formally lasts one month from the date of notification. Where the Commission considers that the proposed concentration is likely to give rise to significant competition concerns so that it raises serious doubts as to its compatibility with the common market, a second more detailed **Phase II** investigation is undertaken. Phase II formally lasts for up to four months at the end of which a decision is made.

Since the introduction of the Merger Regulation in 1990, 4,942 mergers have been notified to DG COMP.[65] The majority of these (about 96%) were cleared, i.e. declared compatible with the common market in Phase I or withdrawn, with only 207 mergers (approximately 4 per cent of all notified mergers) progressing to a Phase II inquiry. For a merger subject to a Phase II investigation, there are three possible outcomes. First, the merger can be cleared in its entirety (Article 8(1)). Secondly, the merger can be cleared subject to the parties giving undertakings to remedy the competitive concerns raised by the investigation (Article 8(2)). Up to the end of April 2012, 51 Phase II investigations have been cleared without conditions and a further 95 with commitments. Article 8(3) provides for a third outcome, namely prohibition. The first merger subject to prohibition was Aerospatiale/de Havilland in 1991. Between then and the end of April 2012, in total 22 concentrations were prohibited, although a number of other mergers which faced the prospect of being prohibited following a Phase II investigation have been withdrawn. Of the 4,942 notified mergers 110 have been withdrawn in Phase I and 37 in Phase II.

A number of the Commission's prohibition decisions have been challenged before the **European Court of First Instance**. Prominent examples in which substantive elements of the Commission's decision were challenged include Airtours/First Choice, Tetra LavallSidel, GE/Honeywell and Sony/BMG. The Court of First Instance's decisions can be further appealed to the **European Court of Justice** (ECJ). The decisions of the latter are final. However, it is important to note that both CFI and ECJ are largely concerned with procedural issues, basing their decisions on the same evidence as the Commission's initial assessment while at the same time granting the Commission a wide degree of discretion in the interpretation of that evidence.

65 This was as of end April 2012 (http://ec.europa.eu/competition/mergers/statistics.pdf).

German Merger Control

In March 2012 the German competition authority (*Bundeskartellamt*, BKartA) issued a guidance document on merger control.[66] Currently, the BKartA (2012: 1) assesses whether mergers create or strengthen a dominant position. In the future, the assessment will be based on the SIEC-test (sigificant impediment to effective competition) once the German competition law (*Gesetz gegen Wettbewerbsbeschränkungen*, GWB) has been revised accordingly. The guidance document (BKartA 2012: 2) makes clear that the "purpose of merger control is to protect **competition as an effective process**. Protecting competition at the same time protects the interests of consumers, not necessarily in the short term but rather in the longer term and on a more permanent basis. The general aim of merger control is to protect competition, not to create advantages for individual competitors or to protect them from competition."

"The criterion for prohibiting a merger under German merger control is the question whether it **creates or strengthens a dominant position** (§ 36 (1) GWB). The law defines the term "dominance" as follows: An undertaking is dominant where, as a supplier or buyer on the relevant product and geographic market, it has no competitors, is not exposed to any substantial competition or has a paramount market position in relation to its competitors (§ 19 (2) GWB)" (BKartA 2012: 3).[67] In this context, intervention "by the BKartA is not contingent upon the proof of an impairment of total welfare or consumer welfare [...]. Rather, it is sufficient to prove that the merger actually **threatens the functioning of competition**" (BKartA 2012: 3-4). The Bundeskartellamt (2012: 4) highlights that such an assessment is not done by comparing the situation prior to the merger with the situation after the merger. The competition authority rather compares the future situation with the merger to a **counterfactual** situation without the merger. "In order to determine the creation of dominance, the magnitude of market power of the merging parties has to have increased as a result of the merger so much that their scope of action can be classified as being no longer sufficiently controlled by competition in future" (BKartA 2012: 4).

Negative effects of market power can be offset by welfare enhancing **efficiency effects** that are considered in the context of § 36 (1) GWB[68] and the ministerial authorization (§ 42 GWB). However, the possibility to claim efficiency effects is somewhat limited under the German dominance test. Among others, this is because the assessment of efficiency gains causes significant

66 http://www.bundeskartellamt.de/wEnglisch/download/pdf/Merkblaetter/2012-03-29_Guidance_final_neu.pdf

67 A more detailed description of dominance is provided in section I .

68 § 36(1) GWB: *Ein Zusammenschluss, von dem zu erwarten ist, dass er eine marktbeherrschende Stellung begründet oder verstärkt, ist vom Bundeskartellamt zu untersagen, es sei denn, die beteiligten Unternehmen weisen nach, dass durch den Zusammenschluss auch verbesserungen der Wettbewerbsbedingungen eintreten und dass diese Verbesserungen die Nachteile der Marktbeherrschung überwiegen.*

transaction costs that appear "to be out of proportion to the added value created by a broader recognition of efficiencies, in particular if one takes into account the experience of other competition authorities and the very limited impact efficiencies have had to date on the outcome of their proceedings" (BKartA 2012: 6).

In **horizontal mergers** dominance may occur in the form of single firm dominance or collective dominance on the supply side or buyer power on the demand side. "The term **single-firm dominance** [*emphasis added*] is used to describe a situation in which one company possesses such a high degree of market power that its behaviour is not sufficiently constrained by competition" (BKartA 2012: 8). The German concept of single-firm dominance corresponds to the unilateral effects in European merger control. The analysis of single-firm dominance starts with an analysis of market shares (and concentration measures like the Herfindahl-Hirschman Index) which provide a rough proxy as to whether a merger could be problematic and, thus, requires further investigation (see sections E and F). However, the assessment of market shares can only be a first step and requires an analysis of the actual conditions of competition in accordance with § 19 (2) GWB (for example, entry barriers, potential competition, demand side substitution, and access to suppliers). As further relevant factors within the market, the BKartA takes into consideration:

- capacities and capacity constraints (see section B.2),
- customer preferences (i.e. substitution patterns; see the below discussion of mergers with differentiated products) and switching costs (see section E.2),
- intellectual property rights and know-how,
- market phase (market shares are more indicative of market power in mature markets than in nascent or dynamic markets),
- access to suppliers (e.g., increased buyer power because of the merger, or vertical backwards integration) and customers (e.g., vertical forward integration, establishment of an effective distribution and/or service network, establishment of an important after sales business),
- corporate (for example, cross-shareholdings) and personal links with other companies, and
- financial resources

Other factors to be considered may lie outside the relevant market (BKartA 2012: 22):

- potential competition (for example, suppliers active on neighboring markets or markets up- or downstream of the relevant market) and barriers to entry (see section E and p. 166 below)
- imperfect substitution

"The market power of companies can also be limited by its customers if they have sufficient buyer power. [...] For buyers to be able to prevent dominance, a number of preconditions have to be fulfilled. Firstly, the buyer has to be a sufficiently important customer [...]. Secondly, the buyer is able to switch to other sources of supply, or to sponsor new entry or to vertically integrate" (BKartA 2012: 29-30).

"The concept of **collective dominance** [*emphasis added*] describes a market situation in which few companies within an oligopolistic setting engage in **tacit coordination or collusion** with the effect that they do not effectively compete against each other, although they would be able to do so" (BKartA 2012: 31). Hence, the German concept of collective dominance corresponds to coordinated effects in European merger control. The assessment of collective dominance by the *Bundeskartellamt* is based on two analytical steps: "First, it will be examined whether there is significant competition between and among the group of coordinating companies ("Innenwettbewerb") [...]. Secondly, a precondition for stable coordination is that competitive constraints by outsiders ("Aussenwettbewerb") do not jeopardize the common strategy. [...] Finally, in the overall assessment of all elements relevant in the particular case it will be determined whether the merger would result in the creation or strengthening of collective dominance" (BKartA 2012: 31). "As for single firm dominance, the GWB also contains a **rebuttable presumption** of collective dominance, based on market share thresholds (§ 19 (3) sentence 2 GWB). According to this provision, companies are presumed to be collectively dominant if three or fewer companies reach a combined market share of 50 percent; or if five or fewer companies reach a combined market share of two thirds" (BKartA 2012: 33).[69]

The *Bundeskartellamt* recognizes that **buyer power** can have pro- as well as anticompetitive effects (see section E.2). In the assessment of demand-side dominance it is particularly relevant to assess whether there is one or several customers that are indispensable for suppliers on a particular product and geographic market (BKartA 2012: 46).

The *Bundeskartellamt* also recognizes the ambivalent nature of **vertical mergers**. These "have a less pronounced and more indirect effect on competition because they do not lead to a reduction of the number of actual competitors in the market. Nevertheless, vertical mergers can also lead to competitive restraints and thus create or strengthen single firm dominance [...] or collective dominance [...]" (BKartA 2012: 48). Single-firm dominance can result in input foreclosure, customer foreclosure (see subsection G.3), or the downstream firm may through its upstream-branch gain access to confidential business information on the activities of its competitors which may impede them in their competitive activities (BKartA 2012: 49). The *Bundeskartellamt* (2012:

69 The economics of collusion is described in sectionError: Reference source not found. Issues related to the existence of a (collectively) dominant position are shown in section I .

57) states that **conglomerate mergers** raise competition concerns less often than other types of mergers but can sometimes also result in single firm dominance or collective dominance.

The prohibition of a merger requires an assessment of the causal link between the merger and the creation or strengthening of a dominant position. In order to assess causation the market conditions that are expected after the merger are compared to a counterfactual, i.e. the probable market conditions without the merger (BKartA 2012: 64). Defining the relevant counterfactual situation is difficult because of the uncertainty associated with predicting the future. In particular, this is important in case of the **failing firm defense** where the merging firms claim that one of them would leave or be forced out of the market without the merger. Even in this case it must be established that there is no less anti-competitive alternative than the merger project.

G.2 Horizontal Mergers

"Horizontal mergers produce two consequences that do not arise in either vertical or conglomerate mergers. They reduce the number of firms active on the relevant market and they result in an increase in market concentration" (Bishop and Walker 2010: p. 354). In this context, the two central questions in the analysis of mergers are: Why do firms merge? And, how does a merger affect consumer surplus. We start below to answer these questions on basis of the Cournot-model as introduced in section B.3 .

"In assessing whether a horizontal merger is likely to give rise to anti-competitive effects (either unilaterally or in a coordinated manner), the competitive assessment is normally undertaken in a two-step approach: first, define the relevant market and calculate pre- and post-merger market shares and the change in market concentration; second, assess the likely impact on competition arising from that change in concentration" (Bishop and Walker (2010: p. 356). The first step – defining the relevant market – is shown in section F . Therefore, in the following, we refer to market definition only shortly and concentrate on assessing the impact of a merger on the market.

The Role of the Relevant Market in Merger Analysis

Market definition plays "an important role in the competitive assessment of mergers, not least because the **Commission [places] significant weight on post-merger market shares** [*emphasis added*]. In addition, experience of the assessment of horizontal mergers under the new substantive test reveals that whether a merger is considered to raise serious doubts continues to be driven primarily by a structural assessment. Moreover, from a practical perspective, defining the relevant market continues to provide a **useful filter** [*emphasis added*] for determining whether a more detailed assessment of the merger is required; put simply, mergers involving firms with low market

shares in a properly defined relevant market cannot give rise to competition concerns and therefore do not require a detailed assessment. Finally, it is worth noting that disregarding market definition raises a real danger that **important competitive constraints are ignored** [*emphasis added*] from the competitive assessment entirely" (Bishop and Walker 2010: p. 361).

"The Horizontal Merger Guidelines state that "[m]arket shares and concentration levels provide useful first indications of the market structure and of the competitive importance of both the merging parties and their competitors". The purpose of providing market share thresholds is (or should be) to provide a clear one-tailed test; if the merging parties' market share is not in excess of the threshold then all competition concerns can be readily dismissed without the need for detailed investigation [**safe harbor threshold**]. [...] The Horizontal Merger Guidelines state that mergers where the postmerger entity would have a limited market share are unlikely to impede effective competition. This is particularly the case where the **market share** [*emphasis added*] of the merging parties does not exceed 25 per cent. [...] The Horizontal Merger Guidelines state that mergers that give rise to a combined market share above 50 per cent are normally considered to be problematic absent evidence to show that actual or potential competitors are able to expand their sales sufficiently. The Horizontal Merger Guidelines [in paras. 17-20] also contain thresholds based on **HHI levels** [*emphasis added*] and the change in HHI where the Commission is unlikely to intervene. These are as follows.

- HHI below 1,000;
- HHI lies between 1,000 and 2,000 and the delta (i.e. the change in the pre-merger and post-merger HHIs) is below 250; and
- HHI lies above 2,000 and the delta is below 150.

These thresholds are set at levels such that many horizontal mergers are likely to exceed them. [...] It is therefore important to recognise, as the Horizontal Merger Guidelines clearly state, that these thresholds only provide an initial indicator of the likelihood of competition concerns and do not give rise to a presumption of either the existence or the absence of competition concerns" (Bishop and Walker 2010: pp. 362-364).

Horizontal Mergers with Homogeneous Products

"A horizontal merger is said to give rise to unilateral effects if the merged firm finds it profitable to increase price regardless of the reactions of the remaining competitors" (Bishop and Walker 2010: p. 366). In the following, we illustrate this effect based on a standard Cournot-model.

"In industries where the product supplied by the firms is homogeneous, firms compete in

quantities with the aim of maximizing profits, and customers do not differentiate between suppliers, competition can be modeled as a Cournot game" (Davies and Garcés 2010: p. 386). Therefore, suppose n symmetric firms produce a homogeneous good at constant marginal cost c. Each firm incurs a fixed cost f. Demand for this good is given by

$$Q = a - p \quad . \tag{42}$$

It is easy to show (see section C.2) that each firm sells an optimal quantity

$$q = \frac{a-c}{n+1} \tag{88}$$

at price

$$p = c + \frac{a-c}{n+1} \tag{89}$$

making an optimal profit in Cournot-competition of

$$\pi = \left(\frac{a-c}{n+1}\right)^2 - f \quad . \tag{90}$$

The consumer surplus in this market is given by

$$CS = 0.5 \cdot \left(\frac{n \cdot (a-c)}{n+1}\right)^2 \quad . \tag{91}$$

Now, suppose that m firms in this market merge. The new number of firms in this market becomes

$$n' = n - m + 1 \quad . \tag{92}$$

Substituting n' for n in equations (89) and (91) indicates that **the merger reduces competition** in the market, thus, raises the market price and reduces consumer surplus. Further, assume for a moment that fixed costs are zero. Then, the post-merger profits of both the merged and the non-merged firms amount to

$$\pi' = \left(\frac{a-c}{n-m+2}\right)^2 \quad . \tag{93}$$

Thus, the profits of the non-merging firms increase because of the reduction of competition, which causes prices to rise and output to decline. The reduction of output has two effects. First, lower output means lower total variable costs of production. Second, firms loose the profit-margin that they had made with each unit of the lost output prior to the merger. The firms choose a price that balances these effects and maximizes profits.

It can be seen that the **merging firms provide a positive externality to the non-merging**

firms. A crucial assumption of this model is that after the merger all firms remain symmetric, i.e. the merged firm looks exactly as any of the non-merged firms. Note that a merger of the participating firms is only profitable if the profit of the merged firm is at least as large as their joint pre-merger profits.

$$m \cdot \pi < \pi' \tag{94}$$

By defining $m=\phi n$ condition (94) can be rewritten for the above Cournot-model as

$$\phi > \frac{3+2n-\sqrt{5+4n}}{2n} \quad . \tag{95}$$

Hence, it can be shown for an industry with n=5 firms that for a merger to be profitable at least ϕ=80% of firms must merge. For industries with more firms an even higher percentage of firms must merge. This is called the **merger paradox**, as in such a world we would see very few mergers which is in opposition to the factual evidence. In the following, we show how the merger paradox can be resolved.

One possible solution to the merger paradox are **fixed cost savings**. To see this assume f>0 again. Without synergies the merged firm would incur fixed costs of mf. However, in case of the existence of **synergies** fixed costs can be reduced to ωmf. "What this means is that the merger allows the merged firms to economize on overhead costs, for example, by combining the headquarters of the [...] firms, eliminating unnecessary overlap, combining R&D functions and economizing on duplicating marketing efforts. These are, in fact, typical cost savings that most firms state they expect to result from a merger" (Pepall et al. 2008: p. 393). To see that in this case a merger of less than 80% of firms can be profitable, consider the following example: a=110, c=10, n=9, f=50, and m=6 i.e. ϕ=66.7%. In this case, a firm makes profits π=50 pre-merger. A merging firm's post-merger profits are π'=400-$\omega m \cdot$50. Using condition (94) it can be shown that this merger is profitable when the fixed costs of the merged entity can be reduced to a level of at maximum one third of the combined pre-merger fixed costs (ω<1/3).

Note that these fixed cost savings do neither affect prices (see equation (89)) nor consumer surplus (see equation (91)). Prices rise and consumer surplus declines because of the reduction of competition. Therefore, a competition authority should not pay attention to merging firms' claim that the merger will reduce fixed costs so that prices can be kept constant or even be lowered (**efficiency defense**).

A further example where on the one hand a merger of less than 80% of firms are profitable and on the other hand an efficiency defense might be valid is the case of **variable cost savings**. To see this, consider the Cournot-model with asymmetric firms from p. 48. Let the industry prior to the

merger be characterized by the following parameters a=110, c_l=8, c_h=55, n=3, and f=0. In this situation, a single firm produces an optimal quantity q_l=49 at marginal costs c_l=10, making profits π_l=2,401. Two firms produce an optimal quantity q_h=2 at marginal costs c_h=55, making profits π_h=4. The pre-merger price is p=57 with consumer surplus being CS=1,404.5. Welfare is at a level of W=3,813.5.

Now, the low-cost firm merges with one of the high-cost firms where production can be re-organized such that the merged firm produces at the lower marginal costs c_l=10. In this post-merger duopoly the efficient firm produces an optimal quantity q_l'=49.67 making profits π_l'=2,466.78. The inefficient firm produces an optimal quantity q_h'=2.67 making profits π_h'=7.11. The merger is profitable because the merged low-cost firm's profits exceed the pre-merger profits of the merging partners. The post-merger price is p'=57.67 with consumer surplus having decreased to CS'=1,369.39. When defining

$$\begin{aligned} p &= [a+c_l+(n-1)\cdot c_h]/(n+1) \\ p' &= [a+c_l+(n-2)\cdot c_h]/n \end{aligned}$$

it can be shown that the post-merger price is only lower than the pre-merger price ($p'<p$) for the economically implausible condition n<0.5. Hence, a merger of the above type raises price and lowers consumer surplus. Welfare rises to a level of W'=3,843.28 because the efficiency-induced increase in profits in stronger than the market power related decrease in consumer surplus. This merger would be prohibited by a competition authority who employs a consumer surplus standard. A competition authority following a welfare standard should allow the merger.

Farrell and Shapiro (1990) show that a merger can reduce prices in a Cournot model when the firms are subject to economies of scale. A merger allows the firms to exploit economies of scale better than any firm could individually. Similarly the production of a higher output might lead to learning effects that reduce variable costs and, thus, raise consumer surplus. For example, assume that the merged firm could produce at marginal costs c_l''=6. It can be shown that this causes the firms to make profits π_l''=2,601 and π_h''=4 with a profit maximizing price p''=57 and consumer surplus CS''=1,404.5. This efficiency effect leaves consumer surplus at its level without the merger and raises producer surplus, i.e. it creates a Pareto improvement. However, Farrell and Shapiro (1990: 114) "find that rather impressive synergies – learning, or economies of scale – are typically necessary for a merger to reduce price." The synergies required to raise welfare are more modest.

Merger Waves and Horizontal Mergers that Create Asymmetry

In addition to the above aspects, the merger might alter the mode of competition in the industry. In particular, the merged firm might act as a **Stackelberg-leader** (Pepall et al. 2008: ch. 16.3). I.e.

firms do not act simultaneously any more as in the above Cournot-model. Instead, the merger gives the merged firm a prominent position in the industry so that this leading firm chooses its output first and the remaining companies follow by choosing an optimal output given the leader's choice. It can be shown that such a change in competition increases the profits of the merged leader and decreases those of the followers. Hence, in this model firms' incentive to merge results from the opportunity to gain a competitive advantage over their competitors creating an asymmetric re-distribution of profits. One also finds that prices will decrease and consumer surplus rises.

However, the decrease in followers' profits may cause some of them to respond to this merger by merging themselves in order to become a leader. Each "additional merger creates two countervailing forces. On the one hand, there are fewer firms in total, which ought to increase profits, but there are also more leaders, which ought to decrease the profits of the leaders. [... It can be shown that] starting from any configuration of leaders and followers, an additional two follower firms always wish to merge" (Pepall et al. 1008: p. 399-400). Hence, this model solves the merger paradox and offers one explanation for **merger waves** that are observed in some industries. Unfortunately, the above finding of the merger lowering prices only applies if the number of leaders is sufficiently small. In their specific example, Pepall et al. (2008: p. 401) show that "a two-firm merger that increases the number of leaders benefits consumers only if the current group of leaders contains fewer than a third of the total number of firms in the industry."

This "model solves the merger paradox and gives rise to a merger wave by assuming an asymmetry between newly merged firms and their remaining unmerged rivals. The former gain membership in the club of industry leaders. However, this is a rather strong assumption. While some mergers may create corporate giants with an ability to commit to large production levels, it is far from obvious that every two-firm merger should have this leadership role regardless of which two firms are joined and irrespective of the number of leaders already present" (Pepall et al. 2008: p. 402). Pepall et al. (2008: p. 402) also present an example for merger waves:

> *"In August of 1998, British Petroleum (BP) announced plans to merge with Amoco another large oil firm although not quite as large as BP. The price tag was $48.2 billion making it, at the time, the biggest industrial merger ever. The new BP-Amoco would control more oil and gas production within North America than any other firm. It would also be the third-largest publicly traded oil firm in the world. (The largest firm of all, Saudi Aramco, is not publicly traded.)*
>
> *Reaction from the rest of the oil industry came swiftly. Within a year, Exxon and Mobil merged in a deal worth $73.7 billion to become the largest publicly traded firm on earth. That was quickly followed by the merger of Phillips Petroleum and*

Conoco. Almost simultaneously, Paris-based Total acquired both PetroFina and Elf to create TotalFinaElf. Soon after, Chevron acquired Texaco for $36 billion. BP then went a step further and acquired Arco for $27 billion. The oil merger wave subsided with the economic decline of 2000-1 but even then, did not die altogether. Chevron acquired Unocal in 2005.

This wave of merger activity concentrated oil and gas refining and marketing into the hands of a noticeably smaller number of firms relative to the situation prior to BP's purchase of Amoco. The BP-Amoco merger was then the catalyst for a major wave of mergers and consolidations. In turn, this suggests that a common motive must be behind all these mergers.

Yet whether this common factor was the naked pursuit of market power or simply the profit-maximizing response of firms to similar problems is difficult to say. The mergers were taken at a time when energy prices were quite low. Oil, for example, was selling at less than $12 per barrel in 1998. Indeed, this low price – and the low energy company profits that went with it – is probably the reason that none of the mergers was seriously challenged by antitrust authorities. However, oil prices and profits have risen dramatically since that time. This could reflect the exploitation of newly enhanced market power. In this connection, a recent study by the Government Accounting Office (GAO) found that increased concentration could account for only few cents of the large rise in wholsesale gasoline prices. The rest appeared to be demand and cost pressures. If this view is correct, lower oil and gas prices are only likely if conservation measures reduce energy demand."

Horizontal Mergers with Differentiated Products

Above, we have assumed quantity-setting firms that supply a homogeneous product. In this case, merging firms run into the merger paradox because quantities are **strategic substitutes**, i.e. when the merging firms reduce output to raise prices the non-merging firms raise output to steel some of the merged firm's business. To see this, recall that best-response functions in Cournot-competition are downward sloping as shown in Figure 12. In the following, we show that the merger-paradox does not occur when firms offer differentiated products and compete in prices. This is because prices are **strategic complements** as can be seen by the upward sloping best-response functions in Figure 9. When a merged firm raises the price of its products its non-merged rivals will respond by also raising their prices, which potentially strengthens the effectiveness of the merger. This can be seen on basis of the following model which is based on Pepall et al. (2008: ch. 16.4).

Suppose that there are three firms in the market, each producing a single differentiated product. Inverse demand for each of the three products is assumed to be given by:

$$\begin{aligned} p_1 &= a - bq_1 - \mu(q_2+q_3) \\ p_2 &= a - bq_2 - \mu(q_1+q_3) \qquad \text{with } \mu \in [0,b] \\ p_3 &= a - bq_3 - \mu(q_1+q_2) \end{aligned} \tag{96}$$

In these inverse demand functions the parameter μ, with $0 \leq \mu \leq b$, measures how similar the three products are to each other. If $\mu=0$ the products are totally differentiated and each product constitutes its own relevant market. Each firm is effectively a monopolist in its relevant market. By contrast, as μ approaches b the three products become increasingly identical, moving us closer to the homogeneous product case. If $\mu=b$ the products are homogeneous and demand takes the well-known form

$$p = a - bq \qquad \text{with } q = q_1 + q_2 + q_3 \quad . \tag{97}$$

Assume that the three firms are Bertrand-competitors, i.e., they compete in prices and set prices simultaneously. The firms are symmetric in marginal costs c that are constant in output. Profit of firm 1 is

$$\begin{aligned} \pi_1 &= (p_1 - c) \cdot q_1 \\ &= (p_1 - c) \cdot \left[\frac{a(b-\mu) - (b+\mu)p_1 + \mu(p_2+p_3)}{(b-\mu)(b+2\mu)} \right] \quad . \end{aligned} \tag{98}$$

There are similar profit-functions for firms 2 and 3. Differentiating (98) with respect to p_1 gives the first-order condition for firm 1.

$$\frac{\partial \pi_1}{\partial p_1} = \frac{a(b-\mu) - 2(b+\mu)p_1 + \mu(p_2+p_3) + c(b+\mu)}{(b-\mu)(b+2\mu)} = 0 \tag{99}$$

Given the symmetry of the firms, in equilibrium all firms charge the same price $p_1^*=p_2^*=p_3^*=p_{nm}^*$, which is given by

$$p_{nm}^* = \frac{a(b-\mu) + c(b+\mu)}{2b} \quad . \tag{100}$$

The corresponding output for each firm is

$$q_{nm}^* = \frac{(a-c)(b+\mu)}{2b(b+2\mu)} \tag{101}$$

with profits being

$$\pi_{nm}^* = \frac{(a-c)^2(b-\mu)(b+\mu)}{4b^2(b+2\mu)} \quad . \tag{102}$$

Now suppose that firms 1 and 2 merge but that the merged and non-merged firms continue to set their prices simultaneously. The two previously independent, single-product firms are now product divisions of a two-product merged firm, coordinating their prices p_1 and p_2 to maximize the joint profit of the two divisions ($\pi_1+\pi_2$) while the non-merged firm chooses p_3 to maximize its profit π_3. This gives the following first-order conditions.

$$\begin{aligned}
\frac{\partial(\pi_1+\pi_2)}{\partial p_1} &= \frac{a(b-\mu)-2(b+\mu)p_1+2\mu p_2+\mu p_3+cb}{(b-\mu)(b+2\mu)} &= 0 \\
\frac{\partial(\pi_1+\pi_2)}{\partial p_2} &= \frac{a(b-\mu)-2(b+\mu)p_2+2\mu p_1+\mu p_3+cb}{(b-\mu)(b+2\mu)} &= 0 \\
\frac{\partial(\pi_3)}{\partial p_3} &= \frac{a(b-\mu)-2(b+\mu)p_3+\mu(p_1+p_2)+c(b+\mu)}{(b-\mu)(b+2\mu)} &= 0
\end{aligned} \tag{103}$$

Solving these for the equilibrium prices gives

$$p_{1,m}=p_{2,m}=\frac{a(2b+3\mu)(b-\mu)+c(2b+\mu)(b+\mu)}{2(2b^2+2b\mu-\mu^2)} \tag{104}$$

for the merged firm and

$$p_{3,m}=\frac{a(b+\mu)(b-\mu)+cb(b+2\mu)}{2b^2+2b\mu-\mu^2} \tag{105}$$

for the non-merged firm. Substituting these prices into the profit equations (98) gives the following profits.

$$\begin{aligned}
\pi_{1,m}=\pi_{2,m} &= \frac{b(a-c)^2(b-\mu)(2b+3\mu)^2}{4(b+2\mu)(2b^2+2b\mu-\mu^2)^2} \\
\pi_{3,m} &= \frac{(a-c)^2(b-\mu)(b+\mu)^3}{(b+2\mu)(2b^2+2b\mu-\mu^2)^2}
\end{aligned} \tag{106}$$

It can be shown analytically[70] that this model has the following properties. As a consequence of the merger (i) the merging firms raise the prices of their goods, (ii) sell a lower quantity than in competition, and (iii) make higher profits. The increase in prices is possible because an increase in the price of, say, good 1 diverts consumers to goods 2 and 3. Unlike in competition, where this diversion effect lowers the profits of the single-product firm 1, the merged firm looses only the profits of those customers who switch to good 3. Those who switch to good 2 still contribute to the profit of the merged entity that controls brand 1 and 2. As a consequence, the merged firm also raises the price of brand 2. (iv) This creates an incentive for firm 3 to also raise its price. However, this increase is not as strong as that of brand 1 and 2's prices. This enables firm 3 to steel some business from the merged firm so that (v) its profits rise because of increased prices and an

70 More easily these features can be shown on basis of a numerical example. E.g., set b=1, μ=0.5, a=110,000, and c=22,000.

increased quantity sold. However, the merged firm's profits rise because of the raised prices and the decrease in variable production costs that comes from restricting its quantity.

This simple framework of price setting in a product differentiated market avoids the merger paradox, suggesting that mergers are both profitable and of potential concern to antitrust authorities unless accompanied by cost efficiencies.

Different Aspects in Analyzing Unilateral Effects

"The standard approach to determining whether a horizontal merger is likely to give rise to unilateral effects involves the assessment of **post-merger market shares** [*emphasis added*]. [...] *Ceteris paribus*, the higher the merged firm's market share, the greater is the benefit on the installed customer base of a price increase or an output restriction. Furthermore, market shares may provide a useful proxy for the strength of competitive constraints provided by each firm in the relevant market, i.e. the larger the market, share of a merging party, the greater its pre-merger competitive constraint is assumed to be. [...]

However, it is also important to consider the increment arising from a merger. A merger that gives rise to a post -merger market share of 60 per cent arising from combining a firm with a 57 per cent market share with one with a 3 per cent market share is likely to have a substantially different effect on post-merger competitive conditions than a merger between two firms both with 30 per cent market share; the latter being more likely to give rise to unilateral effects concerns. [...] However, regardless of its level, the market share threshold is merely a filter and it ought not to be presumed that mergers exceeding the threshold necessarily give rise to competition concerns" (Bishop and Walker 2010: pp. 370-372).

The **Horizontal Merger Guidelines** specify in its paras. 17 and 18 that market shares at or above 50% imply the presumption of a dominant position. Effective competition is presumed not to be impeded if the firms' combined market share does not exceed 25% (safe harbor). When the combined market share lies within these bounds the effects of the merger must be assessed more closely taking into account "the strength and number of competitors the presence of capacity constraints or the extent to which the products of the merging parties are close substitutes" (EU Horizontal Merger Guidelines 2004: para. 17).

In markets with **differentiated products**, interpreting market shares "is rendered more difficult since the very essence of competition between differentiated products implies, that consumers do not consider all products to be equally substitutable. In consequence, products do not all impose the same strength of competitive constraint on each other. Where this is the case, market shares, provide a poor proxy for discriminating between "close" competitors and "not so close"

competitors. A competitor will be said to be, loosely speaking, "close" if following a relative price increase a significant proportion of the resulting lost sales would be gained by that competitor. The concept of "**closeness of competition**" [*emphasis added*] when a merger concerns highly differentiated products can be thought of in terms of product characteristics and geographical location" (Bishop and Walker 2010: pp. 372-373). In more technical terms, the closeness of two products can be measured in terms of **cross-price elasticities** and **diversion ratios** as presented in section F.3 and requires an quantitative/econometric analysis of demand.

Excess capacity is a further element that needs to be considered in the analysis of a merger's effects. To see this, consider a firm with a small market share, which would be considered as imposing no relevant competitive constraint on its rivals when regarding market shares alone. However, if this firm can easily expand its output post-merger it may nonetheless constrain its competitors' pricing behavior.[71] Pricing in **auctions and bidding markets** follows different rules than in posted-price markets and, thus, requires a different assessment. Usually it is the second-placed bidder that constrains the bidding-behavior of the winner of an auction. Therefore, anti-competitive effects are most prevalent when the first- and the second-placed bidders merge. Finally, when assessing mergers one should always take into account **dynamic effects**. Asking what comes next, e.g., includes analyzing whether a merger induces entry into the industry or whether it is likely to drive firms out of the industry. In this context, it is also important to assess what would have happened absent the merger. If, for example, one of the merging partners would have been driven out of the industry anyway, then the merger is not very different from the market structure that would have arisen naturally.

This point is called a **failing firm defense**. The Horizontal Merger Guidelines (para. 90) consider "the following **three criteria** [*emphasis added*] to be especially relevant for the application of a 'failing firm defence'. First, the allegedly failing firm would in the near future be forced out of the market because of financial difficulties if not taken over by another undertaking. Second, there is no less anti-competitive alternative purchase than the notified merger. Third, in the absence of a merger, the assets of the failing firm would inevitably exit the market."

> *"These criteria are appropriately stringent and the failing firm defence is unlikely to prevail in most cases. However, in Kali-Salz/MdK/Treuhand the Commission did accept a failing firm defence to clear the merger between the only two German producers of potash. The Commission argued that absent the merger, Mitteldeutsche Kali AG (MdK) would in the near future exit the market, that Kali-Salz would take over the share of MdK and that there did not exist an alternative purchaser who would give rise to less anticompetitive concerns. Under these*

71 This reasoning is somewhat similar to the theory of contestable markets (section E).

circumstances, the merger could not be said to lead to the creation or strengthening of a dominant position since the same market position would arise in any event" (Bishop and Walker 2010: p. 421).

Analyzing Unilateral Effects by Means of Merger Simulation

The above models put a focus on the **unilateral effects** of mergers, i.e. the effects on prices, quantities, and welfare that arise from the profit-maximization rationale of each individual firm. These unilateral effects stand in contrast to the **coordinated effects** of a merger, i.e. the merger may contribute to making collusion in an industry more likely by, e.g., reducing the number of firms or making them more symmetric which facilitates collusive conduct.

Merger simulation models provide one attempt to quantify the unilateral effects of a merger. Merger simulations "are used for two purposes. First, they can serve as a screening device. In that case a standard model is usually taken as an admittedly very rough approximation to the world with the expectation that the merger simulated with that model provides at least as good a screen as the use of market shares or concentration indices alone and hence is a complementary assessment tool to these simple methods. The second purpose of merger simulation involves building a more substantial model with the explicit aim of providing a realistic basis for a "best guess" prediction of the likely effects of a merger" (Davies and Garcés 2010: p. 382).

When performing a merger simulation, one should follow certain **best practices** related to the choice of assumptions, to the data used, and to the framing of the results within a broader analysis. "[i] Practitioners need to justify their choice of modelling assumptions. [... O]ne should be able to argue why the theoretical assumptions are a reasonable approximation of the facts of the case. [...] The best model may well not be a "standard" one. [... ii] If a merger simulation model is built, then the investigator will have to show that it predicts the facts of the industry reasonably well. In particular, predicted prices, costs, and margin behavior must be consistent with the reality of the industry. [...] Methods to check the validity of simulation models include both the use of "in-sample" and "out-of-sample" predictions. [... iii] When the theoretical framework is chosen, parameters need to be estimated or calibrated. If there are sufficient market data available, econometric estimation may be possible and good practice for econometric and regression analysis applies. If there are insufficient data or indeed insufficient time available for estimation and the model is being used solely as a rough-and-ready screening device, then underlying parameters may be calibrated using the predicted structural relationships between observed variables. [... iv] Finally, one should keep in mind that most merger simulations currently involve static models and do not incorporate dynamic effects. Firms may respond to a merger by issuing new products, repositioning

their current products, or by innovating [...]. Each of these reactions will not be captured by a merger simulation. If there is lot of evidence that the market in question has behaved in the past in a very dynamic fashion and that the competitive environment is subject to constant change, the merger simulation exercise will certainly lose relevance for the medium-term prediction of industry outcomes" (Davies and Garcés 2010: ch. 8.1).

To learn about the method of simulating mergers, consider that the industry of interest meets the assumptions of a Cournot-model as analyzed above. "In any game theoretic context including Cournot, economists characterize firm behavior by their best response functions. Consequently, simulating the effect of a merger involves calculating the best response functions for both the pre-merger and post-merger scenarios and solving for the corresponding equilibrium prices and quantities. In the Cournot model, if firms are symmetric in costs, the only difference between the pre- and post-merger scenarios will be the total number of firms operating in the market and so this is the variable that need to be adjusted in the reaction functions. [...] If firms are heterogeneous, we will, in general, need to solve for equilibrium quantities by solving all N reaction functions. [...] Whether firms are assumed symmetric or not, we will need an estimate of marginal cost(s) as well as parameters of the market demand. Once these parameters are estimated, we can compute the pre-merger quantities and profits using the reaction functions of a market corresponding to the number of firms existing in the pre-merger world. We then compute the post-merger quantities and profits" (Davies and Garcés 2010: p. 387). The possible effects of a merger are then evaluated by a comparison of the pre- and the post-merger situation. Please recall that mergers in Cournot-industries are subject to the merger-paradox. Therefore, Cournot-models (particularly with the assumption of symmetric firms) might not be the best theoretical description of many industries as one would assume firms to merge only if the merger is profitable.

"All merger simulations require that one writes down a structural model involving the following equations: (1) a demand equation or a system of demand equations (one for each product in the market); (2) a cost function or a marginal cost function for each product; (3) a description of the firms' strategic variables (e.g., prices, advertising, or quantities) and their objectives (e.g., to maximize profits); and (4) a description of the way in which all the firms' competing objectives fit together, usually via an equilibrium assumption. [...] We discuss each of the above mentioned elements in turn.

Demand. To write down a demand function one may want to make well-motivated assumptions about consumers' preferences and build up from that level of detail to firms' or market demand curves. [...] There are several standard demand functions that are commonly used to describe consumer preferences and each of them will have implications for the prediction of the

effect of a merger" (Davies and Garcés: p. 401). The **estimation of demand** crucially drives the quality of the simulation-results. An introduction to this complex issue is provided by Davies and Garcés (2010: ch. 9).

"**Costs**. Cost functions can also be explicitly laid out taking into account the technological characteristics of the production process. Are there diseconomies of scale? Do we have constant marginal costs? For the determination of equilibrium prices only marginal costs will typically be relevant, although that is subject to the very important caveat that pricing on that basis nonetheless allows firms to recover their fixed costs so that their economic profits in such an equilibrium are positive. One option is to estimate marginal cost curves directly from industry cost information if this is possible. However, sometimes, given the pricing equations, the market prices and demand parameters, marginal costs can be inferred. In those cases, the accuracy of the cost estimate will be hugely dependent on both the demand estimates and the model of competition being the "right" model. [...]

Strategic Variables and Firm Objectives. The strategic variables are those variables which firms choose in a way that takes into account decisions being taken by rivals. The key strategic variables can usually be inferred from company documents [...].

The Nature of Competition. The last explicit assumption that is needed for merger simulation is a description of the nature of the competition taking place in the industry. [...] At the end of the day, the simulation model is just a theoretical model to which we give particular values to the parameters via estimation or calibration that we hope are either right or sufficiently close to being right to be helpful" (Davies and Garcés 2010: p. 402-404).

Below we argue that **efficiencies** provide a countervailing effect to the negative price- and quantity-effects of mergers. "The empirical assessment of the impact of the efficiencies on prices and quantities requires an estimate of the marginal costs of the post-merger firm. [... T]he analyst who wishes to quantify efficiency effects using a merger simulation model will need to model the competition and recompute market equilibrium prices in order to calculate the level of pass-on of the cost savings to the final consumer" (Davies and Garcés 2010: p. 397).

The Effects of Entry

The Horizontal Merger Guidelines state (para. 68) that when "entering a market is sufficiently easy, a merger is unlikely to pose any significant anti-competitive risk. Therefore, **entry analysis** [*emphasis added*] constitutes an important element of the overall competitive assessment. For entry to be considered a sufficient competitive constraint on the merging parties, it must be shown to be likely, timely and sufficient to deter or defeat any potential anti-competitive effects of the merger."

The effects of entry into a market are discussed in section E.2 above. Please note that the analysis of entry-effects only constitutes a second step in the analysis of mergers. In a first step, one should analyze whether the proposed merger may give rise to competition concerns in the absence of entry. In a second step, one may want to determine whether these anti-competitive effects are likely to be mitigated by actual or potential entry. As the European and the German Horizontal Merger Guidelines state that in order for entry to be considered a sufficient competitive constraint, entry must be shown to be **likely**, **timely**, and **sufficient**, we analyze each of these aspects in turn.

The Horizontal Merger Guidelines (para. 69) state that for entry to be **likely**, it must be sufficiently profitable taking into account the price effects of injecting additional output into the market and the potential responses of the incumbents. Hence, it is necessary to assess the *incentive* of firms to enter the market and the existence of entry barriers. These can be categorized as legal, structural and strategic barriers to entry. "**Legal barriers to entry** exist if laws or government regulations make access to the market difficult or even impossible" (BKartA 2012: 23). This is the case for state-protected monopolies or the existence of patents or other types of intellectual property rights. "**Structural barriers to entry** result from the characteristics of the market, the production process of the goods concerned or on the special market position of a company" (BKartA 2012: 24). Entry is less likely if it would only be economically viable on a large scale, i.e. if significant economies of scale exist, or if it is associated with large sunk costs of, e.g., research and development, advertising or training.[72] "**Strategic barriers to entry** can exist if the well-established firms are able to raise rivals' costs of entry or to lower their expected profits" (BKartA 2012: 26). For example, entry is likely to be more difficult if the incumbents are able to protect their market shares by offering long-term contracts or giving targeted pre-emptive price reductions to those customers that the entrant is trying to acquire. Patenting products or production processes can also constitute a barrier to entry. "Market entry may also be more difficult due to existing **barriers to exit**, if costs which would result from market exit, are already anticipated before entry" (BKartA 2012: 26). For example, high risk and costs of failed entry may make entry less likely. The costs of failed entry will be higher, the higher is the level of sunk cost associated with entry.

In order to deter or counteract any perceived increase in prices following the merger, entrants must be able quickly to affect prices in the relevant market to a significant degree. "What is required for entry to be timely will depend upon the characteristics and dynamics of the particular market concerned in the individual case. The usual duration of customer contracts may give an indication for this assessment" (BKartA 2012: 27). The European Horizontal Merger Guidelines consider entry to be **timely** if it can affect prices in the relevant market within two years (Bishop and Walker 2010: 387).

72 Other barriers to entry, such as learning effects, economies of scope, or network effects are discussed in section E.2 .

"Entry must be of sufficient scope and magnitude to deter or defeat the anti-competitive effects of the merger. Small-scale entry, for instance into some market 'niche', may not be considered **sufficient** [*emphasis added*]" (Horizontal Merger Guidelines 2004: para. 75).

Merger Efficiencies

"Mergers can be strategic in nature, perhaps bringing together a company great at marketing with another whose skills lie in product design or engineering. Perhaps companies may simply realize that by joining production efforts, they can produce output more efficiently than they could as separate companies. Joint production can create synergies, allow the exploitation of economies of scale, and facilitate the better use of expertise. For each of these reasons, mergers may create production efficiencies and actively reduce costs. When those cost reductions are passed on to consumers, they may offset the negative effect of the loss of a competitor on market prices and output" (Davies and Garcés 2010: p. 395). In modern competition policy, it is widely accepted that efficiencies may have a potential countervailing positive effect in the post-merger world so that the merging parties may resort to an **efficiency defense**.

Above we have shown that **marginal cost savings** and/or the **elimination of inefficient competitors** may reduce the negative welfare effects of mergers or even change them to the opposite. Marginal cost savings may occur due to "[e]fficiencies arising from the **better utilisation of tangible assets** [*emphasis added*] include rationalisation through the reallocation of the production and/or benefiting from economies of scale or economies of scope.[73] Efficiencies arising from the **better utilisation of intangible assets** [*emphasis added*] include the sharing of know-how, management expertise, R&D and innovation, adjusting product line and benefiting from increased purchasing power." The Horizontal Merger Guidelines (para. 78) state that for "the Commission to take account of efficiency claims in its assessment of the merger and be in a position to reach the conclusion that as a consequence of efficiencies, there are no grounds for declaring the merger to be incompatible with the common market, the efficiencies have to **benefit consumers**, be **merger-specific** and be **verifiable** [*emphases added*]. These conditions are cumulative." Each of these conditions is evaluated in turn.

In order to **benefit consumers** (Horizontal Merger Guidelines 2004: paras. 79-84), efficiency gains must either lead to lower prices or bring about other benefits such as new or improved products or services resulting from efficiency gains in the sphere of R&D and innovation (see section C.2). A merger may also be considered efficient when it reduces the likelihood of coordinated effects, for example, by making firms more asymmetric. "In general, the later the efficiencies are expected to materialise in the future, the less weight the Commission can assign to

73 See section C.1 .

them. This implies that, in order to be considered as a counteracting factor, the efficiencies must be timely" (Horizontal Merger Guidelines 2004: para. 83).

"Efficiencies are relevant to the competitive assessment when they are a direct consequence of the notified merger and cannot be achieved to a similar extent by less anticompetitive alternatives. In these circumstances, the efficiencies are deemed to be caused by the merger and thus, **merger-specific** [*emphasis added*]. It is for the merging parties to provide in due time all the relevant information necessary to demonstrate that there are no less anti-competitive, realistic and attainable alternatives of a non-concentrative nature (e.g. a licensing agreement, or a cooperative joint venture) or of a concentrative nature (e.g. a concentrative joint venture, or a differently structured merger) than the notified merger which preserve the claimed efficiencies. The Commission only considers alternatives that are reasonably practical in the business situation faced by the merging parties having regard to established business practices in the industry concerned" (Horizontal Merger Guidelines 2004: para. 85).

"Finally, any claims of efficiencies arising from a merger need to be **verifiable** [*emphasis added*] such that the Commission can be reasonably certain the efficiencies are likely to materialise and be substantial enough to counteract a merger's potential harm to consumers. The burden of proof on demonstrating efficiencies lies with the merging parties" (Bishop and Walker 2010: p. 415).

Please note that cost-savings in marginal costs immediately feed through to prices while fixed cost-savings are irrelevant for firms' pricing behavior as long as marginal costs are above average total costs. Therefore, an efficiency defense should only be considered when the claimed efficiencies are likely to reduce marginal costs. In practice, the question may arise what cost-components constitute fixed costs and what cost components are variable. "The more pragmatic, common sense view would be that reductions in costs that are avoidable, or fluctuate with output in the medium term (say over 12 to 24 months), do affect the price formation process and should therefore be taken into account. [... Note that] in practice efficiency arguments that prevail are extremely rare. Indeed, to date, the Commission has not cleared any horizontal merger based on the efficiencies created by that merger. In practice, therefore claims as to the efficiency benefits of a horizontal merger are more likely to provide a rationale for the merger rather than a substantive argument in itself" (Bishop and Walker 2010: p. 416).

Röller (2011) provides some **evidence** on the application and decisiveness of efficiency effects in the 37 phase II merger decisions that have been published between 2004 and 2009. He concludes that "efficiencies have not played a major role in phase II EU merger evaluations since 2004. Out of the 37 published cases, in only 6 cases did efficiencies play any role at all (16%) in the

sense of being claimed by the parties. In all of these 6 cases, the merging parties claimed (static) efficiencies in fixed costs that were not accepted by the Commission. In 5 cases, they also claimed (dynamic) efficiencies in variable costs that were accepted by the Commission in only 2 cases. In none of these cases the claimed efficiencies affected the final decision of the Commission. Röller (2011: p. 191) lists those cases:

- *"T-Mobile Austria / Telering (05): Mobile phone operators. Dynamic efficiencies claimed, consumer benefit not accepted.*
- *Inco / Falconbridge (06): Companies active in the mining, processing and refining of nickel and other metals. Both static and dynamic efficiencies were claimed, but deemed not to be merger specific or benefit consumers.*
- *Ryanair / Aer Lingus (06): Two leading airlines operating from Ireland. Both static and dynamic efficiencies where claimed, but not deemed verifiable or merger specific. However, the efficiency assessment was indicated not to be decisive: " ... even if these criteria were met, the claimed efficiency gains would in all likelihood be insufficient in magnitude to reverse the anti-competitive effects identified ... " (§1127 of decision text)*
- *TomTom / Tele Atlas (07): Vertical acquisition of a navigable digital map provider by a portable navigation devices producer. The claimed static efficiencies were accepted (elimination of double mark-ups), but the dynamic ones not as they were deemed not verifiable.*
- *Nokia / NAVTEQ (08): Vertical acquisition of a navigable digital map database provider by a mobile telephone producer. Static efficiencies were accepted (elimination of double mark-ups), but dynamic efficiencies were deemed not verifiable or merger specific.*
- *KLM / Martinair (08): Dutch airlines active in the transport of passengers and cargo. Both static and dynamic efficiencies where claimed but not accepted due to a lack of verifiability."*

G.3 Non-Horizontal Mergers

This section is concerned with a description of the pro- and anti-competitive effects of vertical and conglomerate mergers and their assessment in EU competition policy. The main pro-competitive effect of vertical mergers consists in the elimination of double mark-ups. However, vertical mergers may also give rise to foreclosure-concerns. Hence, a vertically integrated firm might harm its rivals

in the downstream market by a refusal to supply important inputs (input foreclosure). The vertically integrated firm might also refuse to purchase inputs from its upstream competitors (customer foreclosure). These aspects are detailed in turn after providing a short introduction to non-horizontal mergers and the European Commission's non-horizontal merger guidelines.

Introduction to the Economics and Legal Aspects of Non-Horizontal Mergers

> *"In the fall of 2000, General Electric (GE) and Honeywell International announced that the two companies would merge with GE acquiring Honeywell. GE is, of course, a very well known firm with annual revenues well over $100 billion. Its businesses are involved in everything from lighting and appliances to television programming (it owns NBC) and financial services. GE is also a major supplier of jet engines for commercial aircraft for which its chief competitors are Rolls Royce and Pratt-Whitney. Honeywell was originally a leader in temperature and environmental controls but has, over time, developed into a major aerospace firm whose products included electric lighting, ventilation units, and braking systems for aircraft and also starter motors for aircraft engines of the type GE builds. The deal was approved in the United States. However, in July 2001, the European Commission following the recommendation of Competition Commissioner, Mario Monti, blocked the merger. [emphasis added]*

The proposed GE-Honeywell merger was a marriage of complementary products. The more aircraft engines GE sells the more starter motors and other related aircraft items Honeywell could sell. A merger of GE and Honeywell is a vertical merger. Often vertical mergers are comprised of firms operating at different levels of the production chain, say, a wholesaler and a retailer. [...] Because an upstream-downstream relationship is just one of the many types of complementary relationships that may exist between firms, the term **vertical merger** [*emphasis added*] has come to have the more general interpretation of a merger between any two firms that produce complementary products" (Pepall et al. 2008: p. 430).

In the EU, non-horizontal mergers are assessed under the **non-horizontal merger guidelines**[74] as published in 2008. "Whether a non-horizontal merger is deemed to have a Community dimension is subject to the same turnover thresholds that apply to horizontal mergers. The same procedures and timetables also apply" (Bishop and Walker 2010: 8-001). In its

74 Guidelines on the assessment of non-horizontal mergers under the Council Regulation on the control of concentrations between undertakings. Official Journal 2008/C 265/07. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52008XC1018%2803%29:EN:NOT

introduction, the non-horizontal merger guidelines specify: "**Vertical mergers** [*emphasis added*] involve companies operating at different levels of the supply chain. For example, when a manufacturer of a certain product (the 'upstream firm') merges with one of its distributors (the 'downstream firm'), this is called a vertical merger. **Conglomerate mergers** [*emphasis added*] are mergers between firms that are in a relationship which is neither horizontal (as competitors in the same relevant market) nor vertical (as suppliers or customers). In practice, the focus of the present guidelines is on mergers between companies that are active in closely related markets (e.g. mergers involving suppliers of complementary products or products that belong to the same product range)."

In its overview, the non-horizontal merger guidelines set out that "[n]on-horizontal mergers are generally less likely to significantly impede effective competition than horizontal mergers. First, unlike horizontal mergers, vertical or conglomerate mergers do not entail the loss of direct competition between the merging firms in the same relevant market. As a result, the main source of anti-competitive effect in horizontal mergers is absent from vertical and conglomerate mergers. Second, vertical and conglomerate mergers provide substantial scope for efficiencies. A characteristic of vertical mergers and certain conglomerate mergers is that the activities and/or the products of the companies involved are complementary to each other. The integration of complementary activities or products within a single firm may produce significant efficiencies and be pro-competitive." These **pro-competitive effects of vertical mergers** mainly refer to the **elimination of double marginalization** as is explained below.

The non-horizontal merger guidelines (paras. 17-19) go on noting that "[t]here are two main ways in which non-horizontal mergers may significantly impede effective competition: **non-coordinated effects** and **coordinated effects** [*emphases added*]. [...] Non-coordinated effects may principally arise when non-horizontal mergers give rise to foreclosure. In this document, the term '**foreclosure**' [*emphasis added*] will be used to describe any instance where actual or potential rivals' access to supplies or markets is hampered or eliminated as a result of the merger, thereby reducing these companies' ability and/or incentive to compete." The **anti-competitive effects of vertical mergers,** i.e. an impediment to effective competition by means of the creation or strengthening of a dominant position (also see section I) in particular by means of foreclosure, are presented below. For a more detailed treatment of coordinated effects, please refer to section Error: Reference source not found.

Foreclosure may be defined as a situation, "where the merged firm reduces the ability or incentive of competitors to compete to such an extent that they are marginalised or driven from the market altogether. [… Foreclosure] arises by **refusing to deal** [*emphasis added*] with rival firms, either upstream or downstream firms, or by only dealing with these rival firms on terms less

favourable compared to the pre-merger situation. By foreclosing rival firms, such practices can reduce competition in either the upstream or downstream markets thereby resulting in an increase in price to the detriment of consumers: In the case of conglomerate mergers, leveraging generally arises from the **bundling or tying of products** [*emphasis added*]" (Bishop and Walker 2010: 8-004-005).

Non-horizontal mergers are unlikely to be of concern for the Commission, "where the **market share** post-merger of the new entity in each of the markets concerned is **below 30%** and the post-merger **HHI is below 2 000** [*emphases added*]. In practice, the Commission will not extensively investigate such mergers, except where special circumstances such as, for instance, one or more of the following factors are present:

(a) a merger involves a company that is likely to expand significantly in the near future, e.g. because of a recent innovation;

(b) there are significant cross-shareholdings or cross-directorships among the market participants;

(c) one of the merging firms is a firm with a high likelihood of disrupting coordinated conduct;

(d) indications of past or ongoing coordination, or facilitating practices, are present" (non-horizontal merger guidelines, paras. 25-26).

Pro-Competitive Effects of Vertical Mergers

"When firms occupy different stages of the production stream the convention is to label those firms farthest from the final consumer of the product as upstream and those closest to that consumer as downstream. Film companies and movie theaters are an example. In this case, the film company is the upstream firm and the theater that shows the film is the downstream firm. Manufacturers and retailers have a similar upstream-downstream relation. All such relationships can be usefully viewed through the lens of complementarity. Each firm in the vertical chain provides an essential service to other firms in the chain. Vertical relationships between two firms-each with monopoly power-leads to a loss of economic efficiency in the absence of some mechanism to coordinate the decisions of the two firms. In the case of vertically related firms, this is referred to as the problem of **double marginalization**" (Pepall et al. 2008: pp. 431-432).

The central idea behind the double marginalization problem is the notion that producers of complementary products individually attempt to achieve the greatest gain from serving a particular market. In this context, the incentives of the complementors are contrary to each other: "If your complement or gets less of the pie, that leaves more for you, and vice versa. This tug-of-war

between complementors is evident in the computer business. Since hardware makers complement Microsoft, Compaq's and Dell's entry into the IBM-compatible personal computer market benefited Microsoft. But Microsoft gains even more every time Compaq or Dell starts a price war. When the price of hardware falls, more people buy computers, which leads to more software sales. Microsoft wins. Even people who would have bought computers at the old, higher price now have more money left over to buy software. Microsoft wins again. Complementors may be your friends, but you don't mind if they suffer a little. Their pain is your gain" (Brandenburger and Nalebuff 1996: 36).

More technically, suppose the following situation. There is a **monopolist supplier** in the upstream market, the manufacturer, who produces a unique product at constant marginal costs *c* and sells this product to a downstream firm, the retailer, at a wholesale price *r*. In the downstream market, a **monopolist retailer** resells the product to consumers at the market clearing price *p*. For the sake of simplicity, we assume that the retailer only incurs purchasing costs *r* but no costs of retailing. Consumer demand is described by

$$Q = a - p \quad . \tag{42}$$

To determine the market-equilibrium, we begin with calculating the downstream firm's best response to the upstream manufacturer choosing some value of the wholesale price *r*. We do this by maximizing the retailer's profit function.

$$\begin{array}{rcccccc} & \pi_D(Q,r) & = & (p-r)\cdot Q & = & (a-Q-r)\cdot Q & \\ max.\,\pi_D \quad \Rightarrow & \dfrac{\partial \pi_D}{\partial Q} & = & a-2\mathrm{Q}-r & \stackrel{!}{=} & 0 & (107) \\ & & & MR_D & = & r & \end{array}$$

We find the expected result that the retailer maximizes its profits at the output where its marginal revenue MR_D equals its marginal costs *r*. Thus, given some wholesale price *r*, the retailer optimally demands quantity Q_D

$$Q_D = (a-r)/2 \tag{108}$$

selling the good at price p_D

$$p_D = (a+r)/2 \quad . \tag{109}$$

Equation (108) may be interpreted as the retailer's demand in the upstream market given the wholesale price *r*. Therefore, we maximize the upstream manufacturer's profit subject to this condition.

$$\begin{array}{rcccc} & \pi_U(Q) & = & (r-c)\cdot Q & = & (a-2\mathrm{Q}-c)\cdot Q \\ max.\,\pi_U \;\Rightarrow & \dfrac{\partial \pi_U}{\partial Q} & = & a-4\mathrm{Q}-c & \stackrel{!}{=} & 0 \\ & & & MR_U & = & c \end{array} \tag{110}$$

This gives the manufacturer's profit-maximizing output and price.

$$\begin{array}{l} Q_U=(a-c)/4 \\ r_U=(a+c)/2 \end{array} \tag{111}$$

These relationships are shown in Figure 21 with the light-gray area being the profit of the upstream manufacturer and the dark-gray area being the profit of the downstream retailer.

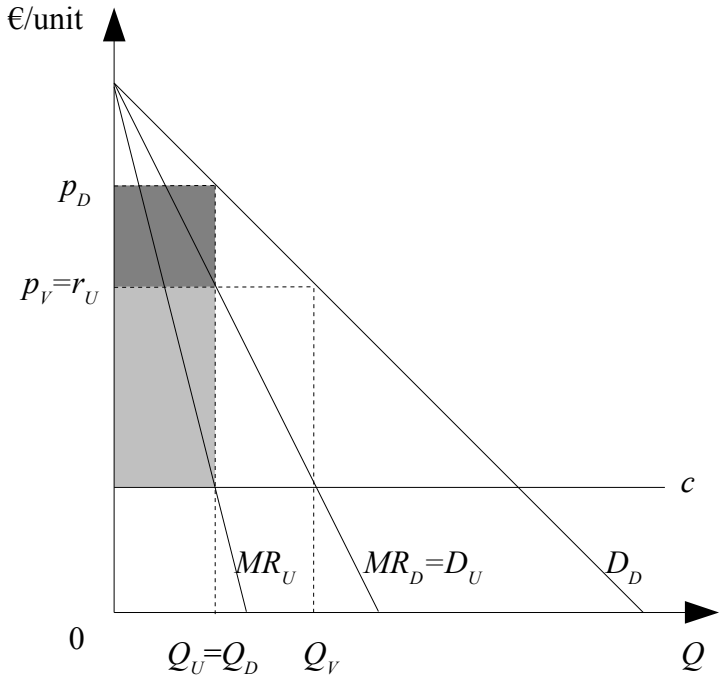


Figure 21: Vertically separated and integrated firms

Now, consider a merger between the upstream-manufacturer and the downstream retailer to a single **vertically integrated monopolist**. Maximizing this monopolist' profit- function yields

$$\begin{array}{rcccc} & \pi_V(Q) & = & (p-c)\cdot Q & = & (a-Q-c)\cdot Q \\ max.\,\pi_V \;\Rightarrow & \dfrac{\partial \pi_V}{\partial Q} & = & a-2\mathrm{Q}-c & \stackrel{!}{=} & 0 \\ & & & MR_V & = & c \end{array} \tag{112}$$

with

$$\begin{array}{l} Q_V=(a-c)/2 \\ p_V=(a+c)/2 \end{array} \quad . \tag{113}$$

The profit of the vertically integrated firm equals the sum of the light-gray and the hatched area in Figure 21. It can easily be checked that the merger of these firms benefits the merging firms via

higher profits *and* the consumers because of lower prices and an increased output.

This result may be attributed to the elimination of **double marginalization** (**increasing pricing efficiency**). The problem of double marginalization can be explained intuitively for, e.g., the computer market. "Each firm's pricing decision imposes an externality on the other firm. A high price for computer hardware reduces demand for PCs. It also reduces demand for programs and operating systems. The hardware manufacturer takes the first effect into account, but not the second. The same is true, of course, in reverse. The software manufacturer does not take into account the impact its price choice has on the demand for hardware. In the noncooperative Nash equilibrium, the prices of both goods are too high. If, say, the hardware firm were to cut its price, this would generate additional demand and additional profit for the software firm. However, since the hardware firm does not receive any of this additional profit, its incentive to reduce price is weakened. This suggests that, with cooperation, both firms would lower their prices and be better off. Consumers, too, would gain as a result of lower prices and expanded output.

One way to achieve the profit and efficiency gains of cooperation is for the two firms to merge. Such a merger creates a single decision-making entity and, therefore, permits the externality to be internalized. The combined hardware and software firm will maximize its total profit by reducing the prices of both complementary goods so as to maximize the joint profit from each. Whenever firms with monopoly power produce complementary products, they have a strong incentive either to merge or to devise some other method to ensure cooperative production and pricing of the complementary goods" (Pepall et al. 2008: p. 430).

"In summary, vertical integration of a chain of producers, each of which has monopoly power, is likely to benefit both firms and consumers by correcting the market failure associated with double (and triple and quadruple ...) marginalization. These benefits are more likely to arise when the technology operated by downstream firms offers limited opportunities for substitution into other inputs" (Pepall et al. 2008: p. 435). Pepall et al. (2008: p. 449) provide an example for the efficiency effects of vertical integration:

Reality Checkpoint

Going Whole Hog on Vertical Integration

Nowhere have the strategic advantages of vertical integration been more aggressively pursued in recent times than in meat and poultry markets. Firms in these markets have made a concerted effort to control all aspects of production from the farm to the store counter or, as they say in the pork business, from "birth to bacon and squeal to meal."

The largest pork producer in both the U.S. and the world is Smithfield Industries. It controls 26 percent of the U.S. market—and control is the right word. Smithfield has either ownership or decision-making rights over every single stage of the production chain up until the product is displayed in a local store. It owns the DNA line for the hogs that it uses and the feed that they eat. It directly owns hundreds of mega hog farms. In addition, thousands of farms work as Smithfield contractors in which case Smithfield still owns the hogs if not the farms themselves.

The advantages of such control are clear. By directing insemination and breeding, Smithfield can maintain a supply of new, and bigger litters that allow it to slaughter more hogs each year without threatening the sustainability of its herds. Feed and genetic control give the firm a tight grip on the leanness and other key features of its pigs. Contracts and ownership also permit Smithfield to design the warehouse barns in which the hogs are raised. This has the added virtue of ensuring that the hogs will mature on schedule and be ready for transport to the pork processing plant in a timely fashion. This is important because the plants are designed to operate efficiently at a particular level of use and even small deviations from this capacity utilization rate can lead to rapidly escalating costs.

Thus, vertical integration has permitted standardization, quality control and, of course, the elimination of double marginalization. As a result, Smithfield has become the world's largest pork producer, annually slaughtering close to 20 million hogs to turn out over 5 billion pounds of pork. Nor is Smithfield the only meat processor that has pursued this strategy. Tyson, which with its poultry, beef, and pork operations is even bigger than Smithfield and is, in fact, the world's largest meat producer, is also highly integrated. In fact, it was Tyson that introduced such large-scale integration to farming when it began to reorganize its poultry business. Other pork firms such as ConAgra and Swift have also organized their operations in this same, vertical manner. It's the swine of the times.

Sources: S. Kilman, "Smithfield to Buy Hog Farmer Premium Standard," *Wall Street Journal*, Tuesday, September 19, 2006, p. A4; and S. Martinez, "A Comparison of Vertical Coordination in the U.S. Poultry, Egg, and Pork Industries," *Current Issues in Economics of Food Markets*, Agriculture Information Bulletin, AIB-747-05 (May, 2002).

As its name suggests, the problem of double marginalization only arises when the firms in the upstream *and* in the downstream market earn a positive margin. If competition in either the upstream or the downstream market drives economic profits down to zero the problem of double marginalization does not occur. Therefore, a merger does not raise welfare and maybe even lowers it because of the potential of market foreclosure as is shown further below. To see that there is no double marginalization if either market level is competitive, assume, first, that the downstream market is characterized by a number of Bertrand-competitors. This implies an equilibrium price in the downstream-market equaling the wholesale price.

$$p_D = r \tag{114}$$

The upstream manufacturer faces a demand by the downstream firms equaling the demand of the end-consumers and behaves optimally by setting Q_V and p_V making the entire monopoly-profit.

Second, assume that the upstream market is characterized by a number of Bertrand-competitors facing a downstream monopolist. In this case, the equilibrium wholesale price equals marginal costs.

$$r = c \tag{115}$$

The downstream monopolist behaves optimally by setting Q_V and p_V making the entire monopoly-profit.

Besides the above **increase in pricing efficiency**, vertical mergers may also increase productive efficiency, prevent profit expropriation, or overcome problems arising from incomplete contracts and/or transaction costs (Bishop and Walker 2010: p. 466).

Productive efficiency may increase because "Non-horizontal integration may result in a more efficient use of inputs and or productive assets, due to the existence of economies of scope or scale, improved managerial or financial efficiency or supply assurance. For example, a vertical merger might improve productive efficiency by increasing coordination of demand and supply between upstream and downstream firms. Integrated firms may also benefit from increased coordination and information flow in the field of R&D, distribution, marketing, and from the adoption of new production technologies" (Bishop and Walker 2010: p. 467).

Vertical integration may also help to **prevent profit expropriation** and thus contributes to product quality and promotes R&D efforts. To see this, think of a proprietor-owned photography shop that offers superb advice when buying a camera. The owner might refuse to sell a certain type of camera when it must be afraid that after offering good advice, the customer buys the product at a lower price at an electronics chain store. Vertical integration may help to mitigate this problem and contribute to promoting this type of camera.

"Efficiencies arising from the existence of **incomplete contracts** [*emphasis added*] and transaction costs represent an important pro-competitive motivation for vertical mergers. Where contracts are not fully specified, a party may be able to exploit contractual loopholes to the disadvantage of their vertical partners. […] Clearly, such concerns do not arise under common ownership. […] In addition, conglomerate mergers may **reduce transaction costs** [*emphasis added*] by providing a one-stop-shopping opportunity for customers. This leads to an efficiency gain whenever transaction costs are non-negligible, either for the seller or for the buyer. "

Anti-Competitive Effects of Vertical Mergers: Input Foreclosure (Theory)

"We now turn [...] to the [...] case in which both upstream and downstream markets are oligopolies. This raises another important issue that needs to be considered explicitly. Beyond the desire to reduce or eliminate double marginalization, there is an additional motive for vertical integration that is more clearly anticompetitive. The motive is the possibility of market foreclosure. That is, the merger of vertically related firms might result in an upstream-downstream company that can either deny downstream rivals a source of inputs, or upstream competitors a market for their products. [...]

The telecommunications industry is one in which foreclosure concerns have been quite real for regulatory authorities in both the United States and in Europe. In this industry, the local telephone network has generally been monopolized by a firm that also competes in the more competitive long-distance market. Since a long-distance provider, such as Sprint Nextel has to gain access to its potential customers by connecting to the local network, the local network provider has the potential to price its long-distance competitors out of the market by charging them a very high price for network access or, in an extreme case, denying them access to the network at all. Accordingly, a major concern of the regulatory authorities has been the prices that suppliers of local telephone networks are allowed to charge for access to the local network. [...] In short, foreclosure arguments suggest that monopoly power in one, say upstream, market may be leveraged into power in another, downstream market" (Pepall et al. 2008: pp. 438-439).

"A merger is said to result in foreclosure where actual or potential rivals' access to supplies or markets is hampered or eliminated as a result of the merger, thereby reducing these companies' ability and/or incentive to compete. [...] Two forms of foreclosure can be distinguished. The first is where the merger is likely to raise the costs of downstream rivals by restricting their access to an important input (**input foreclosure**). The second is where the merger is likely to foreclose upstream rivals by restricting their access to a sufficient customer base (**customer foreclosure**) [*emphases added*]" (non-horizontal merger guidelines paras. 29-30).

To see the rationale of input foreclosure, consider the following example. A **monopolist upstream** produces a good at marginal cost c and sells this good to **two downstream-firms** at wholesale price r. The downstream-retailers resell the good at price p. They only incur marginal costs in the form of the wholesale price r. an compete à la Cournot. Downstream demand for the good is given by

$$Q = a - p \quad . \tag{42}$$

For the case of **vertical separation**, it can be easily checked that the two downstream firms maximize profits by setting an aggregate quantity of

$$Q_D = 2/3 \cdot (a - r) \tag{116}$$

selling the good at price

$$p_D = r + \frac{a-r}{3} \quad , \tag{117}$$

which implies profits worth

$$\pi_D = \left(\frac{a-r}{3} \right)^2 \quad . \tag{118}$$

From equation (116) we can infer the upstream monopolist's derived demand as

$$r = a - \frac{3}{2} \cdot Q \quad . \tag{119}$$

Again, it can easily be checked that the upstream manufacturer maximizes its profits at

$$\begin{aligned} Q_U &= \frac{a-c}{3} \\ r_U &= \frac{a+c}{2} \\ \pi_U &= \frac{(a-c)^2}{6} \end{aligned} \quad . \tag{120}$$

Before moving to the situation of vertical integration, note that competition downstream lowers the downstream price and raises consumer surplus as compared to the above example of monopolies both upstream and downstream. For a=110 and c=10, we observe r=60 in both cases. However, the downstream price amounts to p_D=85 in the monopoly-case and to p_D=76.67 in the duopoly-case.

Now consider the **vertically integrated situation**, where the upstream-monopolist merges with one of the downstream firms. It sells the good to the independent retailer (h) at wholesale price r and to the downstream retailing division (l) of the integrated firm at marginal costs c. Given the sample-calculations on page 49, this implies the following equilibrium at the downstream-market.

$$\begin{aligned} q_l &= \frac{a - 2\mathrm{c} + r}{3} \\ q_h &= \frac{a - 2\mathrm{r} + c}{3} \\ Q &= \frac{2\mathrm{a} - c - r}{3} \\ p_D &= \frac{a + c + r}{3} \end{aligned} \tag{121}$$

Now suppose, the vertically integrated monopolist sets $r>(a+c)/2$ such that the independent retailer is driven out of the market. In this case, it would make profits $\pi_V=(a-c)^2/4$ as implied by equation (113), which in the above example take a value of 2,500. This is greater than the sum of profits of 1.666.67+277.78 that it made in the above competitive situation, i.e. when setting $r=(a+c)/2$ (see equations (118) and (120)). Therefore, choosing a foreclosure-strategy is profitable for the integrated firm.[75] Restricting a downstream firm's access to an important input is called **input foreclosure** (non-horizontal merger guidelines para. 30).

Input foreclosure may either benefit or harm consumers. In the above example, the foreclosure-strategy would result in a downstream price of $p_V=(a+c)/2=60$ (see equation (113)) which is clearly better than $p_D=76.67$ in the duopoly-case. This **efficiency** is caused by the elimination of double marginalization. However, if in the vertically separated situation the two downstream firms would have competed à la Bertrand, prices would have risen from $p=c=10$ to $p_V=60$. Therefore, the Commission specifies in its non-horizontal merger guidelines (para. 31) "whether the increased input costs would lead to higher prices for consumers. Any efficiencies resulting from the merger may, however, lead the merged entity to reduce price, so that the overall likely impact on consumers is neutral or positive. A graphical presentation of this mechanism is provided in" Figure 22.

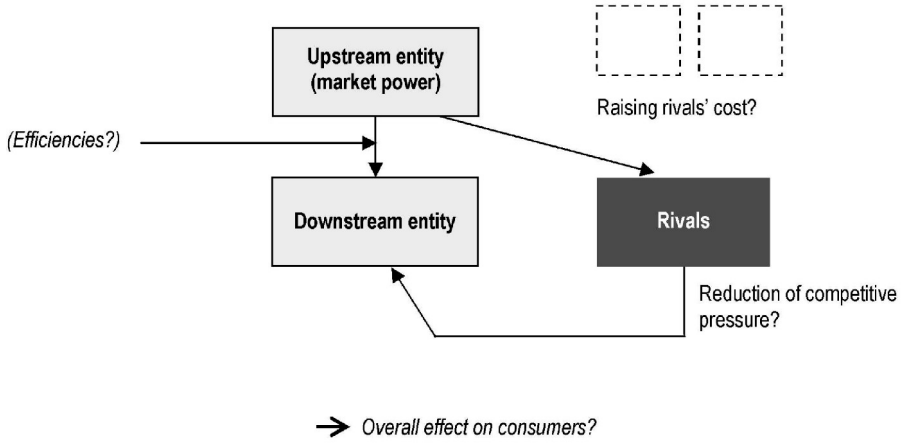


Figure 22: Input foreclosure

Source: Non-horizontal merger guidelines (para. 31)

The picture becomes somewhat more complex when the downstream product is differentiated (see, e.g., Pepall et al. (2008: ch. 17.3.2)). This is, e.g., the case in the production of cars which requires (more or less) homogeneous inputs (such as steel) purchased upstream but constitutes a differentiated product downstream (such as BMW and Mercedes). If there are several

75 Pepall et al. (2008: pp. 439-434) provide a similar argument for a Cournot-duopoly both upstream and downstream.

car-producers and several steel-producers and a single car-producer merges with a single steel-producer it may be profitable to refuse selling steel to the other car-producers. However, this may make it profitable for the other car- and steel-producers to vertically integrate which harms the firm that merged first. Therefore, the vertically integrated firm might prefer supplying the other car producers only at somewhat elevated prices to complete foreclosure, as by this behavior the subsequent vertical mergers can be prevented. Therefore, the existence of further competitors upstream and downstream may reduce the vertically integrated firm's scope for foreclosure and prevents prices in the downstream market from rising too sharply.

Anti-Competitive Effects of Vertical Mergers: Input Foreclosure (Practice and the Law)

"In assessing the likelihood of an anticompetitive input foreclosure scenario, the Commission examines, first, whether the merged entity would have, post-merger, the **ability to substantially foreclose access to inputs**, second, whether it would have the **incentive to do so**, and third, whether a foreclosure strategy would have a significant detrimental **effect on competition downstream** [*emphases added*]. In practice, these factors are often examined together since they are closely intertwined" (non-horizontal merger guidelines, para. 32). A similar assessment is performed in German merger control (BKartA 2012: 50).

The assessment of firms' **ability to foreclose access to inputs** (paras. 33-39) represents a screen for the potential exercise of market power. It refers to the various forms of input foreclosure such as refusal to supply the downstream competitors of the vertically integrated firm, an increase in the price charged from those firms, or a degradation of the supplied input's quality. Such behavior may raise competition concerns only when the input concerned accounts for a majority of costs of the downstream product, if it is a critical component in the downstream production process, or if the cost of switching to alternative inputs is relatively high. Moreover, for "input foreclosure to be a concern, the vertically integrated firm resulting from the merger must have a significant degree of market power in the upstream market" (para. 35). Finally, the Commission will assess whether there are effective and timely counter-strategies that the affected firms might employ.

The assessment of firms' **incentive to foreclose access to inputs** (pars. 40-46) considers whether the merged firm would adopt an anti-competitive strategy given it had the ability to do so. It refers to the effect of foreclosure on the integrated firm's profits. By foreclosing, "the merged entity faces a trade-off between the profit lost in the upstream market due to a reduction of input sales to (actual or potential) rivals and the profit gain, in the short or longer term, from expanding sales downstream or, as the case may be, being able to raise prices to consumers" (para. 40).

Analyzing this trade-off requires measuring the margins and their change both upstream and downstream. Moreover, one needs to determine whether by means of foreclosure a substantial part of downstream customers can be diverted from the dis-advantaged downstream competitors to the downstream division of the merged firm.

Finally, the evaluation of input foreclosure requires an assessment of the **overall likely impact on effective competition** (paras. 47-57). This impact mainly refers to the increase of prices in the downstream market because the merged firm, e.g., increases the costs of its downstream rivals or raises barriers to entry of potential competitors both upstream and downstream. "The effect on competition on the downstream market must also be assessed in light of countervailing factors such as the presence of buyer power or the likelihood that entry upstream would maintain effective competition" (para. 51). Further, the effect of possible efficiencies must be assessed such as the internalization of double marginalization. "A vertical merger may further allow the parties to better coordinate the production and distribution process, and therefore to save on inventories costs" (para. 56).

> Bishop and Walker (2010: 8-015-017) present the TomTom/Teleatlas merger as an example where the Commission analyzed the potential for foreclosure and finally cleared the merger: *"The acquisition of Tele Atlas by TomTom in 2008 was one of the first cases in which the Commission sought to apply the methodology described in the Non-Horizontal Merger Guidelines. TomTom produces portable navigation devices (PNDs), or satnavs as they are more commonly known. TomTom was the leading PND manufacturer in Europe at the time. Tele Atlas produced the maps that are required by PNDs and was the largest of only two producers of "navigable maps" in Europe at the time. At first glance, this merger might appear, to be very problematic; the largest downstream firm was buying the largest upstream firm in an upstream duopoly. However, after a Phase II investigation, the merger was cleared as the Commission decided that there was no plausible theory of harm. Although the Commission investigated both customer foreclosure and input foreclosure, the main competition concerns related to input foreclosure.*
>
> *The theory of harm relating to input foreclosure came in two forms: complete and partial input foreclosure. Complete input foreclosure would involve the merged entity no longer supplying maps to TomTom's rivals. The result would be that Navteq would effectively become the monopoly supplier to TomTom's rivals and so would be able to raise its prices. This would cause TomTom's PND rivals to raise their prices and this would benefit TomTom by reducing the competitive· pressure it faced from*

rivals. It could take advantage of this, by raising its own prices or selling more PNDs, or a mixture of both. Partial input foreclosure was a, similar theory of harm, but under this theory the merged entity would not stop supplying downstream rivals entirely, but would raise its prices to them (or lower the map quality), thus allowing Navteq to also raise its prices and thus reducing the downstream competition faced by TomTom.

This theory of harm' is economically coherent, but whether it is applicable in practice is an empirical matter: Applying, the "ability, incentive, harm to consumers", methodology, the Commission first concluded that the merged entity would have the ability to engage in input foreclosure. The Commission then considered whether the merged entity had an incentive to do so. The theory implies that the merged entity sacrifices profits from its upstream arm (i.e. it reduces map sales) in the expectation of earning increased profits downstream (from higher PND prices and/or more PND sales) that outweigh the lost upstream profits. One factor that is clearly relevant here is the relative margins upstream and downstream. It turned out that downstream margins were considerably higher than upstream margins which suggests that the theory might have been applicable. However, the Commission dismissed the concern, because it found that the effect of higher map prices on PND prices was likely to be very low. 'First,'maps account for only a small proportion of the cost of a PND (less than 10 per cent), so even a significant increase in map prices would have little effect on PND prices. This implies that there would be little reduction in the competitive pressure on TomTom. Secondly, the Commission carried, out an econometric estimation of the cross-price elasticity between TomTom and its rivals and found that it was low, so even a significant increase in the price of rival PNDs would not, significantly increase the demand for TomTom'PNDs. As a result, the Commission concluded that the merged entity would not have an incentive to engage in input foreclosure as it would likely lose more margin from the reduction in upstream sales than it would gain in downstream margin. Given, this conclusion, it was clear that the. merger was not likely to harm consumers.

The Commission also considered the efficiency claims of the parties. As noted in the Non-Horizontal Merger Guidelines; vertical mergers can lead to pricing efficiencies' due to the elimination' of double marginalisation. The Commission accepted that there were likely to be such efficiencies in this case but only after confirming that these efficiencies could not be achieved without the merger.

Specifically, the Commission investigated whether maps could be sold on the basis of a fixed fee plus a variable component related to marginal cost. It concluded that such contracts were not used in the industry and so the price efficiencies were merger specific. The parties also claimed non-pricing efficiencies related to the ability of the integrated company to better update their maps on the basis of information provided by TomTom's users. The Commission accepted that the parties' claim that the merger, would lead to "better maps – faster" was likely to be correct."

Anti-Competitive Effects of Vertical Mergers: Customer Foreclosure

"**Customer foreclosure** [*emphasis added*] may occur when a supplier integrates with an important customer in the downstream market. Because of this downstream presence, the merged entity may foreclose access to a sufficient customer base to its actual or potential rivals in the upstream market (the input market) and reduce their ability or incentive to compete. In turn, this may raise downstream rivals' costs by making it harder for them to obtain supplies of the input under similar prices and conditions as absent the merger. This may allow the merged entity profitably to establish higher prices on the downstream market. Any efficiencies resulting from the merger, however, may lead the merged entity to reduce price, so that there is overall not a negative impact on consumers. For customer foreclosure to lead to consumer harm, it is thus not necessary that the merged firm's rivals are forced to exit the market. The relevant benchmark is whether the increased input costs would lead to higher prices for consumers. A graphical presentation of this mechanism is provided in" Figure 23 (para. 58).

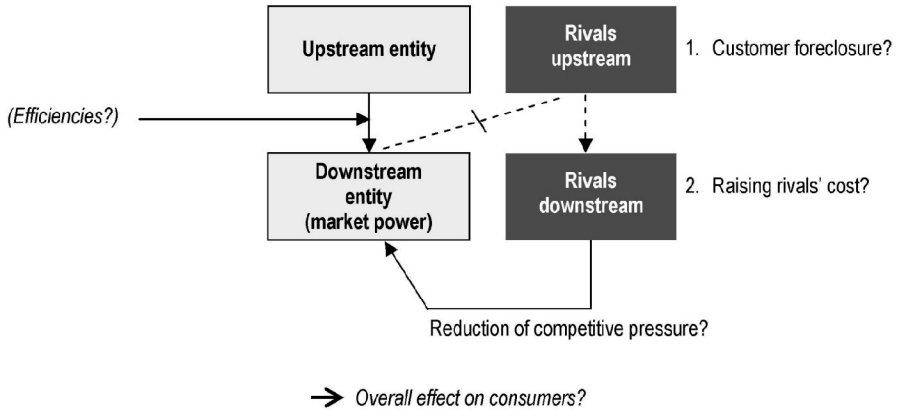


Figure 23: Customer foreclosure

Source: Non-horizontal merger guidelines (para. 58)

These effects are illustrated by the following model. Consider **two upstream firms** produce a homogeneous good at marginal cost *c*. The duopolists compete à la Cournot and sell the good at wholesale price *r* to a **downstream monopolist** who faces demand

$$Q = a - p \tag{42}$$

and sells the good at price *p* to its downstream customers. Profit maximization implies that the downstream monopolist in case of **vertical separation** chooses the following market equilibrium.

$$\begin{aligned} Q_D &= (a-r)/2 \\ p_D &= (a+r)/2 \\ \pi_D &= (a-r)^2/4 \end{aligned} \tag{122}$$

Hence, the upstream duopolists face an inverse demand function $r = a - 2Q$. Profit-maximization implies the following equilibrium in the upstream market.

$$\begin{aligned} Q_U &= 2 \cdot q_i = 2 \cdot \frac{a-c}{6} \\ r_U &= (a+2c)/3 \\ \pi_{U,i} &= (a-c)^2/9 \end{aligned} \tag{123}$$

Now, suppose one of the upstream firms **vertically integrates** with the downstream monopolist. The profit-function of the merged firm becomes

$$\pi_V = (a - q_V - q_{nV} - c) q_V + (a - q_V - q_{nV} - r) q_{nV} \quad . \tag{124}$$

The vertically integrated firm must choose whether it buys some quantity q_{nV} from the independent supplier at price *r*, or whether it sells some quantity q_V that is produced at marginal costs *c* by its own upstream division. Maximizing the integrated firm's profits yields the following first-order condition.

$$\begin{aligned} \frac{d\pi_V}{dq_V} = 0 \quad &\rightarrow \quad a - 2q_V - 2q_{nV} = c \\ \frac{d\pi_V}{dq_{nV}} = 0 \quad &\rightarrow \quad a - 2q_V - 2q_{nV} = r \end{aligned} \tag{125}$$

It is easy to see that the vertically integrated firm only buys from the independent supplier if it sets a wholesale price at the level of the integrated firm's marginal costs ($r=c$). For $r>c$ the vertically integrated firm would maximizing its profits by producing the input itself at marginal cost *c*. For $r<c$ the independent upstream supplier would make losses. Thus, the independent supplier may choose between making zero profits in the market or exiting the market as it is effectively ripped of its customers (**customer foreclosure**). Again, the price and quantity supplied in the downstream market are found as

$$\begin{aligned} Q_V &= (a-c)/2 \\ p_V &= (a+c)/2 \end{aligned} \tag{113}$$

We find that for, e.g., a=110 and c=10 the price in the vertically integrated situation is p_V=60 while it is p_D=76.67 in the situation of vertical separation. Here, vertical integration benefits consumers because of the **efficiency**-effect, i.e. the elimination of double marginalization, and despite of the customer foreclosure of the independent supplier. In this situation, vertical integration is no worse than a situation with Bertrand-competitors upstream and a vertically dis-integrated monopolist downstream. In this case, the downstream monopolist would be supplied by the upstream duopolists at wholesale price r=c and would optimally choose Q_V and p_V (see equation(113)) in the downstream market. In order to evaluate the effects of customer foreclosure, the Commission examines the firms' ability to foreclose access to downstream markets, their incentives to reduce purchases upstream, and whether a foreclosure strategy would have a significant detrimental effect on consumers in the downstream market.

Firms' **ability to foreclose access to downstream markets** (non-horizontal merger guidelines, paras. 60-67) relates to the forms of customer foreclosure as the merged entity's decision to reduce or stop purchasing from independent upstream firms. “For customer foreclosure to be a concern, it must be the case that the vertical merger involves a company which is an important customer with a significant degree of market power in the downstream market. If, on the contrary, there is a sufficiently large customer base, at present or in the future, that is likely to turn to independent suppliers, the Commission is unlikely to raise competition concerns on that ground” (para. 61). In addition to the effects outlined in the above example, customer foreclosure can lead to higher input prices if it prevents the upstream firms to exploit economies of scale or scope (see section C.1). It may also render entry into the upstream market unprofitable or – by reducing the profits of the upstream firms – reduce their ability to innovate or invest in process or product innovations (see section C.2). The ability to foreclose also depends on the upstream firms' potential for counterstrategies such as more aggressive pricing or supplying the input to other markets.

“The **incentive to foreclose [access to downstream markets]** [*emphasis added*] depends on the degree to which it is profitable. The merged entity faces a trade-off between the possible costs associated with not procuring products from upstream rivals and the possible gains from doing so, for instance, because it allows the merged entity to raise price in the upstream or downstream markets. The costs associated with reducing purchases from rival upstream suppliers are higher, when the upstream division of the integrated firm is less efficient than the foreclosed suppliers” (paras. 68-71). Finally, one needs to assess the **overall likely impact on effective competition** with regard to the affected firms' customers.

"Reflecting the fact that customer foreclosure is less likely to give rise to competition concerns than input foreclosure, potential concerns of customer foreclosure have arisen less frequently in the Commission's merger assessments [prior to 2008] than input foreclosure concerns. Moreover, such concerns have yet to be assessed in detail following the publication of the Non-horizontal Merger Guidelines [in 2008]" (Bishop and Walker 2010: 8-026). Therefore, no recent example for customer foreclosure can be provided to date.

Conglomerate Mergers

"The final type of merger to consider is a conglomerate merger. Such mergers bring under common control firms whose products are neither direct substitutes nor complements. The outcome is a set of firms producing a diversified range of products with little or nothing in common" (Pepall et al. 2008: p. 450). "A distinction can be made, between conglomerate mergers that involve **complementary products**, and conglomerate mergers that involve **unrelated products**. [… C]omplementary products [...] are products where a consumer's increased demand for one product generates increased demand for its complement. Unrelated products are those which have no obvious commonality, i.e. they are neither substitutes nor complements for one another" (Bishop and Walker 2010: 8-030).

"**Scope economies** [*emphasis added*] and saving on transactions costs are two possible advantages that may accrue to conglomerate firms. By scope economies we mean that a variety of products or services are more cheaply produced by one firm than by two or more firms [see section C.1]. [...] This line of argument implies that for scope economies to be an important element in conglomerate mergers it is necessary that the firms that merge are related in some respect. Either they sell in similar markets or they have similar production technologies. The data on conglomerates do not appear to be consistent with this hypothesis. A detailed study by Nathanson and Cassano (1982) concludes that there are at least as many conglomerate firms that produce goods with little in common, whether this be technology or the markets at which they are targeted, as there are firms that have relatively low product and market diversity" (Pepall et al. 2008: p. 450).

"By **transaction costs** [*emphasis added*] we mean the costs that are incurred by firms when they use external markets in order to exchange goods and services. These include, for example, the costs of searching for the desired inputs, negotiating supply contracts, monitoring and enforcing these contracts and the risk associated with unforeseen changes in supply conditions. [...] Transaction cost problems are particularly important when the asset involved is knowledge or information intensive. The knowledge of such matters as organizational routines or specialized customer needs is generally embodied in specific individuals or teams employed by the company. It

is difficult to envision contracts to "lease" such personnel. In short, the effort to minimize the transaction costs associated with contracting between firms may explain conglomerate mergers to some extent. Nevertheless this motivation seems unlikely to be the major factor behind such mergers. The reason is that here again, we are talking about some asset that is common to all the lines of production operated by the conglomerate, and such commonality in productive assets does not seem to be a feature of actual conglomerate firms" (Pepall et al. 2008: pp. 450-451).

"The motive for conglomeration may be that it is in the **interest of management** [*emphasis added*] even if it is not in the interest of shareholders. Because management calls the shots, it is the managerial interest that prevails. In any reasonably large public company, ownership, which essentially resides with the shareholders, may be separated from control, which essentially resides with the management team. This separation would not matter too much if management performance could be perfectly observed and monitored by shareholders. Yet perfect monitoring is rare [see section C], and absent such monitoring, management can pursue its own agenda at least to some extent. This would not matter so long as the best interests of management are served by maximizing shareholder wealth. It is precisely the attempt to secure this harmony of interest that lies behind the use of performance-related clauses and payment in stock options in many executive compensation schemes. Still, the match between the interests of shareholders and management is rarely perfect, leaving management with at least some ability to pursue goals other than maximization of shareholders' returns.

Suppose that **management compensation** [*emphasis added*] is based upon company growth. Growth is far from easy to generate internally. It requires that market share be won from competitors who can hardly be expected to sit passively by when they lose customers. Nor is it easy to buy growth through horizontal merger since this is the kind of acquisition that is watched by the antitrust authorities. In these circumstances, we should not be surprised to find management supporting a conglomerate merger, even if this is not necessarily in the best long-term interests of shareholders. Such a merger offers management the desired growth while avoiding antitrust problems. [...]

Management may also pursue conglomeration as a means to minimize risk. When a firm is involved in many distinct markets it avoids putting "all its eggs in one basket." [...] This smoothes the firm's income stream because with many product lines operating, positive and negative shocks tend to cancel each other out. The derived income stream of the firm's executives is also smoother. Even shareholders might prefer this approach if, in the absence of such **diversification** [emphasis added], the firm would have to pay its executives higher salaries to compensate them for the greater risk. This may be particularly true for managers who are heavily invested in the firms so that not

only their labor income but also their capital income is subject to the same risk" (Pepall et al. 2008: pp. 451-453).

Bishop and Walker (2010: 8-031) provide further reasons for conglomerate mergers to be profitable:

- "The merged firm would be more attractive to buyers because it would offer a range of products. […]
- The merged entity will enjoy greater financial resources, thereby strengthening its competitive position vis-a-vis its remaining competitors – the "big is bad" theory of harm.
- The merger increases the threat of a refusal to supply.
- The merger creates greater opportunities for tying products."

Possible anti-competitive consequences of conglomerate mergers mainly relate to the last of these points and may consist in **foreclosure** as "the ability and incentive to leverage a strong market position from one market to another by means of **tying or bundling** [*emphasis added*] or other exclusionary practices" (non-horizontal merger guidelines, para. 93).[76] "Tying usually refers to situations where a product can only be purchased if another product is purchased. Tying can be "**commercial tying**" [*emphasis added*] (known as pure bundling), where the supply of one product is predicated on the purchase of another, or technical. **Technical tying** [*emphasis added*] refers to situations where only certain products are technically compatible with another product. **Bundling** [*emphasis added*] usually refers to the practice of mixed bundling, i.e. situations where it is possible to buy two products separately, but a discount is obtained if both products are purchased at the same time. So while under certain specific circumstances conglomerate mergers may give rise to adverse outcomes for consumers, they are even less likely to give rise to competition concerns than vertical mergers" (Bishop and Walker 2010: p. 454).

The BKartA (2012: 58) also notes that conglomerate mergers can involve **potential competitors** and result in the elimination of the effects of such potential competition. Moreover, **portfolio effects** may allow the merged firm to exercise market power "if consumers find a wider range of products advantageous and prefer to buy these products from one supplier ('one-stop shopping')" (BKartA 2012: 61). Additionally, "a conglomerate merger may also lead to an increase in market power because it strengthens the financial or industry-specific resources of the merged company" (BKartA 2012: 62).

Bishop and Walker (2010: pp. 463-464) present the conglomerate merger of GE and Amersham as an example for the assessment of of such mergers: *"The impact of a*

76 Tying and bundling are described in greater detail in section I .

loss of market share on competitors' ability to compete and the impact of that loss of market share on final consumers was [...] directly addressed in the Commission's assessment of General Electric's acquisition of Amersham. The products affected by the merger were medical diagnostic products. GE produces diagnostic imaging (DI) equipment to produce internal images of the body, while Amersham produces diagnostic pharmaceuticals (DPs) to enhance the clarity of the images produced by DI equipment. DI equipment and DPs can therefore be considered to be complementary products.

The Commission considered whether the merged entity might acquire the ability and incentive to foreclose competition "by leveraging its pre-merger market power from one market to another through exclusionary practices, such as bundling and/or tying." The Commission considered possible theories of harm based on bundling of DI equipment and DPs and on technical tying.

[...] In assessing commercial (or mixed) bundling, the Commission examined whether:

(1) the merged entity would be able to leverage pre-merger dominance in one market into another;

(2) rivals would be unable to respond to the bundling strategy;

(3) the resulting marginalisation of competitors would result in their exit from the market; and

(4) following the exit of these competitors, the merged firm would be able to implement unilateral price increases that would not be undermined in the long run either by new entry or by the re-entry of previously marginalised competitors.

[...] The Commission found that none of these conditions would hold and consequently cleared the merger at Phase 1. The decision explicitly notes that even if commercial bundling were to materialise as a result of the merger, it would not result in the foreclosure of competition. The Commission also dismissed concerns of technical tying whereby the merged entity would develop DPs that are only compatible with its DI equipment. In addition to examining whether such a strategy would actually be possible, the Commission also examined the economic incentives to engage in technical tying and concluded that this would not be a profit-maximising strategy since this would deny the merged entity access to the installed

base of DI equipment."

Lessons Learned

After reading this section you should be able to answer the following questions.

1. What is meant by the term *merger paradox* and how can it be resolved?
2. Explain the terms *strategic complements* and *strategic substitutes*.
3. Name and explain the best practices that should be followed when performing a merger simulation.
4. What is the relevant legislation in the EU that governs merger control?
5. What types of mergers are considered having a *Community dimension*?
6. What role does market definition play in merger control?
7. Describe the substitution effects that occur when firms merge in an industry with differentiated products.
8. What is meant by the term *efficiency defense*?
9. What types of non-horizontal mergers may be distinguished?
10. What is meant by the term *double marginalization*? How can double mark-ups be avoided by means of vertical integration?
11. Describe the meaning of the term and the effects of *input foreclosure*.
12. Describe the meaning of the term and the effects of *customer foreclosure*.
13. Name some economic reasons for conglomerate mergers.

References

Bishop, S. and Walker, M. (2010). "The Economics of EC Competition Law." 3rd edition. Sweet and Maxwell: London

BKartA (2012). "Guidance on Substantive Merger Control." http://www.bundeskartellamt.de/wEnglisch/download/pdf/Merkblaetter/2012-03-29_Guidance_final_neu.pdf

Brandenburger, A.M. and Nalebuff, B.J. (1996). "Co-opetition." Doubleday: New York

Davies, P.J. and Garcés, E. (2010). "Quantitative Techniques for Competition and Antitrust Analysis." Princeton University Press: Princetin and Oxford

Farrell, J. and Shapiro, C. (1990). "Horizontal Mergers: An Equilibrium Analysis." The American Economic Review. Vol. 80 No. 1, pp. 107-126

Pepall, L. and Richards, D. and Norman, G. (2008). "Industrial Organization – Contemporary Theory and Empirical Applications." 4th edition. Blackwell Publishing: Malden, MA

Röller, L.-H. (2011). "Efficiencies in EU Merger Control: Do They Matter?" in: Marktmacht. Ramser, H.J. and Stadler, M. (ed.). Mohr Siebeck: Tübingen

H ART. 101 TFEU – AGREEMENTS BETWEEN UNDERTAKINGS / STRATEGIC INTERACTION IN DYNAMIC INDUSTRIES

The previous sections are primarily concerned with market-outcomes when the firms in an industry compete rather than cooperate. The following characteristics are identified as the main determinants of the market-outcome

- the structure of the industry (see, e.g., section A on perfect competition and monopoly),
- the production technology (see section C), and in connection with the technology
- firms' competitive conduct as, e.g., price- or quantity-setting entities (see section B).

These characteristics are shown to affect firms' market power (see section E), whose persistence depends on

- supply-side features such as the (non-)existence of barriers to entry and
- demand-side features such as buyer power, switching costs, and network effects.

In this section, we analyze market-outcomes when (some) **firms cooperate horizontally** in order to affect product market outcomes. Such cooperation may concern agreements prices, quantities, and/or product quality. The term *horizontal* means that the agreements are made among competitors in the same product market. Horizontal cooperation of this type can be achieved by explicitly agreeing on cooperative conduct (**explicit collusion**) or by a mutual understanding among the firms (**tacit collusion**). Both forms of collusion are typically intended to raise prices and lower output which redistributes rents from buyers to sellers and lowers total welfare.

An introduction to the economics of collusion is presented in subsection H.1 . There, we present an introduction to **dynamic game theory** because collusion is not sustainable in the one-shot games analyzed above. Subsection H.2 presents factors that facilitate collusion and measures of competition policy that are used to cope with cartels because explicit collusion, i.e. the formation of a cartel, is illegal in most jurisdictions.

> Art. 101 I and II TFEU prohibits *"all agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition within the internal market, and in particular those which:*
>
> *(a) directly or indirectly fix purchase or selling prices or any other trading conditions;*

(b) limit or control production, markets, technical development, or investment;

(c) share markets or sources of supply;

(d) apply dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage;

(e) make the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts.

Any agreements or decisions prohibited pursuant to this Article shall be automatically void."

Many agreements between firms are not directly aimed at hardcore anti-competitive behavior (fixing purchase or selling prices, controlling production, or sharing markets). For example, production, purchasing, or research and development agreements can also have pro-competitive effects. In this case, Art. 101 III TFEU specifies:

"The provisions of paragraph 1 may, however, be declared inapplicable in the case of:

- *any agreement or category of agreements between undertakings,*
- *any decision or category of decisions by associations of undertakings,*
- *any concerted practice or category of concerted practices,*

which contributes to improving the production or distribution of goods or to promoting technical or economic progress, while allowing consumers a fair share of the resulting benefit, and which does not:

(a) impose on the undertakings concerned restrictions which are not indispensable to the attainment of these objectives;

(b) afford such undertakings the possibility of eliminating competition in respect of a substantial part of the products in question."

This means that an agreement's anti-competitive effects must be traded-off against its pro-competitive effects. For **horizontal co-operation agreements** this process is described in subsection H.3 . Subsection H.4 is concerned with **vertical restraints**.

H.1 Dynamic Game Theory and Collusion

Collusion Basics

In economics, collusion is a situation where firms' prices are higher than some competitive benchmark (Motta 2004: ch. 4.1.1), i.e. **collusion coincides with an outcome** (high-enough price), and not with a specific form of cooperation through which that outcome is attained. This is important because in competition law the defining characteristic of a cartel is the explicit cooperation among firms no matter whether this agreement is effective in terms of its effects on the market or not. Collusive market outcomes can be attained when firms act through an organized cartel (explicit collusion) or when they act in a purely non-cooperative way (tacit collusion).

Explicitly collusive agreements can take different forms: firms might agree on sales prices, allocate quotas among themselves, divide markets so that some firms decide not to be present in certain markets in exchange for being the sole seller in others, or coordinate their behavior along other dimensions. In any case, the firms may have an incentive to jointly behave as if they were a monopolist and, thus, set monopoly-prices. In section A.3 , we show that the monopoly is characterized by higher prices and, thus, a higher producer surplus than, for example a perfectly competitive market. Similarly, consumer surplus is lower. This increase in prices is generated by a joint reduction in quantity. In other words, collusive practices allow firms to exert market power they would not otherwise have, and artificially restrict competition and increase prices, thereby reducing welfare. Because of these **anti-competitive effects** cartels are illegal in many jurisdictions

The illegality of cartels implies that cartel-members cannot coordinate on a collusive equilibrium by legally binding contracts because every firm has an **incentive to cheat**. Below, we show by which means cartel firms can nonetheless enforce their agreements. We show that collusion cannot be enforced when the firms play a static (one-shot) game like the Bertrand- and Cournot-games presented in section B . However, collusion can be stabilized in dynamic games where firms are assumed to maximize profits over a longer period of time. In this context, we distinguish between games that end at a known point in time (**finite horizon**) and games that (potentially) can be played forever (**infinite horizon**). Such multi-period (or: dynamic) games are analyzed by **dynamic game theory** that is introduced in this section.

The Incentive to Cheat

At the price charged by a cartel, each firm's price-cost margin is relatively large. This gives each individual firm a strong incentive to sell a higher output at only somewhat reduced prices, i.e. to cheat on the agreement (Pepall et al. 2008: ch. 14.1). Yet if every firm acts according to this

incentive and chisels on the agreement by selling a little more, the extra output on the market will not be a little but a lot. Market price will fall and the price-fixing agreement will break down.

These effects can nicely be illustrated for two symmetric firms that compete à la Cournot as shown in Figure 24. The two solid lines are the reaction curves of firms 1 and 2 with the competitive equilibrium-quantities $q_{1,opt}$ and $q_{2,opt}$. If the firms collude, they set the monopoly quantity

$$q_m = \phi \cdot q_1^*(0) + (1-\phi) \cdot q_2^*(0) \quad . \tag{126}$$

This quantity is given by one of the combinations on the dashed line. Let ϕ denote the share of firm 1 in the monopoly quantity and, thus, in profits. For ϕ=0.5 the firms set quantities $q_{1,m}$ and $q_{2,m}$, making supra-competitive profits $\pi_{1,m}=\pi_{2,m}$. In the one-shot game, this is no equilibrium solution because the cooperative output levels do not constitute a pair of best responses. For example, firm 1's best response to firm 2 setting $q_{2,m}$ would be to set $q_{1,d}=q_1^*(q_{2,m})$. By setting this quantity $q_{1,d}$, firm 1 deviates from the collusive agreement. This raises firm 1's profits to a level of $\pi_{1,d}>\pi_{1,m}$. Firm 2's profit $\pi_{2,-d}$ falls even below competitive levels because the increase in aggregate quantity causes a drop in the market-price of the good. The symmetry of firms implies that firm 2 has the same incentive to deviate from the collusive agreement in the one-shot game as firm 1.

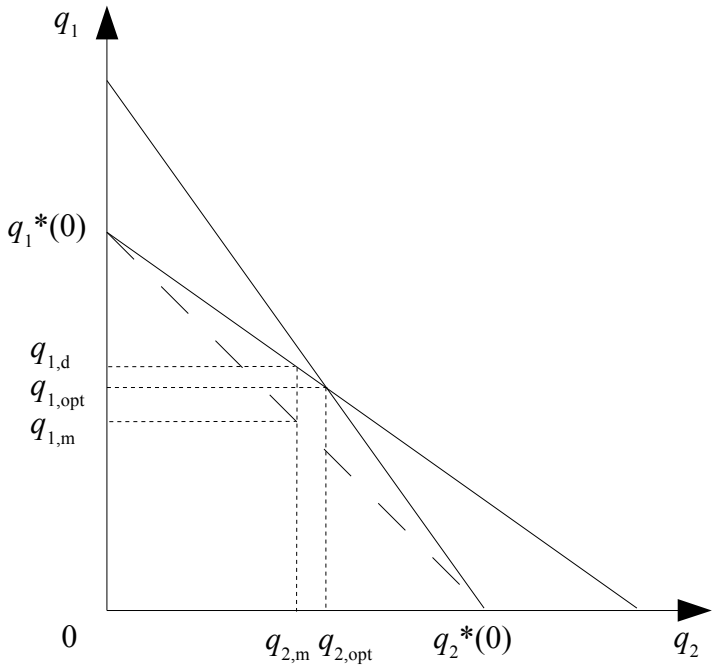


Figure 24: Incentive to Deviate in Cournot-Competition

In the one-shot game, the firms are trapped in a **prisoners' dilemma** as can be seen by the below matrix of payoffs in normal form. Given the inequality $\pi_{i,-d}<\pi_i<\pi_{i,m}<\pi_{i,d}$ one finds that the one-shot game is at equilibrium if both firms act competitively. This is because every firm reasons: "If I cooperate and the other firm cooperates, then we share the monopoly profit. However, if the other firm does not cooperate while I do, then I loose a lot of profit. If, on the other hand, I don't cooperate and the other firm does, then I make a lot of money; and if the other firm does not cooperate, then it's as if we were playing noncooperatively anyway. No matter what the other firm does. I am better of not cooperating." In other words, competing is a dominant strategy.

		firm 2	
		cooperate	*compete*
firm 1	*cooperate*	$\pi_{1,m} \mid \pi_{2,m}$	$\pi_{1,-d} \mid \pi_{2,d}$
	compete	$\pi_{1,d} \mid \pi_{2,-d}$	$\pi_1 \mid \pi_2$

The above reasoning also applies to **Bertrand-competition**. For example, in the symmetric Bertrand-duopoly from section B.2 the colluding firms set the monopoly price $p_1=p_2=p_m$ and make profits $\pi_{1,m}=\pi_{2,m}=0.5\pi_m$, where π_m is the profit of a monopolist. By undercutting this price slightly $p_1=p_m-\varepsilon$ (with $\varepsilon\to 0$) firm 1 wins the entire market and makes deviation profits equaling (almost) monopoly profits ($\pi_{1,d}\approx\pi_m$) while firm 2 makes zero profits ($\pi_{2,-d}=0$). The same gain from deviating would accrue to firm 2. Therefore, in the one-shot game collusion cannot be stabilized and the firms set equilibrium prices at the level of marginal costs $p_1=p_2=c$.

The acknowledgment that any collusive situation naturally brings with it the temptation to deviate from it and therefore to break collusion leads us to the identification of two conditions that must be met for collusion to arise. First, the colluding firms must be able to **detect** in a timely way that a deviation has occurred. Second, there must be a **punishment** for deviating from collusive conduct respectively a reward for adhering to it. A punishment may take the form of rivals producing a higher quantity in the periods following the deviation which lowers the profit of the deviator.

The reasoning, that collusion can only be stabilized if deviations can be punished in subsequent periods, requires modeling **repeated interaction** among the firms. Introducing repetition into a game-theoretic framework adds history as an element to the analysis. When players face each other over and over again, they can adopt strategies that base today's action on the behavior of their rivals in previous periods. This is what rewards and punishment are all about.

We present one such **trigger strategy** below where firms reason: "I play collusively as long as my rival plays collusively. If my rival defects from the collusive solution I will play competitively in all periods after the detection of the deviation." We show that this so-called grim trigger strategy cannot sustain collusion if the game is played for a finite number of periods (**finite horizon games**). However, in **infinite horizon games** collusion is stable if the short-run gain from deviating is smaller than the long-run loss caused by the punishment.[77]

Finite Horizon Games

Recall the assumptions of the basic Bertrand-duopoly from section B.2 (i.e. symmetric firms, homogeneous product, simultaneous price-setting, downward-sloping demand, no capacity-constraints) and relax assumption 6, which states that firms are only interested in the profits of the current period. We assume that firms play this game for T+1 periods, i.e. from period t=0 to period t=T, with T being a finite number (Motta 2004: ch. 8.4.3.1). Every firm i=1,2 is interested in maximizing its present value of profits

$$V_i = \sum_{t=0}^{T} \delta^t \cdot \pi_{i,t} \quad \text{, with} \tag{127}$$

$$\delta = \frac{1}{1+r} \tag{128}$$

being firms' **discount factor**, r being its **discount rate**, and $\pi_{i,t}$ being the profit earned in period t.

At least three situations come to mind to justify the assumption of a finite period where all firms know exactly when the game ends. First, the firms exploit an exhaustible and non-renewable resource such as oil or natural gas. Secondly, the firms operate in a market with proprietary knowledge protected by patents and all patents are awarded for a finite period only. Third, firms are run by management teams who are only employed for a couple of years. When there is a major change in management at one or more of the firms the game is likely to end. Often this end can be foreseen (Pepall et al. 2008: ch. 14.2.1).

As the terminal date of the game is known, it can be solved by **backward induction** (Motta 2004: ch. 8.3.2). The equilibrium-concept is that of a **sub-game perfect Nash equilibrium**, i.e. all players play optimally in each sub-game. In a sequential and finite game, this means to start with determining the Nash-equilibrium at the last period of the game. This is a **sub-game** of the entire game, which comprises all periods. Given the Nash-equilibrium in period T we move to period T-1 and determine the Nash-equilibrium of this sub-game. After that, we sequentially determine the

77 Note that trigger strategies also work if it takes more than one period to detect the deviation. In this context, it is convenient to define a period as the amount of time necessary to detect a deviation.

equilibria of the sub-games in periods T-2, T-3, ..., t=0. Please note that the above game is one of **almost perfect information** because more than one player move in every sub-game, i.e. in a duopoly both firms choose prices simultaneously.

For the above game the solution by backward induction looks as follows. At the last period of the game (t=T) the events of the previous periods are decision-irrelevant and the firms play the one-shot Bertrand game. Therefore, the only equilibrium of the game will be the one-shot Bertrand price.

$$p_{1,T}^*(p_2) = p_{2,T}^*(p_1) = c$$

If the firms set prices above marginal costs it would be profitable for every firm to set its price slightly below that of its rival and win additional demand. In the terminal round, the betrayed company would not be able to retaliate against such behavior in future periods. It does not possess a **credible threat** to prevent such cheating.

In period t=T-1, players know that their current choices will not affect the equilibrium solution in the following period T. Therefore, whatever happened in periods 0, 1, ..., T-2, the game that they are playing at T-1 is effectively the same as if they were playing for the last time, and again the only equilibrium is the Bertrand equilibrium.

$$p_{1,T-1}^*(p_2) = p_{2,T-1}^*(p_1) = c$$

The same reasoning can be applied to all previous periods, leading firms to choose

$$p_1^*(p_2) = p_2^*(p_1) = c \tag{22}$$

in any period t.

This result has first been proven by Nobel Prize winner Reinhard Selten[78] in 1973: If a game with a unique equilibrium is played finitely many times, its solution is that equilibrium played each and every time. Finitely repeated play of a Nash equilibrium is the Nash equilibrium of the repeated game. In our case the Bertrand game is repeated a finite number of times and its outcome is exactly the same as the one-shot game. The firms set prices at the level of marginal cost and making zero economic profits in every period.

Infinite Horizon Games

Now, suppose the number of periods is infinite (T=∞). Games of this type are called infinite horizon games or **supergames**. In the following, we show that in this case firms can sustain prices above marginal costs ($p_{1,m}$=$p_{2,m}$>c) making above-zero profits ($\pi_{1,m}$=$\pi_{2,m}$>0). When both firms sustain this collusive equilibrium forever, each makes an expected present value of collusive profits $V_{i,m}$.

78 http://nobelprize.org/nobel_prizes/economics/laureates/1994/selten-autobio.html

$$\begin{aligned} V_{i,m} &= \sum_{t=0}^{T=\infty} \delta^t \cdot \pi_{i,m,t} = \pi_{i,m} \cdot (1+\delta^1+\delta^2+\delta^3+...) \\ &= \pi_{i,m} + \delta V_{i,m} = \frac{\pi_{i,m}}{1-\delta} \end{aligned} \tag{129}$$

We omit the index t because we assume that the characteristics of the industry remain the same in all periods. Hence firm i makes the collusive profit $\pi_{i,m}$ in every collusive period.

However, both firms have an incentive to undercut and, thus, **cheat** on each other, i.e. by setting a price $p_{1,d}=p_{2,m}-\varepsilon$ (with $\varepsilon \to 0$) firm 1 wins the entire market and makes a deviation-profit $\pi_{1,d} \approx 2\pi_{1,m}$ in the current period. Playing a repeated game allows firm 2 to retaliate against the defector. In particular, suppose that firm 2 uses the following **trigger strategy**. After observing the deviation of firm 1, firm 2 reverts to the competitive equilibrium and sets $p_2=c$ forever, which implies ($\pi_1=\pi_2=0$). Hence, firm 1's present value of profits becomes

$$V_{1,d} = \pi_{1,d} + \sum_{t=1}^{T=\infty} \delta^t \cdot \pi_{1,t} = \pi_{1,d} + \delta \cdot V_1 \quad . \tag{130}$$

This game cannot by solved by backward induction because it is impossible to determine a terminal period. Firm 1 rather chooses the action (cheating or colluding) that provides the greater present value of profits. It plays the collusive equilibrium if it earns a collusive present value of profits that is at least as high as that after a deviation.

$$V_{1,m} \geq V_{1,d} \tag{131}$$

This condition can be rewritten as

$$\begin{aligned} \pi_{1,d} - \pi_{1,m} &\leq \delta (V_{1,m} - V_1) \\ &\leq \frac{\delta (\pi_{1,m} - \pi_1)}{1-\delta} \end{aligned} \quad . \tag{132}$$

Hence, firm 1 faces a trade-off, it adheres to the collusive agreement if the immediate, one-period gain from deviation (left-hand side of (132)) is smaller than the future loss in profits caused by the punishment through firm 2 (right-hand side of (132)). This condition can be rewritten as

$$\begin{aligned} \delta &\geq \frac{\pi_{1,d} - \pi_{1,m}}{V_{1,m} - V_1} \\ &\geq \frac{\pi_{1,d} - \pi_{1,m}}{\pi_{1,d} - \pi_1} \equiv \bar{\delta} \end{aligned} \quad . \tag{133}$$

Hence, the collusive equilibrium can be sustained if the firms value future profits high enough relative to current profits, i.e. if their true discount-factor δ is at least as high as the critical discount factor $\bar{\delta}$. This result is called the **Folk theorem** for infinitely repeated games: Suppose that an infinitely repeated game has a set of payoffs that exceed the one-shot Nash equilibrium payoffs for

each and every firm. Then any set of feasible payoffs that are preferred by all firms to the Nash equilibrium payoffs can be supported as subgame perfect equilibria for the repeated game for some discount rate sufficiently close to unity.

In the above Bertrand-example, it is easy to show that the **incentive compatibility constraint** (133) becomes $\delta \geq 0.5$. Given (128), the collusive equilibrium will be played as long as the discount rate does not exceed a critical level of 100% (in this example, i.e. $r \leq 1$), i.e. firms do not discount the future too heavily. It can be shown that the short-run gain from deviating (left-hand side of equation (132)) is larger in Bertrand- than in Cournot-competition because in Bertrand-competition the deviator is not capacity-constrained and wins a larger share of the market. However, the punishment, i.e. reversion to the one-shot Nash-equilibrium, is harsher in Bertrand-competition because firms would make zero economic profits. However, in Cournot-competition the competitive profits are above zero.

The assumption of an infinitely lived market is not always realistic. However, the above type of analysis also applies when the survival of a market is uncertain. To see this, let us denote the probability of a market to break down in some period by ρ. In this case, firm i makes zero profits. With probability (1-ρ) the market continues to exist and the firm makes profit $\pi'_{i,t}$. The probability that the market still exists t periods from now is $(1-\rho)^t$. Hence, we can write the firm's present value of profits (see equation (127)) as

$$V_i = \sum_{t=0}^{T} \delta^t \cdot (1-\rho)^t \cdot \pi'_{i,t} = \sum_{t=0}^{T} \gamma^t \cdot \pi'_{i,t} \quad \text{with} \tag{134}$$

$$\gamma = \frac{1-\rho}{1+r} = \frac{1}{1+r'} \quad \text{and} \tag{135}$$

$$r' = r + \frac{\rho(1+r)}{1-\rho} \quad . \tag{136}$$

Equation (135) indicates that the uncertainty about the survival of the market lowers firms' discount factor from a level of δ to γ. This can either be done (i) by explicitly including the probability of survival in the numerator of the discount factor (second term in (135)) to obtain the **probability-adjusted discount factor** γ or (ii) by increasing the discount rate to a level of r' as shown in equation (136). After re-formulating the discount factor in this way, an incentive compatibility constraint for cartels can be found for this new situation as shown in equation (133).

As an alternative to a punishment scheme the firms may employ a **compensation scheme**. If the cartel firms detect that one of their fellow co-conspirators has sold a higher output or set a lower price than agreed upon, the deviator is required to pay a compensation to the “truthful” cartel members, i.e. to make a **side payment**. Alternatively, a member who has sold more than agreed

upon may be required to buy output of cartel firms who have undersold. Levenstein and Suslow (2011: 476) find that "the implementation of punishments is considerably less common than compensation schemes, occurring in only 19 percent of the cartels in [their] sample." Moreover, they (Levenstein and Suslow 2011: 482) find that cartels with a compensation scheme are less likely to die a natural death, for example because of demand fluctuations, than cartels who employ a retaliatory punishment scheme.

H.2 (Tacit) Collusion

Explicitly collusive agreements are prohibited in many jurisdictions. For example, in Europe **Art. 101 TFEU** (Treaty on the Functioning of the European Union) prohibits *"all agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition within the internal market".*This prohibition applies most directly to cartels. However, many agreements between firms are not directly aimed at anti-competitive behavior such as fixing purchase or selling prices, controlling production, or sharing markets. For example, production, purchasing, or research and development agreements can also have pro-competitive effects. Such potentially pro-competitive horizontal agreements are analyzed in section H.3 .

In the following, we show how collusion is addressed in competition policy. In particular, we analyze how explicit collusion (cartels) can be detected, broken up, and deterred by competition authorities. Moreover, we show what factors facilitate collusion and what role explicit communication plays in stabilizing collusive agreements. The section concludes with an overview on models that examine the effect of demand movements on the stability of cartels.

Explicit Collusion, Tacit Collusion, and Standards of Proof

Economics defines collusion as a situation where firms' prices are higher than some competitive benchmark. In the economic definition of the term, collusion coincides with an *effect* (high-enough price), not with a specific *form* through which that outcome is attained, i.e. an organized cartel (explicit collusion) or by an implicit, non-cooperative understanding of the firms (tacit collusion). "**Tacit collusion** [*emphasis added*] exists where in the absence of any formal attempts to implement a collusive outcome, firms understand that if each firm competes less vigorously they might all be able to enjoy higher prices and higher profits. For example, a firm may realise that cutting prices will lead to rival firms following suit. Hence, the best the firm can do given the likely reactions of its rivals is to maintain prices at the current level. Equally, a firm may believe that if it raises its

price, its rivals will also raise their prices, thus making the price rise profitable" (Bishop and Walker 2010: p. 164).

In subsection H.1 we are concerned with self-enforcing, tacitly collusive agreements while competition law penalizes explicit collusion. For example, this can be seen by Art. 101 I TFEU that prohibits agreements between firms that *have as their object* the prevention, restriction or distortion of competition. This raises two issues. First, according to this legal definition cartels can exist even if the market outcome is not collusive according to the economic definition of collusion. Second, competition policy relies on economic models of tacit collusion in order to analyze explicitly collusive agreements. However, the "factors which imply that tacit collusion is more likely are similar to those which make sustainable cartel behaviour more likely. For this reason, those industries in which firms have formed, or attempted to form, cartels are more often viewed as being prone to tacit collusion" (Bishop and Walker 2010: p. 165). Below, we present **factors that facilitate collusion**.

It "should be noted that in general firms should prefer to collude tacitly rather than explicitly since explicit cartel behaviour lays firms open to more serious legal penalties, such as large fines or [...] criminal penalties. This implies that firms should only engage in explicit collusion where tacit collusion would not be successful. Where conditions are such that it is relatively easy for firms to tacitly collude successfully (i.e. to achieve close to the monopoly equilibrium), we should not expect to see explicit cartels" (Bishop and Walker 2010: p. 165).

Even if tacit collusion is not addressed by Art. 101 TFEU, it can be captured in a variety of ways in EU competition policy:

1. Tacit collusion in the form of **coordinated effects** is addressed within EU **merger control** (see section G).
2. "In the past, there was a debate as to whether the concept of collective dominance was wider than the concept of tacit coordination that underpins the notion of coordinated effects. At the European level this debate is effectively over; it is generally accepted that the concepts of **collective dominance** [*emphasis added*] and tacit coordination are equivalent from an economic perspective" (Bishop and Walker 2010: p. 356). Therefore, tacit collusion may be incriminated within **Art. 102 TFEU** (see section I).
3. Tacit collusion (conscious parallel behavior) may be regarded as a **softening of competition** and, thus, falls within the scope of the EU's **Guidelines on Vertical Restraints** (para. 100) (see section H.4).

This raises a difficult policy issue. "How can one **discriminate between anti-competitive tacit**

behaviour and normal effective competition [*emphasis added*]? [...] Given the difficulties in discriminating between tacit collusion and effective competition, policy should not as a general rule attempt to intervene in a market on suspicion that firms are competing tacitly without there being very good reasons to believe that the market outcome is substantially different to the effectively competitive outcome. It should also be noted that it is not clear what remedy the Commission could impose in the event of a finding of tacit collusion. The Commission dislikes (with good reason) price regulating firms and it does not impose structural remedies in non-merger cases. The implication of this is that the principal policy weapon against tacit collusion is likely to be merger control [i.e. analyzing the potential for **coordinated effects**] and is therefore preventive in nature rather than corrective" (Bishop and Walker 2010: pp. 165-166).

Let us return to **explicit collusion and cartels**. Self-enforcing collusive outcomes (tacit collusion) can be difficult to achieve so that firms may have to communicate to sustain the collusive market outcome. For example, it can be difficult for the firms to determine the optimal price or the optimal quantity without explicit communication. This is particularly a problem if the industry is subject to changing business conditions, for example demand shocks, which require frequent adjustments of price and/or quantity.[79] With **explicit collusion** firms talk to each other and coordinate on their collusive strategies without having to experiment with the market, which is costly. Furthermore, if there are some shocks which modify market conditions, communication will allow the firms to change to a new collusive price without the risk of triggering a period of punishment.

Institutional arrangements to sustain collusion take on a variety of forms. For example, **market allocation** schemes, according to which a firm sells in a certain region (or serves customers of a certain type) have the advantage of allowing for prices to adjust to new demand or cost conditions without triggering possible price wars. As long as each firm does not serve market segments allocated to other cartel members, prices can be varied without disrupting the collusion. This explains why such collusive schemes are often used. For example, the garbage-cartel in New York (see below) used to allocate customers to trash haulers. Levenstein and Suslow (2011: 457, 475) report that eighty percent of the cartels in their sample allocated geographic markets or assigned specific customers to cartel members. Moreover, rules for geographic market allocation are used more frequently by international cartels, rather than the simple production quotas favored by domestic cartels.

Explicit Enforcement of Collusive Agreements by Information Exchange

Collusion can only be sustained if firms can detect deviations. If deviations cannot easily be

79 The case of demand-shocks is analyzed in greater detail below.

detected the firms must engage in more complex trigger strategies such as the one proposed by Green and Porter (1984; see the below subsection on collusion over the business cycle). In general, imperfect information causes collusion to be less effective. Firms can circumvent this problem by **exchanging information**, i.e. by colluding explicitly, on prices and quantities. This point is illustrated in greater detail in section H.3 .

The exchange of information on **past prices and quantities** facilitates collusion, as it allows to identify deviators and better target punishments, which then become more effective and less costly for the punishing firms. However, information exchange may also have effects that are beneficial to welfare such as serving demand more accurately or devising incentive schemes for firms' personnel based on their relative productivity.

The exchange of information on **current prices and quantities** allows firms to explicitly coordinate on the desired outcome. When the industry is subject to dynamic changes, the firms might otherwise miss this solution. Hence, information exchange on current prices and quantities is most likely to be detrimental to welfare.

Announcement of **future prices or production plans** might help collusion, in that it also might allow firms to better coordinate on a particular equilibrium. **Private announcements** of such information are directed only to competitors. It is hard to imagine any efficiency reason behind such announcements. Most likely, they merely help rivals to coordinate on a particular collusive price. **Public announcements** of, e.g., prices are seen by rival firms as well as consumers. On the one hand, this helps collusion as it facilitates the monitoring of deviations. On the other hand, it makes collusion more difficult because consumers can more easily shop around for the best offer.

Competition Authorities and Leniency

An extreme example of the illegality and criminal nature of cartels is the New York garbage cartel.

> *Between 1956 and 1995, trash haulers in New York were dominated by the Mafia, creating massive barriers to entry for other companies. Moreover, these trash haulers had agreed on a customer-allocation scheme that allowed them to charge prices which were approximately 40% above competitive levels, resulting in an absolute overcharge of $600 million annually.*
>
> *"The scheme works like this: Carters own customers' locations, known as "stops." Carters who "steal" stops (i.e., compete) are known as "outlaws." If you're an outlaw, you get four choices: Return the customer, keep the customer and pay a one-time multiple of 30 to 60 times monthly revenues to the previous carter, swap the*

customer for another, or do nothing and get your face batted in."[80] *The effectiveness of this customer-allocation scheme was also increased by the rule that trash haulers had to list their customers with New York City's Department of Consumer Affairs. This facilitated the monitoring of deviations from the collusive agreement.*

Unlike the vast majority of cartels the enforcement of this cartel was not restricted to trigger strategies and/or side payments only. The cartel was stabilized by the family ties among the firms as well as violent acts such as the torching of deviators' trucks. Enforcement in this particular cartel also included (threats of) physical violence to customers who solicited bids from other than their current trash haulers such as 'You will have your head split open if you continue'. Such threats caused some customers to stay with their trash hauler even after the conviction of the cartel.

Entry into this market was almost impossible. When Browing-Ferris Industries (BFI) entered this market, one of its executives found the severed head of a German shepherd on his lawn having a note in his mouth saying 'Welcome to New York'. BFI had massive trouble to attract customers even at savings of 30-60%. BFI then decided to collaborate with New York's district attorney in breaking up the cartel. In addition, the district attorney was approached by an "outlaw" carter that had come to him to complain about the cartel. This insider allowed the district attorney to collect data on the cartel for about one year.

Even in 1996 – three years after its entry and after the trials – BFI had not won more than 1% of the market. However, its entry had brought prices down to 10% of their level in 1991. Later in 1996 two of BFI's rivals entered into the New York-market, too.[81]

Although the physical violence associated with this cartel is a rare exception, this example makes clear that cartels can have massive effects on prices and welfare. Therefore, this subsection is concerned with presenting methods for detecting and breaking cartels up.

A competition authority may detect cartels by explicitly searching for them respectively by searching for evidence of market power. In this context, one should note that competition authorities usually have a hard time to assess the degree of market power by solely regarding the market-outcome. Therefore, evidence of market power is difficult to establish just from analyzing, for example, the **level of prices** (Motta 2004: ch. 4.4.1). This is for the following reasons. (i) Price-data

80 http://money.cnn.com/magazines/fortune/fortune_archive/1996/01/15/207168/index.htm

81 http://money.cnn.com/magazines/fortune/fortune_archive/1996/05/27/212867/index.htm

might not always be available. The available data often refers to list-prices rather than actual transaction-prices which are affected by negotiations among sellers and buyers and discounts. (ii) It is usually unclear what the monopoly-price and the competitive price in the industry are, and how the observed price relates to them. (iii) Even if this interval of prices was known there is no clear answer to the question what observed price-level may be termed *too high* and, thus, collusive. Rather than analyzing price levels, competition authorities might look at the **evolution of prices**. The plain idea is that price-movements should be quite similar in case of collusion. However, this idea does not take into account that common shocks, which affect all firms in an industry, such as shocks to input-prices, technology or demand would also cause a similar movement of prices even in a perfectly competitive situation.

If the competition authority suspects collusion in an industry, it may engage in **dawn raids** where the employees of a competition authority along with policemen search for evidence of collusion in the headquarters of firms suspected of collusive conduct. By thus, they seize documents which might help prove collusive agreements. To some extent, the mere possibility of surprise inspections suffices to destabilize cartels as is shown in the following.

Suppose, at the end of period t the competition authority detects an existing cartel with probability ρ. After discovery of the cartel, the cartel breaks down forever (with the firms making competitive profits π_i) and each cartel-firm must pay a fine F. Hence, the present value of profits after discovery by the competition authority becomes

$$V_{i,ca} = \frac{\pi_i}{1-\delta} - F \quad . \tag{137}$$

After a cartel has been established, firm i makes the collusive profit $\pi_{i,m}$ in the current period. In future periods, the firm makes the present value of profits $V_{i,ca}$ when the cartel is detected at the end of the current period, which occurs with probability ρ, and the present value of collusive profits $V_{i,m}$ when the cartel is not detected, which occurs with probability 1-ρ. Hence, the present value of collusive profits $V_{i,m}$ can be expressed as

$$V_{i,m} = \pi_{i,m} + \delta \cdot \left(\rho \cdot V_{i,ca} + (1-\rho) \cdot V_{i,m}\right) = \frac{\pi_{i,m} + \delta \cdot \rho \cdot \left(\pi_i / (1-\delta) - F\right)}{1 - \delta \cdot (1-\rho)} \quad . \tag{138}$$

It can be shown that the present value of collusive profits in presence of a competition authority (equation (138)) is smaller than in the absence of a competition authority (equation (129)). This makes collusion less profitable (in expectation) and, thus, **deters** the formation of new cartels.

The present value of deviation-profits remains

$$V_{i,d} = \pi_{i,d} + \frac{\delta \pi_i}{1-\delta} \quad .$$

The condition $V_{i,m} \geq V_{i,d}$ implies that the cartel is stable when condition (139) is satisfied.

$$\delta \geq \frac{\pi_{i,d} - \pi_{i,m}}{(1-\rho)\cdot(\pi_{i,d} - \pi_i) - \rho \cdot F} \equiv \bar{\delta}_{ca} \tag{139}$$

Let us compare the critical discount factor when a competition authority is present ($\bar{\delta}_{ca}$) to the critical discount without a competition authority ($\bar{\delta}$, ρ=0, F=0).

$$\delta \geq \frac{\pi_{1,d} - \pi_{1,m}}{\pi_{1,d} - \pi_1} \equiv \bar{\delta} \tag{133}$$

We find that $\bar{\delta}_{ca} > \bar{\delta}$ applies for $\rho > 0$[82] and $F > 0$. Hence, cartel-firms have a greater incentive to deviate from the cartel-agreement in the presence of a competition authority than in its absence. The intuition for this result is the following. The existence of a competition authority raises the probability that the cartel stops to exist in future periods which decreases the severity of the punishment following a deviation. The existence of a competition authority **destabilizes** existing cartels.

In addition to actively searching for collusive agreements. Competition authorities provide incentives for cartel-members who are unsatisfied with the agreement to reveal it to the authority (Pepall et al. 2008: ch.15.4 and 15.5). In this context one may, e.g., think of former employees of a cartel-firm or non-colluding competitors as in the case of the New York garbage cartel.[83] One such possibility are **leniency-programs** that many competition authorities have offered since the 1990s. The basic idea of such leniency-programs is to create disaccord between the cartel-members. In Europe, this is done by awarding cartel-firms full amnesty for disclosing a cartel and partial amnesty – i.e. a partial reduction of fines – for providing important information necessary to prosecute the cartel's members.

The incentive to "save" the fine may be an important driver for firms to reveal the collusive agreement especially when market-conditions or the nature of the collusive agreement change which makes collusion less profitable. This effect appears to be a main reason for leniency being one of the most effective tools of competition authorities in detecting collusive agreements. However, leniency programs also have an adverse effect because waiving the fine, i.e. setting F=0,

82 In order to further increase its effectiveness in detecting and prosecuting cartels, the German *Bundeskartellamt* established a 'special commission cartels' (*Sonderkommission Kartelle*) in 2002. (http://www.bundeskartellamt.de/wDeutsch/Kartellverbot/kartellverbotW3DnavidW2633.php)

83 http://select.nytimes.com/gst/abstract.html?res=F6061FF934590C708EDDAF0894DD494D81&scp=2&sq=to%20prosecutors%20breakthrough%20after%205%20years%20of%20scrutiny&st=cse

increases the present value of cartel profits (see equation (138)). By reducing the deterrence-effect of fines leniency programs may make the formation of new cartels more profitable.

The effect of competition authorities on the stability of cartels is substantial. Levenstein and Suslow (2011: 466, 479, 480) report that about 80 percent of the cartels in their sample ended with antitrust intervention. Moreover, they find a large, significant increase in the probability of breakup by antitrust enforcement in the post-1995 period, i.e. after competition authorities in Europe and USA had intensified their efforts to break up cartels. Antitrust enforcement is also found to raise the probability that cartels collapse on their own as is predicted by the above model.

Factors that Facilitate Collusion

In the following, we present some stylized industry and firm characteristics that facilitate collusion. These are not only relevant for the analysis of cartels but also for the analysis of collective dominance / coordinated effects in merger control (see section G and BKartA 2012: 34-43). Any factor that facilitates collusion must do one of two things. It must either reduce the critical probability-adjusted discount factor δ (see equation (133)), or it must reduce the likelihood of profitable cheating by cartel members.

Collusion is typically easier in more **highly concentrated industries**. To see this, consider the Bertrand-example from subsection H.1 and assume the number of the symmetric firms in this industry is n. When a firm i deviates from an existing collusive agreement, its deviation profit $\pi_{i,d}$ equals n times its profit in the cartel ($\pi_{i,m}$).

$$\pi_{i,d} = n \cdot \pi_{i,m}$$

This indicates that a firm's short-run gain from deviation is the higher the more firms are in this industry. Plugging the above condition into equation (133) for the critical discount factor yields

$$\delta \geq 1 - \frac{1}{n} \quad .$$

This shows clearly that a cartel is the more stable the fewer firms are active in this industry. Additionally, a larger number of firms also raises the efforts that need to be undertaken to coordinate the cartel and divide its spoils.

Levenstein and Suslow (2011: 470) find that the cartels in their sample occur predominantly in highly concentrated industries. Two-thirds of the cartels were in industries with a concentration ratio (CR_4) of 75 percent or more. However, their results on the effect of industry concentration on cartel stability do not support our above prediction of a destabilizing effect. They (*ibid.*, p. 479, 481) neither find a significant impact of concentration on the likelihood of detection by a competition

authority nor on cartel stability.

In section E.2 , we argue that firms can only exercise market power if the market is protected by **barriers to entry**, because the high cartel-profits would motivate other firms to enter the industry. This could mean two things. Either, the entrant could behave as a non-collusive maverick and steal market shares from the cartel by setting prices below the collusive price. Or, the entrant would be included in the cartel which means to divide the monopoly-profit by an additional member. In both cases, the cartel becomes less profitable for the previous cartel-members so that it becomes more likely that one of them deviates from the collusive agreement. Consequently, cartels may be assumed to be the more effective the lower the probability that other firms enter the cartelized market.

If a firm A has participation in a competitor B (**cross-ownership**), even without controlling it, the scope for collusion will be enhanced. First, if a representative in a firm is sitting in the board of directors of a rival firm, it will be easier to coordinate pricing and marketing policies. This makes it easy to monitor a rival's behavior and is an important facilitating factor for collusion. Second, if the profits of firm B enter the profit-function of firm A, firm A has an incentive to maximize joint profits which may imply a collusive outcome.

Frequent and regular orders facilitate collusion. With infrequent orders it takes longer to punish a firm that cheats on the cartel agreement, making cheating more attractive. Moreover, if the size of the orders varies, defection is particularly profitable in presence of temporarily large orders (see the below sub-section on demand shocks). More evenly sized orders reduce that risk of defection while a high frequency of orders helps collusion because it allows for a timely punishment.

Buyer power (see section E.2) is a further important factor. On the one hand, a purchasing agreement downstream may break the market power of an upstream-seller. On the other hand, such a purchasing agreement may also facilitate the coordination of these firms' selling activities. Buyer power can also be exercised by large customers of a cartel. On the one hand, the incentive to cheat is high when orders are large. On the other hand, collusion in the upstream market can be stabilized if a large downstream buyer negotiates a low price for himself but makes the cartel firms sell to his competitors at the elevated, collusive prices. Hence, the effect of buyer concentration or buyer power on the stability of a cartel is unclear. In fact, Levenstein and Suslow (2011: 481) do not find a statistically significant effect of buyer concentration on the stability of collusion.

The below discussion of collusion over the business cycle shows that cartels tend to be **more sustainable in growing markets** and more likely to be unstable in declining markets. In both cases, firms can earn a short-run deviation-profit by defecting, followed by eternal reversion to the

competitive equilibrium. In declining markets this punishment is relatively low because the sacrificed cartel-profits are low themselves. However, deviating in a growing market comes along with sacrificing even higher cartel-profits in the future. These points are illustrated in greater detail below.

Collusion is more likely to be sustainable when firms are **symmetric**. Symmetry can concern different dimensions such as market shares, number of varieties in the product portfolio, costs and technological knowledge, or capacities. When two firms have different costs it will be more difficult to formulate a collusive agreement that they both find satisfactory. Detailed negotiations over prices and market shares are much more straightforward when firms are similar. In addition, when firms distribute the cartel-profit asymmetrically, the firm with the lower share in profits usually finds it more profitable to deviate than the firm with the larger share, as the earlier can make a greater additional deviation-profit. To see this, suppose a Bertrand duopoly with monopoly-profits π_M. When the firms collude, firm 1 receives a share s_1 of π_M for whatever reason ($\pi_{1,m}=s_1\pi_M$). Correspondingly, firm 2 receives a share $s_2=1-s_1$. When either of the two firms deviates from the collusive agreement, it can make a deviation profit $\pi_{1,d}=\pi_{2,d}=\pi_M$. Using equation (133), we find the following stability-conditions for the two firms

$$\delta \geq 1 - s_i \quad . \tag{140}$$

It is easy to check that this condition supports the above proposition that the firm with the lower share finds it more profitable to deviate from the collusive agreement.

It is often claimed that **multi-market contact** facilitates collusion. This is not entirely obvious to see, because there are two effects of a deviation. Suppose two duopolists are colluding in two different markets, e.g., the markets for an identical product in Europe and USA. Now, a deviation in one market can be punished in two markets which stabilizes the agreement. However, defection in both markets poses a remarkable incentive for deviation. To see these effects more clearly, suppose that two Bertrand-duopolists collude in identical markets E and U with firm 1 possessing market shares $s_{1,E}=60\%$ and $s_{1,U}=40\%$. When regarding both markets separately the above incentive compatibility constraint (140) applies for both markets. Firm 1 anticipates that defection in just one market will result in punishment in both markets. Therefore, it does best to deviate in both markets and its joint incentive compatibility constraint becomes the average of the constraints in both individual markets.

$$\delta \geq 1 - \frac{s_{i,E} + s_{i,U}}{2} \tag{141}$$

Given the above asymmetric market shares, the cartel would be stable for $\delta \geq 0.5$. This makes collusion more stable as the stability condition in each separate market would be $\delta \geq 0.6$ (see

equation (140)).

Product homogeneity is generally believed by practitioners, competition authorities, and judges to facilitate collusion. The main argument underlying this proposition (and empirical finding) is that monitoring of deviations is easier in markets for homogeneous products because one only needs to observe a single price. In markets for differentiated products the price for each variant must be monitored. Besides this monitoring-effect the effects of product differentiation on cartel-stability are more complex. Recall that a cartel is stable when condition is satisfied.

$$\delta \geq \frac{\pi_{1,d} - \pi_{1,m}}{\pi_{1,d} - \pi_1} \equiv \bar{\delta} \tag{133}$$

Now, consider the effects of product differentiation on the critical discount factor $\bar{\delta}$ (Ross 1992). "As we introduce product differentiation we can see it having two opposing effects on cartel stability. First, the segmentation of the markets through differentiation means that the gains from cheating are less. Now a cheater cannot capture the entire market with a trivial price cut. This reduces the denominator in [(133)]. However, the Bertrand-Nash punishments are less severe with differentiation; the Nash equilibrium will leave all firms with positive profits in this case. With punishments less significant, the numerator in [(133)] falls as well" (Ross 1992: p. 4). The aggregate effect depends pretty much on consumers' reaction. Ross (1992) finds that greater product homogeneity (i) *decreases* cartel stability when products are rather differentiated (i.e., the deviation-effect dominates) and (ii) increases cartel stability when products are already relatively homogeneous (i.e., the punishment-effect dominates). Paha (2010) adds to this literature by allowing cartels to be incomplete, i.e. the cartel does not comprise all firms in an industry. He finds that greater homogeneity (i) increases cartel stability especially when the joint market share of the cartel-members is small and (ii) decreases cartel stability when the joint market share is large. In the latter case the cartelists are particularly effective in raising the price of their goods which makes deviation quite profitable.

The above discussion makes clear that **monitoring** is a very important element for the stability of collusive agreements. A cartel-firm's ability to shade its deviation quite effectively reduces its risk of being punished and, thus, makes it more likely to deviate.[84] When this effect is anticipated at the time of the cartel-formation it can even prevent the creation of a cartel, as deviations cannot effectively be punished. In this context, **most-favored-customer** (also referred to as meet-the-competition, or price-matching) **clauses** are likely to stabilize cartels. These clauses guarantee a customer that a firm will match any lower price offered by another seller. Therefore, customers would contribute to the stability of a cartel by reporting the deviator from a collusive

84 This monitoring-effect is considered in the EU-guidelines on the assessment of horizontal co-operation agreements in a chapter on **information exchange** (see section H.3).

agreement to the other cartel-members. Moreover, the deviator would not even receive the full deviation-profit because at least some customers would buy the product from its co-conspirators who meet the deviator's low price. Even in the absence of a cartel most-favored-customer clauses can be anti-competitive because they make it more costly for firms to give price discounts.

To summarize, cooperative price-fixing agreements are facilitated when an industry exhibits characteristics that make the detection and the deterrence of cheating easier. Such factors include the presence of only few firms, selling homogeneous products on a reasonably frequent basis and relatively stable market conditions.

Demand Shocks and Collusion over the Business Cycle

In the above cases, market demand is assumed to be the same in all periods. However in reality, demand can be subject to, for example, temporary shocks or movements over the business cycle. In the following, we analyze how such shocks affect the stability of cartels.

We start with the case of **temporary unobservable shocks to demand** (Green and Porter 1984). This model rests on the assumption that firms (i) can well observe the market price but (ii) do not observe demand, i.e. they have **imperfect information** about demand. (iii) Firms' decision variable is output and (iv) all firms supply a homogeneous product. The effect of this information structure can be seen in Figure 25. When demand is at the level D_1 the optimal cartel price would be $p_{1,m}$. However, at the lower demand D_2 the market-price at quantity q_m would be p_2.

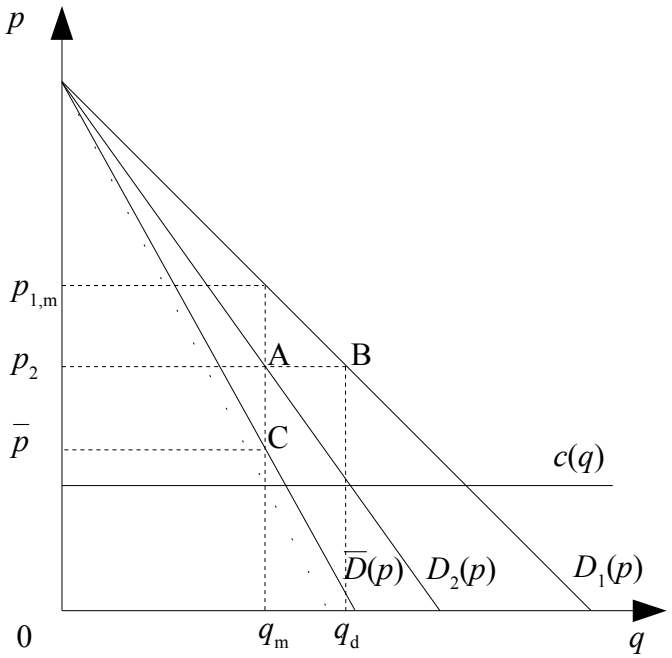


Figure 25: Green and Porter (1984)

If firms observe a price of p_2 under the above assumptions, they cannot discriminate if this lower price is a result of (i) a drop in demand (point A) or (ii) the deviation of at least one firm (point B) so that the cartel-output has risen to q_d while demand is still at the level of D_1. This situation is difficult for the following reasons. First, if the firms concluded that the lower price has been brought about by a deviation (point B) they could play a trigger strategy and revert to the competitive equilibrium forever. However, if in fact the low price is a result of a lower demand (point A) this means to sacrifice profits. Second, if the firms concluded that the lower price is due to a drop in demand while in fact it is a consequence of deviation, further deviations are likely to undermine the cartel.

Green and Porter (1984) propose that firms can solve this problem in the following way. The firms agree on a price $\bar{p}$ that is rather unlikely to be observed when all cartelists stick to the agreement and set q_m. Observing $\bar{p}$ and q_m at the same time (point C) would require a massive reduction of demand, which is rather unlikely. Hence, when the market-price drops below the threshold level $\bar{p}$ the cartel-firms assume a deviation from the collusive agreement and revert to the competitive equilibrium. However, there is still some chance that demand is really that low and punishment is unjustified. Therefore, the cartelists do not revert to the competitive equilibrium forever but for a predefined number of periods τ only. When these τ periods have passed, they revert back and set the collusive quantity q_m. Given a known distribution of demand-shocks the cartelists determine a combination of $\bar{p}$ and τ such that the firms are indifferent between colluding and deviating.

When temporary demand shocks cannot be observed by the firms, we find that the collusive equilibrium cannot be sustained when demand is low. Mapping prices over time would imply an evolution of prices where periods of high, collusive prices are followed by periods of low, competitive prices. The above discussion suggests that this **price-war** like pricing-pattern does *not* imply a breakdown and re-establishment of the cartel but is rather necessary to maintain the effectiveness of the cartel.

Now, suppose that **observable, temporary shocks to demand** occur (Rotemberg and Saloner 1986). In this case, "**oligopolies find implicit collusion [...] more difficult when their demand is relatively high** [*emphasis added*]. The reason for this is simple. When demand is relatively high and price is the strategic variable, the benefit to a single firm from undercutting the price that maximizes joint profits is larger. [...] On the other hand, the punishment from deviating is less affected by the state of demand if punishments are meted out in the future, and demand tends to return to its normal level" (Rotemberg and Saloner 1986: 390).

To see this, suppose that demand in an industry is subject to observable demand shocks ε_t

that last one period and shift the demand-curve $p(q,\varepsilon_t)$ up or down. All firms are assumed to produce a homogeneous product and compete in prices. Recall that in this case a cartel is stable as long as condition (132) is satisfied.

$$\pi_{i,d,t} - \pi_{i,m,t} \leq \frac{\delta(\pi_{i,m} - \pi_i)}{1-\delta} \tag{132}$$

Please note, that we have brought the time-indices back in. The right-hand side of equation (132) shows the present value of lost profits when a firm deviates in period t and is punished by eternal reversion to the Nash-equilibrium. Because future demand shocks are unknown with $E(\varepsilon_t)=0$, the right-hand side of equation (132) remains the same irrespective of the realization of ε_t in the current period t. However, the current profits rise is the level of ε_t, i.e.

$$\left(\frac{\partial \pi_{i,d,t}}{\partial \epsilon_t} > 0\right) \wedge \left(\frac{\partial \pi_{i,m,t}}{\partial \epsilon_t} > 0\right)$$

which simply means that a high level of demand implies a high level of profits. Similarly, it can be shown that the gain from deviation, i.e. the entire left-hand side of (132), rises in ε_t. This means that a **deviation is more likely when demand is high**.

In addition to this finding, Rotemberg and Saloner (1986) propose an additional strategy for preventing deviations from the collusive agreement. Suppose that ε^* is the maximum demand-shock where firms would just not want to deviate from the collusive agreement, i.e.

$$\pi^*_{i,d} - \pi^*_{i,m} \leq \frac{\delta(\pi_{i,m} - \pi_i)}{1-\delta} \quad .$$

Therefore, the firms set the monopoly-price whenever the observable demand shock is below or at this critical level ($\varepsilon_t \leq \varepsilon^*$). When the demand-shock is above this critical level ($\varepsilon_t > \varepsilon^*$), the firms simply reduce the collusive price to a level where stability-condition (132) is satisfied in equality. As a result, we would expect to see a reduction in prices when demand is high (**countercyclical pricing**). Therefore, observing countercyclical pricing in real industries might be considered evidence of collusion, because competitively behaving firms would price procyclical.

Haltiwanger and Harrington (1991) extend this model to **observable, correlated demand shocks** as one would observe in business cycles. Such demand is shown in Figure 26. This business cycle can be perfectly observed and anticipated by the firms. Because demand is the same in periods $\underline{t}$ and $\bar{t}$. Therefore, the short-run gain from deviation is the same at these two periods in the upturn and in the downturn. However, the expected punishment differs at these two points in time. Deviating at period $\underline{t}$ implies a relatively harsh punishment. This is because reverting to the competitive equilibrium in boom-periods means forgiving high collusive profits. However,

punishment in period $\bar{t}$ is much weaker. This is because reverting to the competitive equilibrium in periods of recession means sacrificing only low collusive profits. This would mean that **cartels are more likely to break down in downturns than in upturns**.

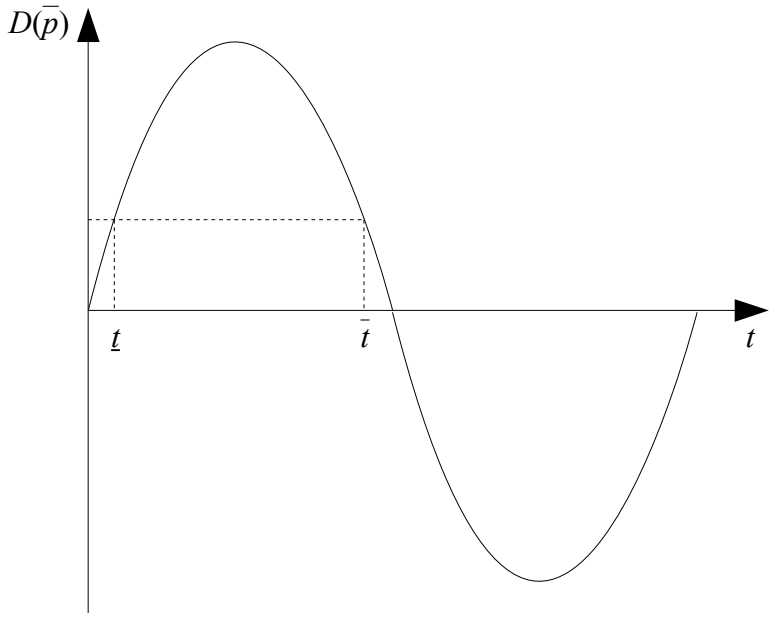


Figure 26: Haltiwanger and Harrington (1991)

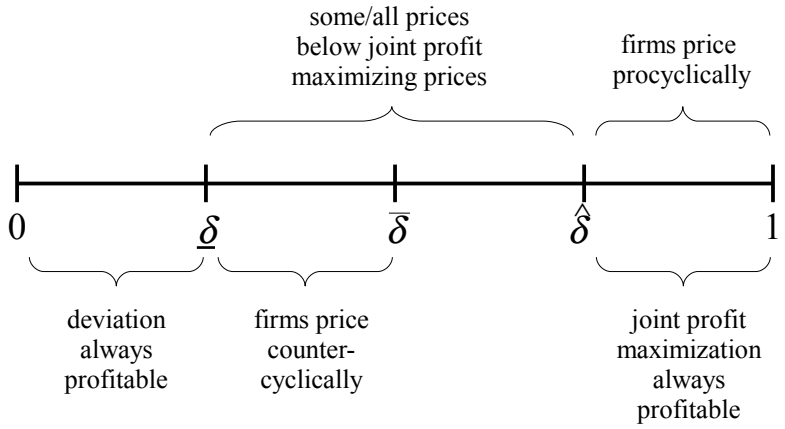


Figure 27: Discount factors

With regard to the cyclicality of the price-path, Haltiwanger and Harrington (1991) show the following. When firms are sufficiently patient (δ is above $\hat{\delta}$) collusion can always be sustained at the monopoly-price. Hence, prices moves procyclically. When firms are very impatient (δ is below $\underline{\delta}$) collusion can never be sustained and the resulting competitive price moves procyclically. When firms are rather patient ($\delta \in [\bar{\delta}; \hat{\delta}]$) collusion can be sustained only if the price is sometimes set below the monopoly-price. However, these price reductions are so faint that prices still move pro-cyclically. When firms are rather impatient ($\delta \in [\underline{\delta}; \bar{\delta}]$) collusion can only be sustained if prices are

often and markedly lowered below the monopoly-price. As a result, a countercyclical price-path emerges.

Summarizing we can say the following. The short-run gain from deviation is high when demand is high. The (expected) long-run loss from deviating depends on the nature of demand movements (demand-shocks vs. anticipated business cycles). Therefore, it is not always clear whether firms are more likely to defect from the collusive agreement when demand is high or low. Moreover, firms' information structure is an important determinant for their optimal punishment strategy.

Miscellaneous

The above discussion makes clear that the grim trigger strategy introduced in subsection H.1 is not necessarily optimal. Eternal reversion to the competitive equilibrium can be a too harsh punishment that even harms the punishing companies more than necessary. Motta (2004: ch. 4.3) provides an introduction to additional **punishment strategies** such as stick and carrot strategies. In this case, firms choose as punishment a market outcome that is characterized by below-competitive profits and return to collusion as quickly as possible.

Moreover, the assumption of eternal reversion to the one-shot Nash-equilibrium is not always realistic because firms can, theoretically, meet after a defection and **re-negotiate** the collusive agreement. The possibility of re-negotiations softens the intensity of a punishment and, thus, would make deviations more likely. However, in a changing business-environment with varying demand- and/or cost-conditions re-negotiations can be necessary to sustain collusion.

H.3 Horizontal Co-Operation Agreements (Art. 101 III TFEU)

This section is concerned with the assessment of Horizontal Co-Operation Agreements in European competition policy according to the guidelines on this topic (EC 2011). The focus of these guidelines is on outlining efficiencies that may reduce the anti-competitive effects of horizontal agreements that fall under Art. 101 I TFEU. Such efficiencies are assessed under Art. 101 III TFEU.

Horizontal Co-Operation Agreements

The relevant rules regarding the applicability of Art. 101 TFEU to **horizontal co-operation agreements** have recently been laid down in newly revised guidelines (EC 2011). In these guidelines the European Commission provides a framework for the assessment of the most common types of horizontal agreements taking into account their pro- and anti-competitive effects as explained below.

The term **horizontal** in the meaning of these guidelines refers to, first, agreements among actual or potential competitors. Second, it also refers to non-competitors if these are "active in the same product markets but in different geographic markets" (EC 2011: para. 1).

"Two companies are treated as **actual competitors** [*emphasis added*] if they are active on the same relevant market [see section F]. A company is treated as a **potential competitor** [*emphasis added*] of another company if, in the absence of the agreement, in case of a small but permanent increase in relative prices it is likely that the former, within a short period of time, would undertake the necessary additional investments or other necessary switching costs to enter the relevant market on which the latter is active" (EC 2011: para. 10).

The economic effects of horizontal co-operation agreements are fairly similar to those of **joint ventures**. The latter are agreements between competitors that create a new entity that carries out some activities instead of the partners. Such full-function joint ventures perform "on a lasting basis all the functions of an autonomous economic entity" (EC 2011: para. 6) and, thus, are subject to merger control in the European Union. "There is often only a fine line between full-function joint ventures that fall under the Merger Regulation and non-full-function joint ventures that are assessed under Article 101" (EC 2011: para. 21). In the following, we refer to non-full-function joint ventures as horizontal co-operation agreements.

The effects of both horizontal co-operation agreements and (full-function) joint ventures lie somwehere between those of cartels and mergers. On the one hand, a commercialization agreement whose only purpose is to set prices or quantities in the final market without any additional activity is nothing other than a cartel between the two competitors. On the other hand, consider two competitors that give all their research, production and sales assets in a certain sector to a newly created firm whose ownership they share. Since the partners cease any independent business in the sector, the joint venture is akin to a merger.[85]

Pro- and Anti-Competitive Effects, and Basic Principles for the Assessment of Horizontal Agreements

The **assessment under Article 101 TFEU** consists of two steps (EC 2011: para. 20).

1. One has to assess under article 101 I TFEU if an agreement between undertakings, which is capable of affecting trade between Member States, has an

   (i) **anti-competitive object** (so-called restrictions of competition by object (EC 2011: para.

85 In the interview linked below the former BMW chief-economist Helmut Becker comments on the effects of cooperation in the automobile-industry and stresses the similarity of joint ventures and mergers: http://www.manager-magazin.de/unternehmen/artikel/0,2828,688540,00.html

24)) or

(ii) actual or potential **restrictive effects on competition** (EC 2011: para. 26). "Restrictive effects on competition [...] are likely to occur where it can be expected with a reasonable degree of probability that, due to the agreement, the parties would be able to profitably raise prices or reduce output, product quality, product variety or innovation" (EC 2011: para. 28).

2. If the agreement is found to be restrictive of competition within the meaning of Art. 101 I TFEU, one has to determine the pro-competitive benefits produced by that agreement and to assess whether those pro-competitive effects outweigh the restrictive effects on competition. The balancing of restrictive and pro-competitive effects is conducted exclusively within the framework laid down by Article 101 III TFEU.

If the pro-competitive effects do not outweigh a restriction of competition, Article 101 II TFEU stipulates that the agreement shall be automatically void.

Horizontal co-operation agreements may have **anti-competitive effects** in several ways (EC 2011: paras. 33-47). Hence, the assessment of restrictions of competition by object requires paying attention to, e.g., the following points.

- The agreement may be exclusive in the sense that it **limits the possibility of the parties to compete** against each other or third parties. "This is, for example, the case if the parties agree to fix prices or output or to share markets, or if the co-operation enables the parties to maintain, gain or increase market power and thereby is likely to give rise to negative market effects with respect to prices, output, product quality, product variety or innovation" (EC 2011: para. 3).[86] Competitors can also benefit from the reduction of competitive pressure that results from the agreement and may therefore find it profitable to increase their prices (umbrella-effect). The reduction in those competitive constraints may lead to price increases in the relevant market.
- The agreement may require the parties to contribute such assets or affect the parties' financial interests in such a way that their **decision-making independence is appreciably reduced**.
- The agreement may also (i) lead to the disclosure of strategic information thereby increasing the likelihood of coordination among the parties or (ii) achieve significant commonality of costs, so that the parties may **more easily coordinate market prices and output** (see section H.2).

86 These points have been illustrated in greater detail in sections C.2 and H.2 above.

"Horizontal co-operation agreements can lead to substantial economic benefits, in particular if they combine complementary activities, skills or assets. Horizontal co-operation can be a means to share risk, save costs, increase investments, pool know-how, enhance product quality and variety, and launch innovation faster" (EC 2011: para. 2). The assessment of such **pro-competitive effects** is done according to the four cumulative conditions as formulated in Art. 101 III TFEU (EC 2011: para. 49). These conditions are all necessary conditions, so an agreement that fails anyone of them cannot be allowed.

(1) The agreement leads to **efficiency gains**, i.e. it must contribute to improving the production or distribution of products or contribute to promoting technical or economic progress.

(2) The restriction of competition must be **indispensable** to the attainment of these efficiency gains.

(3) Consumers must receive a fair share of the resulting benefits. The efficiency gain must be **passed on to consumers** so that they are at least compensated for the restrictive effects of the agreement.

(4) The agreement must **not** afford the parties the possibility of **eliminating competition** in respect of a substantial part of the products in question.

"There are a number of comments relating to these conditions that should be noted. First, the **efficiencies** [*emphasis added*] could be cost savings or they could be related to the quality of the products available (e.g. a better version of an existing product or a new product). [...] The Commission requires that any claimed efficiency must be clearly linked to the agreement, it must be verifiable and it must be quantifiable. These requirements are consistent with those imposed by the Commission in the assessment of efficiencies in horizontal mergers.

Secondly, the Commission has made clear that it considers a "fair share" for consumers as meaning that consumers are no worse off as a result of the agreement that they are without the agreement. So the requirement is that consumer welfare remains the same or is improved. This implies that it is not acceptable for consumers to be harmed but for producers to gain more benefit than the harm caused to consumers. That is, the Commission is clear that it is applying a **consumer welfare standard** [*emphasis added*], not a social welfare standard.

Thirdly, consumers can benefit either from lower prices or from a new or improved product. Where the efficiency is a cost efficiency, it is necessary for it to be **passed on to consumers** [*emphasis added*] in the guise of lower prices.

Fourthly, analysing whether competition has been eliminated in a substantial part of the market essentially requires a **market power analysis** [*emphasis added*] [...]. If the agreement is

unlikely to lead to the creation of market power, then it will not eliminate competition to any significant degree" (Bishop and Walker 2010: pp. 161-163). In the following, we analyze several types of horizontal agreements with respect to their pro- and anticompetitive effects.

Information Exchange

As a new issue, the guidelines contain a chapter on **information exchange**. "Information exchange can only be addressed under Article 101 if it establishes or is part of an agreement, a concerted practice or a decision by an association of undertakings" (EC 2011: para. 60). Principally, information exchange can take two forms. "Firstly, data can be **directly shared** [*emphasis added*] between competitors. Secondly, data can be **shared indirectly** [*emphasis added*] through a common agency (for example, a trade association) or a third party such as a market research organisation or through the companies' suppliers or retailers" (EC 2011: para. 55).

As an **anti-competitive effect**, the exchange of market information may facilitate collusive market outcomes (EC 2011: para. 65). "One way is that through information exchange companies may reach a **common understanding** [*emphasis added*] on the terms of coordination, which can lead to a collusive outcome on the market. [...] Another channel [...] is by **increasing the internal stability of a collusive outcome** [*emphasis added*] on the market. In particular, it can do so by enabling the companies involved to monitor deviations. [...] Both exchanges of present and past data can constitute such a monitoring mechanism." Moreover, information exchange may make the market sufficiently transparent so that the colluding companies can (i) more easily prevent the entry of new competitors into the market and (ii) put unaffiliated competitors at a competitive disadvantage (**anti-competitive foreclosure**) (EC 2011: paras. 66-69).

These anti-competitive effects are particularly likely if the parties exchange individualized intentions concerning **future conduct regarding prices or quantities**. The exchange of such information is considered a restriction of competition by object (EC 2011: paras. 73-74). The guidelines name the the following types of strategic information. It "can be related to prices (for example, actual prices, discounts, increases, reductions or rebates), customer lists, production costs, quantities, turnovers, sales, capacities, qualities, marketing plans, risks, investments, technologies and R&D programmes and their results. Generally, information related to prices and quantities is the most strategic, followed by information about costs and demand" (EC 2011: para. 86). The Commission (EC 2011: paras. 90-94) considers particularly the **frequent** exchange of **current or future**, **non-public** data the most critical.

As **pro-competitive effects**, the commission notices (EC 2011: para. 57) that information exchange may generate the below types of efficiency gains.

(1) Information exchange may solve problems of **information asymmetries**, thereby making markets more efficient. One example for such information exchange is the collection of data on risks in the insurance industry by industry organizations such as the GDV (*Gesamtverband der Deutschen Versicherungswirtschaft*) in Germany. By collecting data it facilitates insurers' task to calculate risks and charge adequate premiums.

In this context, the guidelines (EC 2011: para. 89) specify that collection "and publication of **aggregated market data** [*emphasis added*] (such as sales data, data on capacities or data on costs of inputs and components) by a trade organisation or market intelligence firm may benefit suppliers and customers alike by allowing them to get a clearer picture of the economic situation of a sector. [... T]he exchange of aggregated data is unlikely to give rise to restrictive effects on competition. Conversely, the exchange of **individualised data** [*emphasis added*] facilitates a common understanding on the market and punishment strategies by allowing the coordinating companies to single out a deviator or entrant. "

(2) Moreover, companies may improve their internal efficiency through **benchmarking** against each other's best practices.

(3) Sharing of information may also help companies to save costs by **reducing their inventories**, enabling quicker delivery of perishable products to consumers, or dealing with unstable demand etc.

Another beneficial effect relates to a reduction of deadweight loss in case of **cost-uncertainty**. To see this, suppose that a monopolist must determine a quantity for its good at the beginning of the period. At this point in time the firm does not know whether it incurs low marginal costs c_L or high marginal costs c_H. This situation is shown in Figure 28. The best the firm can do, is to set a quantity $\bar{q}$ which is optimal for the expected value of marginal costs $\bar{c}$.

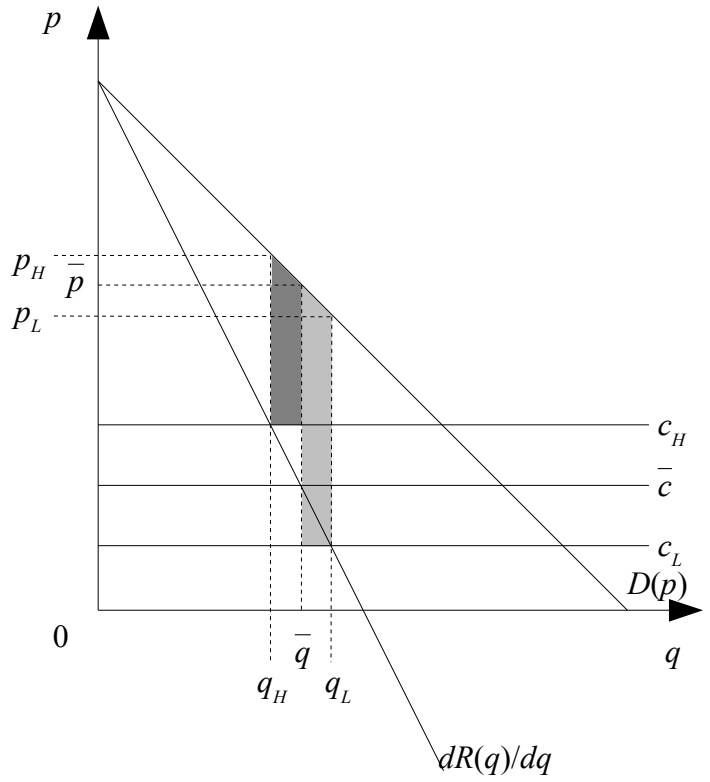


Figure 28: Welfare-Effects of Cost-Uncertainty

When the realization of marginal costs is c_L, setting $\bar{q}$ implies a deadweight-loss equaling the light-gray trapezoid. When the realization of marginal costs is c_H, setting $\bar{q}$ implies an increase in welfare equaling the dark-gray trapezoid. Consequently, firms' imperfect information can increase or decrease consumer surplus. Important is the assessment of the net-effect. In this example, the negative effect of imperfect information outweighs its positive effect. Hence, an exchange of information, which eliminates this imperfect information, has a positive effect on welfare.

(4) Furthermore, information exchanges may directly benefit consumers by r**educing their search costs and improving choice**.

These pro-competitive effects must be assessed along with the anti-competitive effects on a case-by-case basis. In this context, "it is important to assess the restrictive effects of the information exchange in the context of both the initial market conditions, and how the information exchange changes those conditions" (EC 2011: paras. 75-76). These conditions basically refer to the above factors that facilitate collusion (see section H.2).

A more complete treatment of the effects of information sharing is provided by Kühn and Vives (1994). They show that the welfare-effects of information sharing are far from clear-cut. This is because of three effects of information sharing.

(1) Information sharing will make the decisions of the firms more precisely geared to the environment. This typically raises firms' profits.

(2) The change in actions induced by information acquisition will have external effects on consumers in the market. The above monopoly-example shows that consumers sometimes benefit from firms' suboptimal behavior when information is imperfect. The net-effect of information sharing depends on the shape of demand and firms' mode of competition, i.e. if they compete in quantities or prices.

(3) Moreover, information acquisition by a firm may have external effects on other firms that compete in the same market.

Given these separate effects, the aggregate effect of information sharing on consumer surplus cannot be said to be clearly positive or negative.

To see the effect of information sharing in an oligopoly, consider the situation in Figure 29. Again, a firm (now a Cournot-oligopolist) faces some residual demand D and sets its output at the beginning of the period without knowing whether marginal costs will be high c_H or low c_L. Not knowing anything about the realization of costs, the firm would base its decision on mean residual demand $\overline{D}$ and optimally set quantity $\overline{q}$. It is possible to show that this firm is affected by the information of the other firms. To see this, suppose that the firm's competitors acquire information in order to perfectly predict the realization of marginal costs.

When marginal costs are low, the firm's rivals set a higher quantity so that residual demand drops to D_L with the market price being p_L. In comparison to a situation of full information, we can show the following:

1. The informed firms make higher profits and the uninformed firm lower profits than in a situation of perfect information about costs.
2. As the uninformed firm sets a lower than optimal quantity the market price under imperfect information is higher and consumer surplus is lower than under perfect information.

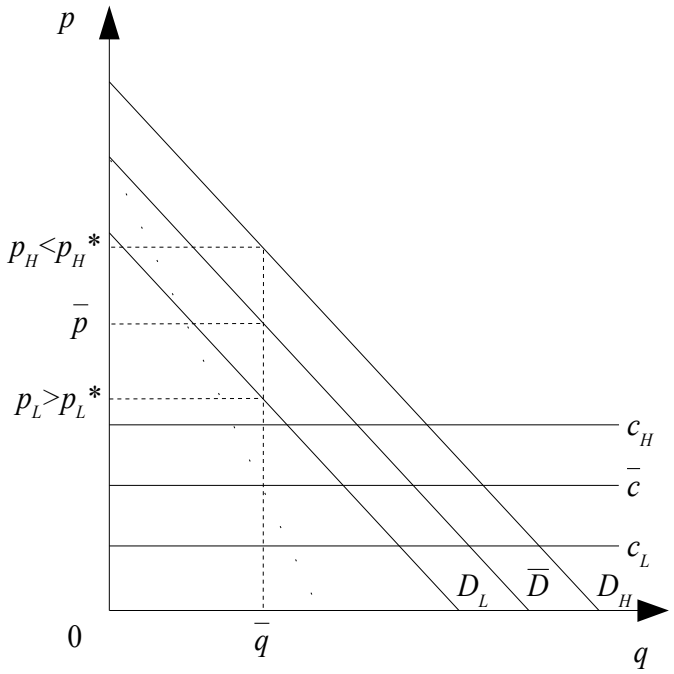


Figure 29: Profit-Effects of Asymmetric Information-Sharing

When marginal costs are high, the firm's rivals set a lower quantity so that residual demand increases to D_H with the market-price being p_H. In comparison to a situation of full information, we can show the following:

3. The uninformed firm sets a higher than optimal quantity which lowers the market price and benefits consumers.
4. This harms its competitors who make lower than optimal profits and benefits the uninformed firm who makes higher than optimal profits.

In total we find:

5. The uninformed firm's increase in profits when marginal costs are high is lower than its decrease in profits when marginal costs are low. Hence, information exchange benefits the firm.
6. Moreover, information exchange benefits the consumers because the decrease in consumer surplus when marginal costs are low exceeds the increase in consumer surplus when marginal costs are high.
7. However, information exchange is not in the interest of the informed firms as they would loose more profits in times of low marginal costs than they could gain in periods of high marginal costs. More intuitively, information exchange would eliminate their informational

advantage.

Research and Development Agreements

The EC's (2011: para. 141) guidelines on horizontal co-operation agreements specify that many "R&D agreements [...] bring about **efficiency gains** [*emphasis added*] by combining complementary skills and assets, thus resulting in improved or new products and technologies being developed and marketed more rapidly than would otherwise be the case."[87] The main **anti-competitive concerns** of R&D co-operation are the following. "First, it may reduce or slow down innovation, leading to fewer or worse products coming to the market later than they otherwise would. Secondly, on product or technology markets the R&D co-operation may reduce significantly competition between the parties outside the scope of the agreement or it may make anti-competitive coordination on those markets likely, thereby leading to higher prices" (EC 2011: para. 127). The earlier point refers to the inverted u-shaped relationship between industry-concentration and the incentive to innovate (see section C.2), i.e. moving form point A to B in Figure 30. The latter point refers to the effects of collusion as explained in section H.2 .

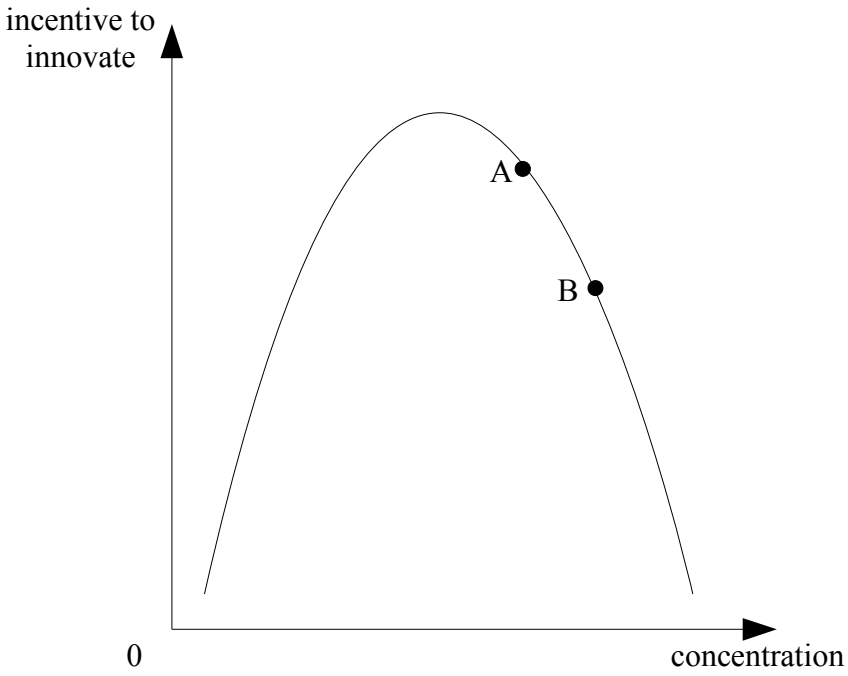


Figure 30: Inverted U-Shaped Relationship

The guidelines (2011: para. 112) specify that R&D may result in a product (or technology) which competes in an existing product (or technology) market (**existing product markets**) or in an entirely new product which creates its own **new product market**. This difference is important for

87 A detailed definition of R&D agreements in terms of European law is provided in Art. 1 of the R&D Block Exemption Regulation (EC 2010.a).

determining the relevant market. In either case, the relevant market is determined according to the principles outlined in section F .

When R&D agreements do not truly concern joint R&D but serve as a tool to engage in a **cartel** (aimed at price fixing, output limitation or market allocation) they are automatically void according to Art. 101 I TFEU (EC 2011: para. 128). In this context, the guidelines (EC 2011: para. 132) also note that "R&D co-operation which does not include the joint exploitation of possible results by means of licensing, production and/or marketing rarely gives rise to restrictive effects on competition within the meaning of" Art. 101 I TFEU. However, when there are pro- *and* anti-competitive effects, these must be weighed against each other according to Art. 101 III TFEU. This particularly affects the assessment of **competition in innovation** and of **competition in the output market**.

Regarding competition in innovation, the guidelines distinguish between the case where (i) innovation is structured in such a way that one can identify **R&D poles** and the case where (ii) innovative efforts are not clearly structured so that R&D poles cannot be identified. "Competing R&D poles are R&D efforts directed towards a certain new product or technology, and the substitutes for that R&D, that is to say, R&D aimed at developing substitutable products or technology for those developed by the co-operation and having similar timing" (EC 2011: para. 120). The first case (R&D poles can be identified) can, for example, be motivated by the pharmaceutical industry. In this case, it will be analyzed whether the cooperation leaves a sufficient number of R&D poles such that the competition in innovation is not negatively affected. In the second case (R&D poles cannot be identified), such an assessment cannot be performed.

"R&D agreements between competitors are covered by the **R&D Block Exemption Regulation** [*emphasis added*] provided that their combined market share does not exceed 25% and that the other conditions for the application of the R&D Block Exemption Regulation are fulfilled" (EC 2011: para. 134). This means that the above weighing of pro- and anti-competitive effects must only be performed if the agreement's combined market share exceeds 25%. When the combined market share is below 25%, it is presumed "that the positive effects of research and development agreements will outweigh any negative effects on competition" (EC 2010.a: Intro. para. 4). In this case, the R&D agreement is exempted from the application of Art. 101 I TFEU for seven years (EC 2010.a: Art. 4).

In the following, we present a model to illustrate some economic effects of research and development (R&D). R&D is often characterized by spillovers and non-rivalry that need to be taken into account when analyzing the effect of R&D joint ventures. **Spillovers** occur because technology and know-how often flow from one firm to another, for instance through imitation, reverse-

engineering and workers' mobility. This reduces the extent to which firms can appropriate the results of their R&D efforts, which in turn reduces their incentive to invest in R&D. R&D is also **non-rival** in the sense that it can be used by other parties without its value being diminished. Knowledge might be costly to create the first time, but once it exists, its diffusion does not modify its nature. In this sense, ex post one would like R&D to spread as much as possible in society in order to avoid duplication of R&D-costs.

We analyze the effects of R&D co-operations by means of a model for a homogeneous good (Pepall et al. 2008: ch. 22.5). We use the Cournot-model outlined on p. 48, with demand

$$Q = a - p \quad . \tag{42}$$

For simplicity, we restrict the number of firms to n=2. These firms are ex ante identical and produce the good at marginal costs c. In the output-market, the firms are assumed to compete á la Cournot. The firms' costs can be reduced by R&D (process innovation), but it is possible that the knowledge developed by one firm can spill over to its rival. Specifically, if firm 1 undertakes R&D at intensity x_1 and firm 2 undertakes R&D at intensity x_2, the marginal costs of the two firms become

$$\begin{aligned} c_1 &= c - x_1 - \beta\, x_2 \\ c_2 &= c - x_2 - \beta\, x_1 \end{aligned} \quad . \tag{142}$$

Here, $0 \leq \beta \leq 1$ measures the degree to which R&D activities spill over to the other firm (β=0: no spillovers; β=1: perfect spillovers). Research is costly and is assumed to exhibit diseconomies of scale

$$r(x_i) = x_i^2/2 \quad . \tag{143}$$

We solve this game by backwards induction, i.e. we start with determining the equilibrium in the output-market and, then, determine firms' optimal level of R&D when the firms do not cooperate. Using the results from the box on p. 49 the firms choose optimal quantities

$$\begin{aligned} q_1 &= \frac{a - 2c_1 + c_2}{3} \\ q_2 &= \frac{a - 2c_2 + c_1}{3} \end{aligned} \tag{144}$$

making profits

$$\begin{aligned} \pi_1 &= \left(\frac{a - 2c_1 + c_2}{3}\right)^2 - \frac{x_1^2}{2} \\ \pi_2 &= \left(\frac{a - 2c_2 + c_1}{3}\right)^2 - \frac{x_2^2}{2} \end{aligned} \quad . \tag{145}$$

Plugging (142) in (144) and (145) yields

$$\begin{aligned} q_{1,c} &= \frac{a-c+x_1(2-\beta)+x_2(2\beta-1)}{3} \\ q_{2,c} &= \frac{a-c+x_2(2-\beta)+x_1(2\beta-1)}{3} \end{aligned} \tag{146}$$

and

$$\begin{aligned} \pi_{1,c} &= \left(\frac{a-c+x_1(2-\beta)+x_2(2\beta-1)}{3}\right)^2 - \frac{x_1^2}{2} \\ \pi_{2,c} &= \left(\frac{a-c+x_2(2-\beta)+x_1(2\beta-1)}{3}\right)^2 - \frac{x_2^2}{2} \end{aligned} \tag{147}$$

.

Equation (146) indicates that the output of each firm is an increasing function of its own R&D expenditures x_i. By contrast, the effect of the rival's R&D effort on a firm's production can go either way. For example, consider firm 1. For $\beta>0.5$ the spillover from firm 2's R&D efforts reduces firm 1's costs so strongly, that it increases its output. For $\beta<0.5$ the spillover reduces firm 1's costs only weakly. However, the research gives firm 2 a competitive advantage over firm 1. Because of this increase in competition, firm 1 reduces its quantity. The same effect is present in the profit-equations (147).

Regarding the investment-stage, every firm chooses the level of R&D activity that maximizes its profit given the research effort of its rival. So we can identify the **research intensity reaction functions** for each firm by differentiating profit function (147) w.r.t. the research intensity.

$$\begin{aligned} x_1 &= \frac{2(2-\beta)[a-c+x_2(2\beta-1)]}{[9-2(2-\beta)^2]} \\ x_2 &= \frac{2(2-\beta)[a-c+x_1(2\beta-1)]}{[9-2(2-\beta)^2]} \end{aligned} \tag{148}$$

When research spillovers are low ($\beta<0.5$), the research intensity reaction functions for the two firms are downward sloping, indicating that the research expenditures of the two firms are **strategic substitutes** – more research reduces the amount of research by the other. The intuition is that in this case the increased research effort by one firm, primarily reduces its costs and so gives it a competitive advantage with respect to the other rival firm. By contrast, when research spillovers are high ($\beta>0.5$), the research intensity reaction functions are upward sloping, meaning that the research expenditures of the two firms are **strategic complements**, i.e. an increase in research intensity by one of the firms induces an increase in research intensity by the other.

The **Nash-equilibrium** is found at the intersection of the two response functions.

$$x_{1,c}=x_{2,c}=\frac{2(a-c)(2-\beta)}{9-2(2-\beta)(1+\beta)} \tag{149}$$

This is decreasing in β, implying that increased research spillovers decrease each firm's chosen research intensity. The intuition underlying this result is that the higher the spillover the more is either firm interested in free-riding on the other's research efforts rather than doing research itself. The solution for research effort x_i implies output levels and profits for the two firms of

$$q_{1,c}=q_{2,c}=\frac{3(a-c)}{9-2(2-\beta)(1+\beta)} \quad \text{and} \tag{150}$$

$$\pi_{1,c}=\pi_{2,c}=\frac{(a-c)^2[9-2(2-\beta)^2]}{[9-2(2-\beta)(1+\beta)]^2} \quad . \tag{151}$$

Now, consider the case of a **R&D co-operation**. This has two effects on the above model. First, the firms may choose their R&D-efforts such as to maximize their joint profits (see equation (147))

$$\begin{aligned}\pi_{1,RJV}+\pi_{2,RJV} \quad = \quad & \left(\frac{a-c+x_1(2-\beta)+x_2(2\beta-1)}{3}\right)^2-\frac{x_1^2}{2}+\ldots \\ & \ldots\left(\frac{a-c+x_2(2-\beta)+x_1(2\beta-1)}{3}\right)^2-\frac{x_2^2}{2}\end{aligned} \quad . \tag{152}$$

Second, they agree on sharing the outcome of their research efforts, i.e. β=1. Maximizing the joint profits (152) under this condition yields the following optimal research efforts, quantities, and profits

$$\begin{aligned} x_{1,RJV}=x_{2,RJV} &= 4(a-c) \\ q_{1,RJV}=q_{2,RJV} &= 3(a-c) \\ \pi_{1,RJV}=\pi_{2,RJV} &= 9(a-c)^2 \end{aligned} \quad . \tag{153}$$

It can easily be checked that research efforts, quantities, and profits exceed those in the competitive case (see equations (149)-(151)). Therefore, the research joint venture benefits the firms as well as the consumers. This is because of two effects. First, the joint venture internalizes the effect of one firm's research on the other firm's profit as reflected in the joint profit maximization. Second, the agreement on sharing the spoils of research (β=1) increases the effectiveness of research and, thus, prevents double research.[88]

To summarize, without co-operation in research the market alone is unlikely to give rise to socially optimal levels of research. Research co-operations might help to cope with these problems and create appropriate levels of R&D. If firms collaborate in research, they will share the cost of R&D, thereby increasing their incentives to invest; they will also have immediate access to R&D output, thereby increasing diffusion; finally they can coordinate their effort, thereby avoiding

88 Note that this model only yields plausible results (specially prices above zero) if c in the base situation without research is sufficiently large with regard to a.

duplication of investments.

Production Agreements

"**Production agreements** [*emphasis added*] vary in form and scope. They can provide that production is carried out by only one party or by two or more parties. Companies can produce jointly by way of a joint venture, that is to say, a jointly controlled company operating one or several production facilities or by looser forms of co-operation in production such as subcontracting agreements where one party (the 'contractor') entrusts to another party (the 'subcontractor') the production of a good" (EC 2011: para. 150).

Regarding the **main competition concerns** "[p]roduction agreements can lead to a direct limitation of competition between the parties [**explicit collusion**]. [...] Production agreements may also result in the coordination of the parties' competitive behaviour as suppliers leading to higher prices or reduced output, product quality, product variety or innovation [**tacit collusion**]. [...] Production agreements may furthermore lead to **anti-competitive foreclosure** [*emphasis added*] of third parties in a related market (for example, the downstream market relying on inputs from the market in which the production agreement takes place). For instance, by gaining enough market power, parties engaging in joint production in an upstream market may be able to raise the price of a key component for a market downstream. Thereby, they could use the joint production to raise the costs of their rivals downstream and, ultimately, force them off the market" (EC 2011: paras. 157-159).

"A production agreement is unlikely to lead to restrictive effects on competition if the parties to the agreement do not have **market power** [*emphasis added*] in the market on which a restriction of competition is assessed. [...] Companies are unlikely to have market power below a certain level of market share. [...] In any event, if the parties' combined **market share does not exceed 20%** [*emphasis added*] it is likely that the conditions of Article 101(3) are fulfilled. However, if the parties' combined market share exceeds 20%, the restrictive effects have to be analysed [...]. Generally, a production agreement is more likely to lead to restrictive effects on competition in a concentrated market than in a market which is not concentrated. Similarly, a production agreement in a concentrated market may increase the risk of a collusive outcome even if the parties only have a moderate combined market share. Even if the market shares of the parties to the agreement and the market concentration are high, the risks of restrictive effects on competition may still be low if the **market is dynamic** [*emphasis added*], that is to say, a market in which entry occurs and market positions change frequently" (EC 2011: paras. 168-173).

With regard to **efficiency gains** the guidelines specify that "[p]roduction agreements can be

pro-competitive if they provide efficiency gains in the form of **cost savings** or better production technologies. [...] Producing jointly can also help companies to **improve product quality** if they put together their complementary skills and know-how. Co-operation can also enable companies to **increase product variety** [...]. Efficiency gains attained by indispensable restrictions need to be **passed on to consumers** in the form of lower prices or better product quality or variety to an extent that outweighs the restrictive effects on competition. [...] If the parties to the production agreement achieve savings in their variable costs they are more likely to pass them on to consumers than if they reduce their fixed costs. Moreover, the higher the market power of the parties, the less likely they will pass on the efficiency gains to consumers to an extent that would outweigh the restrictive effects on competition" (EC 2011: paras. 183-187).

Purchasing Agreements

"Joint purchasing agreements usually aim at the creation of buying power which can lead to lower prices or better quality products or service for consumers. However, buying power may, under certain circumstances also give rise to competition concerns" (EC 2011: para. 194). Joint purchasing agreements may involve horizontal and vertical agreements. The effects of horizontal agreements have been analyzed in section E.2 . The effects of vertical agreements will be analyzed in section H.4 .

In section E.2 , we show that buyer power is detrimental to welfare if the sellers at the purchasing market act competitively. Moreover, the detrimental effects to welfare are the more severe when the purchasing agreement possesses a dominant position at the selling market. In this context, the European Commission (2011: paras. 201+204) notes that "[i]f the parties have a significant degree of market power [...] on the selling market [...], the lower purchase prices achieved by the joint purchasing arrangements are likely not to be passend on to consumers." Moreover, "joint purchasing arrangements are less likely to give rise to competition concerns when the parties do not have market power on the selling market". The European Commission (2011: para. 208) presumes that purchasing agreements do not possess market power if their joint market share at the purchasing and selling markets does not exceed 15% (safe harbor). In the remaining cases, the joint purchasing agreement must provide a detailed assessment of its effects on the market (EC 2011: para. 209).

In addition to these effects, joint purchasing agreements may facilitate collusion among the participating firms. This is because joint purchasing aligns each other firm's costs and, thus, facilitates the monitoring of their decisions at the output-market. This is important for detection deviations from collusive behavior. Moreover, "the implementation of a joint purchasing

arrangement may require the exchange of commercially sensitive information such as purchase prices and volumes. The exchange of such information may facilitate coordination with regard to sales prices and output and thus lead to a collusive outcome on the selling market" (EC 2011: para. 215).

Agreements on Commercialization

"Commercialisation agreements involve co-operation between competitors in the selling, distribution or promotion of their substitute products" (EC 2011: para. 225). This may, e.g., involve cooperation in advertising, market-sharing agreements leading to a collusive market-outcome or vertical agreements including joint production joint commercialization (see subsection H.4). Therefore, the main competition concerns (EC 2011: paras. 230-233) are (i) price fixing, (ii) output limitation, (iii) market allocation, and (iv) exchange of strategic information.

However, the guidelines (EC 2011: para. 240) specify that commercialization agreements will not raise competition concerns if the parties have a combined market share lower than 15%. In this case, they are unlikely to possess market power. Moreover, commercialisation "agreements can give rise to significant efficiency gains […] stemming from economies of scale or scope, especially for smaller producers" (EC 2011: para. 246).

> *One example for the exchange of strategic information in case of commercialization agreements is the German food industry[89]. Representatives of several food producers (Kraft Foods, Unilever, Dr. Oetker, Mars, and Henkel) had met for years to exchange information about selling agreements and negotiations with important retailers. This information exchange mainly concerned sweets, ice cream, ready-to-eat meals, frozen pizza, animal food, and detergents.*
>
> *From an economic point of view, this market can be characterized by an oligopoly upstream and an oligopoly downstream. The downstream retailers may, thus, be assumed to possess buyer power (see section E.2) that harms the food companies. The exchange of information might have been used to reduce the negative impact of retailers' buyer power on the profits of the food companies. The resulting increase in prices of the above goods is likely to have been passed on to final consumers. As a consequence, the food companies were fined in March 2011 with an aggregate fine of about 38m EUR.*

89 http://www.bundeskartellamt.de/wDeutsch/aktuelles/presse/2011_03_17.php

Standardization Agreements

"**Standardisation agreements** [*emphasis added*] have as their primary objective the definition of technical or quality requirements with which current or future products, production processes, services or methods may comply. Standardisation agreements can cover various issues, such as standardisation of different grades or sizes of a particular product or technical specifications in product or services markets where compatibility and interoperability with other products or systems is essential. The terms of access to a particular quality mark or for approval by a regulatory body can also be regarded as a standard. Agreements setting out standards on the environmental performance of products or production processes are also covered by this chapter. [...] In certain industries companies use **standard terms** [*emphasis added*] and conditions of sale or purchase elaborated by a trade association or directly by the competing companies [...]. When such standard terms are widely used within an industry, the conditions of purchase or sale used in the industry may become de facto aligned. Examples of industries in which standard terms play an important role are the banking (for example, bank account terms) and insurance sectors" (EC 2011: paras. 257-260).

"Standardisation agreements may produce their effects on four possible markets, which will be defined according to the Market Definition Notice. First, standard-setting may have an impact on the **product or service market** or markets to which the standard or standards relates. Second, where the standard-setting involves the selection of technology and where the rights to intellectual property are marketed separately from the products to which they relate, the standard can have effects on the relevant **technology market**. Third, the **market for standard-setting** may be affected if different standard-setting bodies or agreements exist. Fourth, where relevant, a distinct **market for testing and certification** may be affected by standard-setting" (EC 2011: para. 261).

Concerning the **main competition concerns** of standardization agreements the guidelines specify that "[s]tandardisation agreements usually produce significant **positive economic** effects [*emphasis added*], for example by promoting economic interpenetration on the internal market and encouraging the development of new and improved products or markets and improved supply conditions. Standards thus normally increase competition and lower output and sales costs, benefiting economies as a whole. Standards may maintain and enhance quality, provide information and ensure interoperability and compatibility (thus increasing value for consumers). Standard-setting can, however, in specific circumstances, also give rise to **restrictive effects on competition** [*emphasis added*] by potentially restricting price competition and limiting or controlling production, markets, innovation or technical development. This can occur through three main channels, namely reduction in price competition, foreclosure of innovative technologies and exclusion of, or

discrimination against, certain companies by prevention of effective access to the standard.

First, if companies were to engage in anti-competitive discussions in the context of standard-setting, this could reduce or **eliminate price competition** [*emphasis added*] in the markets concerned, thereby facilitating a collusive outcome on the market.

Second, standards that set detailed technical specifications for a product or service may **limit technical development and innovation** [*emphasis added*]. While a standard is being developed, alternative technologies can compete for inclusion in the standard. Once one technology has been chosen and the standard has been set, competing technologies and companies may face a barrier to entry and may potentially be excluded from the market. In addition, standards requiring that a particular technology is used exclusively for a standard or preventing the development of other technologies by obliging the members of the standard-setting organisation to exclusively use a particular standard, may lead to the same effect. The risk of limitation of innovation is increased if one or more companies are unjustifiably excluded from the standard-setting process.

[...] Third, standardisation may lead to anti-competitive results by **preventing certain companies from obtaining effective access to the results of the standard-setting process** [*emphasis added*] (that is to say, the specification and/or the essential IPR for implementing the standard). [...] **Intellectual property laws** [*emphasis added*] and competition laws share the same objectives of promoting innovation and enhancing consumer welfare. IPR promote dynamic competition by encouraging undertakings to invest in developing new or improved products and processes. IPR are therefore in general pro-competitive. However, by virtue of its IPR, a participant holding IPR essential for implementing the standard, could, in the specific context of standard-setting, also acquire control over the use of a standard. When the **standard constitutes a barrier to entry** [*emphasis added*], the company could thereby control the product or service market to which the standard relates. This in turn could allow companies to behave in anti-competitive ways, for example by 'holding-up' users after the adoption of the standard either by refusing to license the necessary IPR or by extracting excess rents by way of excessive royalty fees thereby preventing effective access to the standard. However, even if the establishment of a standard can create or increase the market power of IPR holders possessing IPR essential to the standard, there is no presumption that holding or exercising IPR essential to a standard equates to the possession or exercise of market power. The question of market power can only be assessed on a case by case basis" (EC 2011: paras. 263-269).

"**Standard terms** [*emphasis added*] can give rise to restrictive effects on competition by **limiting product choice and innovation** [*emphasis added*]. If a large part of an industry adopts the standard terms and chooses not to deviate from them in individual cases [...], customers might have

no option other than to accept the conditions in the standard terms. [... T]here is a serious risk that standard terms relating to price would **restrict price competition** [*emphasis added*]. Moreover, if the standard terms become industry practice, access to them might be vital for entry into the market. In such cases, refusing access to the standard terms could risk causing **anti-competitive foreclosure** [*emphasis added*]" (EC 2011: paras. 270-271).

"Where participation in standard-setting is **unrestricted** and the procedure for adopting the standard in question is **transparent**, standardisation agreements which contain **no obligation to comply** with the standard and provide **access to the standard on fair, reasonable and non-discriminatory terms** [FRAND commitment] will normally not restrict competition within the meaning of Article 101(1)" (EC 2011: para. 280).

"Whether standardisation agreements may give rise to restrictive effects on competition may depend on whether the members of a standard-setting organisation remain **free to develop alternative standards or products** that do not comply with the agreed standard [...]. The assessment whether the agreement restricts competition will also focus on **access to the standard**. Where the result of a standard [...] is not at all accessible, or only accessible on discriminatory terms, for members or third parties [...] this may discriminate or foreclose or segment markets according to their geographic scope of application and thereby is likely to restrict competition. However, in the case of several competing standards or in the case of effective competition between the standardised solution and non-standardised solution, a limitation of access may not produce restrictive effects on competition. If **participation in the standard-setting process** is open in the sense that it allows all competitors (and/or stakeholders) in the market affected by the standard to take part in choosing and elaborating the standard, this will lower the risks of a likely restrictive effect on competition by not excluding certain companies from the ability to influence the choice and elaboration of the standard [...]. In certain situations the potential negative effects of restricted participation may be removed or at least lessened by ensuring that stakeholders are **kept informed and consulted** on the work in progress [...]" (EC 2011: paras. 292-299).

"As long as participation in the actual establishment of **standard terms** [*emphasis added*] is **unrestricted** for the competitors in the relevant market [...], and the established standard terms are **non-binding** and **effectively accessible** for anyone, such agreements are not likely to give rise to restrictive effects on competition. [...] Firstly, standard terms for the sale of consumer goods or services where the standard terms define the scope of the product sold to the customer, and where therefore the risk of limiting product choice is more significant, could give rise to restrictive effects on competition within the meaning of Article 101(1) where their common application is likely to result in a de facto alignment. This could be the case when the widespread use of the standard terms

de facto leads to a limitation of innovation and product variety. [...] Secondly, even if the standard terms do not define the actual scope of the end-product they might be a decisive part of the transaction with the customer for other reasons. An example would be online shopping where customer confidence is essential (for example, in the use of safe payment systems, a proper description of the products, clear and transparent pricing rules, flexibility of the return policy, etc). As it is difficult for customers to make a clear assessment of all those elements, they tend to favour widespread practices and standard terms regarding those elements could therefore become a de facto standard with which companies would need to comply to sell in the market. Even though non-binding, those standard terms would become a de facto standard, the effects of which are very close to a binding standard and need to be analysed accordingly" [EC 2011: paras. 301-307].

"**Standardisation agreements** [*emphasis added*] frequently give rise to significant **efficiency gains** [*emphasis added*]. For example, Union wide standards may facilitate market integration and allow companies to market their goods and services in all Member States, leading to increased consumer choice and decreasing prices. Standards which establish technical interoperability and compatibility often encourage competition on the merits between technologies from different companies and help prevent lock-in to one particular supplier. Furthermore, standards may reduce transaction costs for sellers and buyers. Standards on, for instance, quality, safety and environmental aspects of a product may also facilitate consumer choice and can lead to increased product quality. Standards also play an important role for innovation. They can reduce the time it takes to bring a new technology to the market and facilitate innovation by allowing companies to build on top of agreed solutions. [...] The use of **standard terms** [*emphasis added*] can entail economic benefits such as making it easier for customers to compare the conditions offered and thus facilitate switching between companies. Standard terms might also lead to efficiency gains in the form of savings in transaction costs and, in certain sectors (in particular where the contracts are of a complex legal structure), facilitate entry. Standard terms may also increase legal certainty for the contract parties" (EC 2011: paras. 308 and 312).

Block Exemptions

The Commission may adopt block exemption regulations (BER) which define certain categories of agreements that generally fulfill the conditions of Article 101 III TFEU. Such BERs may be criticized as putting too much weight on the form of an agreement in order to fall within the scope of a certain BER rather than concentrating on the economic effects of the agreements (*per se* rules vs. rule of reason). "In recognition of the inadequacies of the form-based approach to block exemptions, the Commission has implemented over the last decade a move towards an economics-based approach that focuses on market power and market outcomes. In 1999, the Commission

implemented a block exemption covering vertical agreements together with Guidelines that set out the principles for their assessment under Article [101]. Under this block exemption, vertical agreements entered into by firms with market shares below 30 per cent gain automatic exemption form the Article [101] prohibition. A similar market share threshold approach is used in the guidelines on horizontal agreements. The use of market share thresholds is intended to eliminate the regulatory burden of notifying agreements from those firms that are unable to act anti-competitively" (Bishop and Walker 2010: p. 159).

H.4 Vertical Restraints

"**Vertical agreements** [*emphasis added*] are agreements between firms at different levels in the production and supply chain and include agreements between manufacturers and retailers, manufacturers and distributors, distributors and retailers and so on. [A more detailed definition of vertical relationships is provided in section G.3 .] Vertical agreements in general contain restrictions imposed by one party on another. On occasion, these restrictions can fall foul of Article [101].

There are a wide variety of vertical restraints employed by firms that may or may not give rise to competition concerns. Some manufacturers distribute their products to selected outlets only ("**selective distribution**"). This is typically the case with branded products where the manufacturer is concerned with the environment in which its product is sold. Some retailers sell the products of only one manufacturer ("**exclusive dealership**"). Some retailers are given a guarantee by the manufacturer that no other retailers within their geographic area will be supplied by that manufacturer ("**exclusive territories**"). On occasion, manufacturers insist that their product is sold for a certain minimum amount ("**resale price maintenance**"). Other manufacturers may insist that the retailer sells a certain minimum amount of their product ("**quantity forcing**") [*emphases added*].

Many of these restrictions (resale price maintenance, quantity forcing, exclusive territories and exclusive dealership) are substitutes, albeit in some cases imperfect ones, for **vertical integration** [*emphasis added*]. For example, selective distribution is only a partial substitute for vertical integration. Both vertical integration and selective distribution lead to a restriction in the number of retailers being supplied with the product. However, whereas a vertically integrated firm might choose not to retail the products of its rivals, under selective distribution retailers would be free to select products from a number of different manufacturers" (Bishop and Walker 2010: pp. 187-189). The above treatment of vertical integration (see section G.3) suggests that vertical agreements can have pro-competitive (by eliminating double marginalization) or anti-competitive effects (by raising concerns of foreclosure). The principles for the assessment of vertical agreements

under Article 101 I TFEU are laid down in the EU Commission's 2010 **Guidelines on Vertical Restraints**[90].

The following discussion shows that vertical restraints can have both strong pro-competitive and anti-competitive effects. These can hardly be stated as general rules but depend strongly on supply and demand conditions both upstream and downstream. Therefore, vertical restraints should be assessed following a **rule of reason** approach rather than applying per se rules. It should be noted that vertical restraints are generally less harmful than horizontal restraints which is also recognized in the EU's Guidelines on Vertical Restraints (para. 98).

The Scope of Art. 101 I TFEU

The Guidelines on Vertical Restraints specify in para. 9 that "vertical agreements entered into by non-competing undertakings whose individual market share on the relevant market does not exceed 15% are generally considered to fall outside the scope of" Art. 101 I TFEU. The same applies for vertical agreements between **small and medium-sized enterprises** unless those collectively hold a dominant position in the relevant market (para. 11). **Agency agreements** and **subcontracting agreements** also fall outside the scope of Art. 101 I TFEU.

Further exemptions are provided by the Commission's **Block Exemption Regulation** concerning the application of Art. 101 III TFEU.[91] "Under Article 3 of the Block Exemption Regulation, the market share of both the supplier and the buyer are decisive to determine if the block exemption applies. In order for the block exemption to apply, the **market share** [*emphasis added*] of the supplier on the market where it sells the contract products to the buyer, and the market share of the buyer on the market where it purchases the contract products, must each be **30% or less** [*emphasis added*]" (Guidelines on Vertical Restraints, para. 87).

Enforcement Policy in Individual Cases

Analyzing vertical restraints under the enforcement policy of the EU, first, requires determining whether the vertical agreement falls within the scope of Art. 101 I TFEU and whether the block exemption does not apply. Second, it shall be individually assessed whether the conditions of Art. 101 III TFEU are satisfied (Guidelines on Vertical Restraints, para. 96).

"The assessment of whether a vertical agreement has the effect of restricting competition will be made by **comparing the actual or likely future situation** on the relevant market with the vertical restraints in place **with the situation that would prevail in the absence of the vertical**

90 http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52010SC0411:EN:NOT

91 http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32010R0330:EN:NOT

restraints in the agreement. In the assessment of individual cases, the Commission will take, as appropriate, both actual and likely effects into account. For vertical agreements to be restrictive of competition by effect they must affect actual or potential competition to such an extent that on the relevant market negative effects on prices, output, innovation, or the variety or quality of goods and services can be expected with a reasonable degree of probability. The likely negative effects on competition must be appreciable. Appreciable anticompetitive effects are likely to occur when **at least one of the parties has or obtains some degree of market power** and the agreement contributes to the creation, maintenance or strengthening of that market power or allows the parties to exploit such market power. Market power is the ability to maintain prices above competitive levels or to maintain output in terms of product quantities, product quality and variety or innovation below competitive levels for a not insignificant period of time. The degree of market power normally required for a finding of an infringement under Article 101(1) is less than the degree of market power required for a finding of dominance under Article 102 [*emphases added*]" (Guidelines on Vertical Restraints, para. 97).

The EU Guidelines on Vertical Restraints (para. 100) list the following **anti-competitive effects** of vertical restraints that are discussed in greater detail below:

- **Anticompetitive foreclosure** of other suppliers or other buyers by raising barriers to entry or expansion
- **Softening of competition** between the supplier and its competitors and/or **facilitation of collusion** amongst these suppliers, often referred to as reduction of **inter-brand competition**
- Softening of competition between the buyer and its competitors and/or facilitation of collusion amongst these competitors, often referred to as reduction of **intra-brand competition** if it concerns distributors' competition on the basis of the brand or product of the same supplier
- The creation of **obstacles to market integration**

In this context, the EU stresses that "[e]xclusive arrangements are generally more anti-competitive than non-exclusive arrangements. [Moreover, v]ertical restraints agreed for non-branded goods and services are in general less harmful than restraints affecting the distribution of branded goods and services. In general, a combination of vertical restraints aggravates their individual negative effects" (Guidelines on Vertical Restraints, paras. 103-105).

The EU Guidelines on Vertical Restraints (paras. 106-107) list the following **pro-competitive effects** of vertical restraints that are discussed in greater detail below:

- **Reduction of free-rider problems** such as the free-riding on others' promotional activities or their pre-sales services that are free for the buyers
- Vertical agreements may be necessary
  - Such that first time investments are made that are required when **opening up or entering a new market**
  - To guarantee that a product is sold by retailers only who provide the adequate service or only sell other high quality products (**certification free-rider issue**)
  - Such that client-specific investments (e.g. for special equipment or training) is made (**hold-up problem**)

In the context of free-riding, vertical agreements can help to prevent under-investment only if, first, the investment is relationship-specific, i.e. it cannot be used after termination of the agreement. Second, it is a long-term investment that cannot be recouped in the short term. Third, the investment is asymmetric insofar as one party to the contract invests more than the other. Additionally, vertical agreements may help to solve the following issues:

- The problem of **double marginalization** is reduced or eliminated.
- Concentration of sales to a limited number of retailers allows to exploit **economies of scale in distribution**.
- A retailer, whose quality as a borrower is unkown to a bank, may receive a loan more easily if it engages in a vertical agreement with a high-quality manufacturer so that **capital market imperfections** due to this imperfect information can be overcome.
- Vertical agreements, e.g. in the form of franchising agreements, may contribute to ensure **uniformity and quality standardization** of the retailers and, thus, create a brand image.

The assessment of the pro-competitive effects of vertical agreements is done on basis of the framework provided in **Art. 101 III TFEU**. Efficiencies are only noted as such if the vertical agreement is **indispensable** for their attainment. “The condition that consumers must receive a fair share of the benefits implies that consumers of the products purchased and/or (re)sold under the vertical agreement must at least be compensated for the negative effects of the agreement. In other words, the efficiency gains must **fully off-set the likely negative impact** [*emphasis added*] on prices, output and other relevant factors caused by the agreement” (Guidelines on Vertical Restraints, para. 126).

Intra-Brand Competition: Double Marginalization

The term **intra-brand competition** refers to the relationship between firms which produce and distribute the same product or brand. In the following, we will see that vertical agreements are often pro-competitive when regarded from the viewpoint of intra-brand competition.

To see the pro-competitive effects of vertical agreements, recall the above model with a monopolist supplier upstream and a monopolist retailer downstream (see section G.3). Without the vertical agreement these two firms face a **double marginalization** problem and set the following price and quantity in the downstream market (see equations (108), (109), and (111)).

$$\begin{aligned} Q_D &= (a-c)/4 \\ p_D &= (3a+c)/4 \end{aligned} \tag{154}$$

What a vertical agreement can do is eliminate the double marginalization problem resulting in a higher downstream quantity and a lower downstream price.

$$\begin{aligned} Q_V &= (a-c)/2 \\ p_V &= (a+c)/2 \end{aligned} \tag{113}$$

This benefits the consumers as well as the firms.

There are a number of vertical restraints that could be used in this case (see Bishop and Walker (2010: p. 192) and Motta (2004: p. 308 and 311-312)). For instance, the upstream firm could impose the resale price on the retailer (**resale price maintenance**, price ceiling, price cap). Imposing the retail price $p=p_V=(a+c)/2$ on the downstream firm will maximize the surplus of the vertical structure. The way in which the upstream and downstream firms share the surplus will then be determined by the wholesale price r. If the upstream firm has all the bargaining power, then it will fix $r=p_V=(a+c)/2$ and will get all the producer surplus. More generally, the higher r (with $r \in [c,p_V]$) the higher the share of the surplus going to the upstream firm. An identical outcome would be the upstream firm sets a **price ceiling** $\bar{p}=p_V=(a+c)/2$. This obliges the downstream firm to sell at a price $p \leq \bar{p}$. For any wholesale price $r \in [c,p_V]$ the downstream firm would then choose precisely $p=\bar{p}$ (and again the actual r determines the division of the surplus).

Alternatively, the manufacturer could impose the restraint that the downstream firm had to buy a certain given number of units from the upstream firm (**quantity forcing**, **quantity fixing**) where the number of units was set equal to the level $Q_V=(a-c)/2$ that would push the retail price p_D down to the downstream firm's marginal cost r. As in the previous case, the level of the wholesale price determines the distribution of the producer surplus. If the upstream firm has all the bargaining power, it will choose $r=p_V$ and appropriate all the profit of the vertical structure (Motta 2004: p. 312).

Yet another possibility to restore the vertically integrated outcome is for the manufacturer to use **non-linear pricing** (a fixed component T plus a variable component r for each unit bought) in order to make the retailer the *residual claimant* of all the profit generated in the market. By setting the variable component identical to the manufacturer's own cost, $r=c$, the retailer would face the maximization problem

$$\pi_D=(a-Q-c)\cdot Q-T \quad . \tag{155}$$

Therefore, it would effectively behave in the same way as a vertically integrated firm, and would choose the optimal final price $p_V=(a+c)/2$. The retailer would then make the maximum profit. However, part of or all such profit can be appropriated by the manufacturer through the **franchise fee** T. In general, the distribution of the profit depends on the relative bargaining power of the two firms. If it is the manufacturer who has all the bargaining power the manufacturer can make exactly the same profit as if it owned the retailer (Motta 2004: p. 308 and 312).

Note that despite the removal of double marginalization prices will not drop to the price level that would prevail if the upstream and the downstream market were perfectly competitive ($p=c$). The vertical agreement only removes the effects of market power in the downstream market but retains the market power in the upstream market. Hence, the market outcome (113) is the same as in the case of a vertically integrated monopolist. “This implies that when assessing the competitive effect of a vertical restraint, it is important to understand the nature of competition at each vertical level” [Bishop and Walker 2010: p. 193].

Intra-Brand Competition: Provision of Services and Other Efficiency Reasons for Vertical Restraints

The above example shows that in vertical relationships situations can arise where the best interest of the retailer is not in the best interest of the manufacturer and, in particular, of the society in general. Many of these instances are the result of the retailer not being able to appropriate all of the benefits of investment that he undertakes in his store. This leads the retailer to under-invest relative to the level that he would invest if he could appropriate all of the benefits of the investment. This **under-investment** might be harmful to the manufacturer, and to society, particularly if it leads to a lower demand for the products.

A classic example of this type of inefficiency arises when consumers value **pre-sa1es service** in a shop (see Bishop and Walker (2010: p. 197) and Motta (2004: p. 314)). Note that now the above model of vertically related industries is extended to multiple retailers downstream. Pre-sa1es service can increase the demand for some products, particularly if they are rather complex goods whose quality is not immediately apparent (e.g. expensive consumer electronics).

Manufacturers want retailers to offer pre-sales service, such as knowledgeable sales staff, as it increases the demand for their products. However, if most retailers offer pre-sales service, it is profitable for some retailers to offer no pre-sales service and to free-ride on the service (i.e. the positive externality) offered by other retailers. Consumers can get pre-sales service at one retailer and then buy the product at another retailer that is able to offer a lower price because it does not offer any pre-sales service and so does not incur the cost of offering pre-sales service. The problem for the manufacturer is that this type of behavior reduces the incentive of any retailers to offer pre-sales service. The result is that too little pre-sales service, from the manufacturer's and society's perspective, is offered.

Vertical restraints might restore incentives for the retailers to invest in services. For instance, suppose that the producer divides the city in different areas, and appoints an exclusive distributor in each area (**exclusive territories**). This would reduce the possibility of consumers visiting several shops and therefore reduces the risk that a retailer will be undercut by a free- riding competing shop. Hence, each retailer will have a higher incentive to offer brand-supporting services. Another possibility is for the producer to maintain all the shops in the city, but fix the resale price (**resale price maintenance**), or impose a price floor, to avoid the problem of undercutting and to allow the retailers to recoup (part of) the investment.

A similar argument applies in case of **quality considerations**. For example, customers might have a preference for the product being available in a store a couple of times in different sizes. Or they like the product being distributed exclusively by a few specialized shops only and, thus, having a luxury brand image. In these cases it might be profitable for the firms and beneficial for the customers if the good is offered within a **selective distribution** scheme (Guidelines on Vertical Restraints, paras. 174-188). Only a few shops are licensed to sell the product which ensures that it is displayed as intended by the manufacturer and as desired by the customer. Selective distribution ensures that no further shops, who do not provide the desired quality of distribution, free-ride on the externality provided by their rivals by offering the good at a lower price.

The above considerations may lead to concerns about **free-riding among producers**. Suppose a retailer carries manufacturer A's brand. Therefore, producer A may invest into the retailer's services, “such as technical support, promotion, training, equipment and financing. To the extent that such investments favour not a particular brand but the retail outlet in general, other producers would also benefit from them. This gives rise to a free-riding problem that may be solved by resorting to exclusive dealers (i.e., retailers cannot stock products from competing brands) [...]. **Exclusive dealing** [*emphasis added*] might also push a retailer to sell a brand more aggressively than if it devoted its marketing effort among different brands, thereby raising competition” (Motta

2004: p. 335).

Vertical restraints may also be used to allow the manufacturer to capture **economies of scale**. A manufacturer may want to avoid supplying many outlets with a small amount of stock and would instead prefer to supply only a few outlets but with more stock each. A vertical restraint such as a quantity forcing requirement can solve this problem. Note that although this seems to be a legitimate use of a vertical restraint, it may be anti-competitive if the result is that so few retailers are supplied that intra-brand competition is significantly reduced in a market where inter-brand competition is weak.

Vertical restraints can also be used to avoid **opportunistic behavior** by one or other party to a vertical relationship. Some vertical relationships require relationship-specific investment, i.e. investment which looses most of its value outside a particular relationship, because they are tailored and dedicated to a particular partner. Parties may be unwilling to make that investment unless they can be reassured that the other party will not try to expropriate the value of the investment ex post. An example of opportunistic behavior might be the electricity generation plant built next to a steel plant that decides to raise prices to the steel plant because, once the steel plant's location decision has been made, the steel plant has no alternative suppliers. Knowing this danger, the steel company will not build its plant unless it is sure that it will not suffer from opportunistic behavior. A **long-term supply agreement** could be used to avoid this problem.

Intra-Brand Competition: The Commitment Problem

Vertical restraints can have an adverse effect on welfare when they help a manufacturer (more generally, an upstream firm) to keep prices high whereas without them it would not be able to commit to high prices. To understand why such a **commitment problem** (Motta 2004: pp. 338-347) arises, consider the following example.

Consider an industry with a **monopolist manufacturer upstream** and **two retailers downstream** who compete à la Cournot. This is the modeling structure proposed on p. 179 above. **Without any vertical agreement**, the equilibrium of this game is given by

$$\begin{aligned} Q_U &= \frac{a-c}{3} \\ r_U &= \frac{a+c}{2} \\ \pi_U &= \frac{(a-c)^2}{6} \end{aligned} \tag{120}$$

in the upstream market and by

$$\begin{aligned} Q_D &= \frac{a-c}{3} \\ p_D &= \frac{2a+c}{3} \\ \pi_{D,1}=\pi_{D,2} &= \left(\frac{a-c}{6}\right)^2 \end{aligned} \tag{156}$$

in the downstream market.

Now consider the situation where the monopolist manufacturer can sign **observable** (and legally enforceable) **contracts** with the retailers. In this case, they might agree on a franchise contract that each of the two retailers buys half of the quantity that a vertically integrated monopolist would sell (see equation (113))

$$q_{V,1}=q_{V,2}=(a-c)/4 \tag{157}$$

and pays a price of $r=c$ to the manufacturer. In doing so, each of the downstream retailers makes a profit

$$\pi_{V,1}=\pi_{V,2}=(a-c)^2/8 \quad . \tag{158}$$

It is easy to show that the sum of profits with this vertical agreement is above the sum of profits without it.

$$\pi_{V,1}+\pi_{V,2}>\pi_{D,1}+\pi_{D,2}+\pi_U \tag{159}$$

In order to be a Pareto-improvement for all firms, each of the downstream retailers would have to pay a franchise fee $T \in [(a\text{-}c)^2/12;(a\text{-}c)^2/8]$ to the upstream manufacturer. The exact size of the franchise fee will depend on the distribution of bargaining power among those firms. As proposed above, this vertical agreement also raises consumer surplus as the market price drops to $p_V=(a+c)/2$.

Now suppose the **contracts** were **not publicly observable** and the monopolist could renegotiate the contract with, say, retailer 2. Moreover, assume the manufacturer has full bargaining power and can appropriate all the profits being generated downstream via the franchise fee T. Hence, the manufacturer's profit π' consists of the franchise fee collected from retailer 1 $T_1=(a\text{-}c)^2/8$ plus the profit that retailer 2 can make if it is supplied by the manufacturer with some quantity q_2.

$$\begin{aligned} \pi' &= (p - r)\cdot q_2 + T_1 \\ &= (a-q_1-q_2 - c)\cdot q_2 + (a-c)^2/8 \\ &= (a-(a-c)/4-q_2 - c)\cdot q_2 + (a-c)^2/8 \end{aligned} \tag{160}$$

Maximizing π' with respect to q_2 yields an optimal quantity of

$$q_2'=(3/8)\cdot(a-c)>q_{V,2} \tag{161}$$

implying a lower downstream price causing a loss for retailer 1. Therefore, the contract offer ($T_i=(a\text{-}$

$c)^2/8$, $q_i=(a-c)/4$) cannot be an equilibrium, since each retailer would anticipate that if it signed such a contract, the manufacturer would have an incentive to offer larger output to the rival, which in turn would create losses for it. If contracts are not observable a welfare-enhancing franchising contract would not be signed and the above vertically separated equilibrium would be played. Similarly, without the monopolist's **ability to credibly commit to a contract** it would be unable to exercise its monopoly power.

The monopolist may engage in vertical restraints to commit to its contracts and exploit his market power which is likely to reduce welfare in this market. One way to do so is to engage in exclusive dealing and guarantee one of the retailers the exclusive right to serve this (geographic) market segment (**exclusive territories**). In the region protected by the exclusivity clause, competition among the potential retailers will bring them to pay a franchise fee up to $T_i=(a-c)^2/4$, i.e. the profits of a vertically integrated monopolist, to have the opportunity to be the only dealer selling the good. With full bargaining power, this will allow the manufacturer to appropriate all the monopoly profit. In the above example, this vertical restraint is welfare enhancing by eliminating double marginalization. However, if competition downstream would have been fierce (for example Bertrand-competition) absent the vertical agreement, the exclusivity clause does not improve welfare but contributes to engage in input foreclosure as one of the retailers would not be supplied any more. This emphasizes the notion that vertical restraints are rarely clearly pro-competitive or anti-competitive. Their effects should rather be assessed on a case-by-case basis.

As an alternative to assigning exclusive territories, the manufacturer might engage in **resale price maintenance** and require the retailers to sell the good (at least) at the profit-maximizing monopoly price (price floor). By supplying each of the retailers with have the quantity of a vertically integrated monopolist (see equation (157)) competition among retailers will drive the market price down to the monopoly price. The manufacturer may also gain credibility by including **most-favored customer** (or most-favored nation, MFN) **clauses** in its contracts with the retailers. Suppose that the manufacturer was able to credibly commit to and enforce a clause stating that whenever it offers a price discount to one retailer, all other retailers are also entitled to it. This would remove any temptation to renege on a previously signed contract with some retailers. Consider for instance the franchise example at the beginning of this section. If after having signed a contract with one franchisee for a price of $\pi/2$, the manufacturer sold a franchise to a second retailer for the price of $\pi/2$, under MFN it would have to reimburse the first retailer of $\pi/2$. Clearly, there would be no point in reneging on the promise and offering the franchise to a second retailer.

Above, the term **franchising** mainly refers to non-linear contracts that include a fixed (franchising) fee. In a wider definition "[f]ranchise agreements contain **licences of intellectual**

property rights relating in particular to trade marks or signs and know-how for the use and distribution of goods or services. In addition to the licence of IPRs, the franchisor usually provides the franchisee during the life of the agreement with **commercial or technical** assistance [*emphases added*]. The licence and the assistance are integral components of the business method being franchised. The franchisor is in general paid a franchise fee by the franchisee for the use of the particular business method. Franchising may enable the franchisor to establish, with limited investments, a uniform network for the distribution of its products. In addition to the provision of the business method, franchise agreements usually contain a combination of different vertical restraints concerning the products being distributed, in particular selective distribution and/or non-compete and/or exclusive distribution or weaker forms thereof" (Guidelines on Vertical Restraints, para. 189).

"This section has shown that vertical mergers and vertical restraints that affect only intra-brand competition are mostly efficiency-enhancing. [...] The main conclusion of this section is therefore that vertical restraints which affect intra-brand competition do not raise many welfare problems [...]. Another important implication of the analysis carried out here is that vertical restraints are often substitutable – at least to some extent – with each other" (Motta 2004: p. 347).

Inter-Brand Competition: Relaxation of Competition

Inter-brand competition becomes an issue when there are multiple suppliers of a good upstream (see the model on p. 185 above). In such a setting, vertical agreements may have **anti-competitive effects** such as the **relaxation of competition** between retailers and manufacturers, the facilitation of **collusive behavior** or **entry-deterrence**. In this context, the main competitive concerns in the upstream-market are increases in wholesale prices, a limitation in the choice of products, lower product quality, or a reduction in the level of product innovation. The main competitive concerns in the downstream-market relate to an increase in retail-prices, limitations in the choice of service combinations and distribution formats, reductions of the quality and availability of retail services or in the level of innovations of distribution (Guidelines on Vertical Restraints, para. 101). Again, the main conclusion of the following paragraphs is that one should only worry about vertical restraints when they involve firms endowed with large market power.

With regard to inter-brand competition vertical agreements can cause a **relaxation of competition**. To see this, consider the following model structure (Motta 2004: p. 351). **Two upstream manufacturers** U_1 and U_2 produce a differentiated product. The goods are sold via **two downstream retailers** R_1 and R_2 who compete in prices (**Bertrand-competition with differentiated products**). Manufacturer U_1 distributes its good via retailer R_1. Manufacturer U_2

distributes its good via retailer R_2. These two vertical chains may now decide whether

(i) both manufacturers shall **vertically integrate** with their respective retailer and compete only in the downstream market as duopolists or whether

(ii) they shall remain independent and choose a **non-linear pricing contract** that is perfectly observable and not renegotiable. In this case, the upstream manufacturers possess full bargaining power.

In case (ii), the firms specify the terms of the pricing contract (i.e. wholesale price r and fixed transfer T) in stage 1. In stage 2, the retailers compete downstream in prices, realize profits and pay the transfer $t_i=q_i r_i+T_i$ to the manufacturer.

In the vertically integrated case (i), each manufacturer would optimally set a wholesale price equaling its marginal costs ($r_i=c_i$) and choose prices in the downstream market accordingly. It can be shown that the manufacturers in the vertically separated case (ii) can do even better than mimicking the vertically integrated equilibrium. This is because goods in Bertrand-competition are strategic complements. If one of the retailers raises the price of its good the other responds by raising its respective price in return. This softens competition among the retailers and raises their profits. As the manufacturers possess full bargaining power they extract the retailers' entire profits via the fixed transfer and, thus, benefit from the increased profits themselves. Hence a relaxation of competition downstream contributes to a relaxation of competition upstream.

The means by which the manufacturers can induce the retailers to soften competition and increase prices downstream is to set wholesale prices above marginal costs ($r_i>c_i$). As a consequence, the retailers will choose higher prices downstream than they would in the vertically integrated situation. Here, a vertical agreement in the form of a **non-linear pricing contract with $r_i>c_i$** is to the detriment of consumers and benefits the upstream manufacturers.

A similar result can be obtained for the assignment of **exclusive territories** (Motta 2004: p. 354). To see this, consider the above example and assume that each manufacturer does not sell its product via one retailer but multiple retailers. This adds a component of intra-brand competition downstream causing low price levels. By assigning the retailers an exclusive territory a manufacturer creates monopolist retailers in local sub-markets. This relaxes downstream competition which may benefit manufacturers upstream.

It should be noted that a relaxation of competition benefits a manufacturer the less the fewer market power it possesses. A manufacturer with a high level of market power can – e.g. by increasing its wholesale price r – trigger a stronger response on downstream prices than a powerless firm. As market power decreases when competition in a market increases we expect vertical

restraints to be **less of a welfare concern when competition among manufacturers is intense**.

The above results are strongly dependent on the mode of competition. If the downstream retailers **compete in quantities** (Motta 2004: 356) a completely different result can be shown to arise. In this case, quantities as retailers' strategic variables are strategic substitutes. When retailer 1 sets a lower quantity in order to raise the downstream price, retailer 2 will respond by increasing its own quantity and exploit the higher price. It can be shown that manufacturers would optimally set a **non-linear pricing contract** with a wholesale price below their production costs ($r<c$). This would trigger aggressive competition downstream causing lower prices and higher welfare than in a vertically integrated situation. Note that the upstream manufacturers would not make economic losses despite setting $r<c$ as they also receive the fixed transfer payment from the downstream retailers.

The Guidelines on Vertical Restraints (paras. 151-167) refer to the above case as an **exclusive distribution agreement**, i.e. “the supplier agrees to sell its products to only one distributor for resale in a particular territory. [...] The possible competition risks are mainly reduced intra-brand competition and market partitioning, which may facilitate price discrimination in particular. When most or all of the suppliers apply exclusive distribution, it may soften competition and facilitate collusion, both at the suppliers' and distributors' level. Lastly, exclusive distribution may lead to foreclosure of other distributors and therewith reduce competition at that level” (para. 151). “The market position of the supplier and its competitors is of major importance, as the loss of intra-brand competition can only be problematic if inter-brand competition is limited. The stronger the position of the supplier, the more serious is the loss of intra-brand competition. Above the 30% market share threshold, there may be a risk of a significant reduction of intra-brand competition” (para. 153).

Inter-Brand Competition: Collusive Devices

In some cases, vertical agreements can be used as collusive devices. First, **resale price maintenance** (RPM) may act as a collusive device because it increases price observability. “Absent RPM, when shocks in the retail markets occur, final prices will tend to change, making it more difficult for manufacturers to distinguish changes in retail prices that are caused by different retail conditions from cheating on the cartel. RPM makes collusion more likely by eliminating the retail price variation” (Motta 2004: p. 359).

Moreover, it can be shown that agreements on the commercialization through a **common agency** (also see p. 234 below) can serve as a collusive device. This is obvious when two competing manufacturers sell their products through a common retailer and leave the pricing choice

to the retailer who will, thus, choose the jointly profit-maximizing monopoly prices. The same is true when the manufacturers set a non-linear pricing contract such as to extract the retailer's profits via the fixed fee and engage in resale price maintenance. The manufacturers will maximize their own profits by, again, setting the downstream price at the jointly profit-maximizing levels.

Inter-Brand Competition: Foreclosure

The concerns of foreclosure that arise in the context of vertical agreements are similar to those raised by vertical mergers (see pp. 179-188). For example, an **exclusive dealing contract** may result in input foreclosure. However, a rational buyer might not be willing to accept a contract that requires her to pay a higher price as in competition (Motta 2004: p. 363). Suppose for instance in a certain industry there is an incumbent monopolist with marginal costs c_h, a potential entrant with marginal costs c_l and $c_h>c_l$ and only one buyer who demands $Q=a-p$. By accepting an exclusive dealing contract, the buyer would commit to buy from the incumbent monopolist even if entry occurs. This rules out entry, and the buyer will end up paying the monopoly price $p_m=(a+c_h)/2$ for the good enjoying a consumer surplus $CS_m=(a-c_h)^2/8$ (area A). By rejecting the contract offer the buyer would trigger entry.

If the two **firms compete in quantities** the market price will drop to $p_e=(a+c_l+c_h)/3$ while consumer surplus rises at $CS_e=(2a-c_l-c_h)^2/18$ (area A+B+C). Sure enough, the incumbent might offer a compensation to the buyer to persuade her to accept exclusivity. However, the incumbent is willing to pay a compensation no higher than its monopoly profit $\pi_m=(a-c_h)^2/4$ (area B+D). However, the buyer - by accepting the exclusivity contract - would lose all the consumer surplus (CS_e-CS_m, area B+C) that arises by buying at lower prices. Hence, in the above example a rational buyer would only accept an exclusive dealing contract if area D is larger than area C. In this case, both parties might benefit from such a contract. The customer wins by receiving a transfer payment. The monopolist wins by remaining in the market if its monopoly profit (net of the transfer payment to the customer) exceeds its profit in the Cournot-duopoly.

Depending on the shape of demand, the size of marginal costs and the mode of competition of the manufacturers, it is easy to provide examples where the customer would not enter into an exclusive dealing contract. This is, e.g., the case when the two firms would enter into **Bertrand-competition** after entry of the more efficient firm (Motta 2004: p. 367). In this case, the more efficient entrant would set a price $p_e'=c_h-\varepsilon$ and drive the incumbent out of the market. The resulting increase in consumer surplus equals the area B+C+D+E+F. This is clearly larger than the incumbent monopolist's profit prior to entry B+D. In this case, a rational buyer would not enter into an exclusive dealing contract with the incumbent.

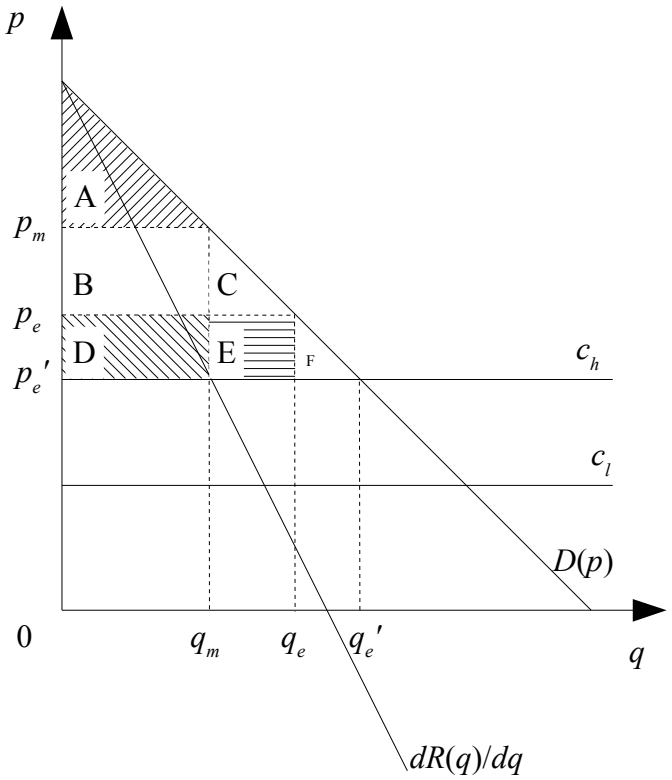


Figure 31: Exclusive Dealing and Foreclosure

However, exclusive dealing contracts can serve as a **barrier to entry** if there is **uncertainty about the costs** and, thus, the price of the entrant (Motta 2004: p. 368). In this case, an exclusive dealing contract might serve as some form of insurance. The incumbent monopolist commits to selling the product at some lower price than without the contract. However, if the entrant decides to enter the industry and the customer, who accepted the exclusive dealing contract, buys from the entrant it must pay a contractual penalty to the incumbent. The buyer will accept the contract if in expectation she will be better off with it than without. After the entrant's costs and price are known the buyer may win from the contract when the entrant sets a high price. However, she looses when the entrant would set a low price. It can be shown that such a contract in presence of uncertainty about the entrant's cost realization deters entry unless the entrant is much more efficient than the incumbent.

In practice, exclusive dealing may arise in the form of **single branding**, i.e. a retailer may only store products of a certain brand. “The possible competition risks of single branding are foreclosure of the market to competing suppliers and potential suppliers, softening of competition and facilitation of collusion between suppliers in case of cumulative use and, where the buyer is a retailer selling to final consumers, a loss of in-store inter-brand competition” (Guidelines on Vertical Restraints, para. 130). Single branding is typically unlike to create competition concerns if

a large number of different brands competes in engaging in an exclusive dealing contract with a retailer. “In cases where the market share of the largest supplier is below 30% and the market share of the five largest suppliers is below 50%, there is unlikely to be a single or a cumulative anti-competitive effect situation” (Guidelines on Vertical Restraints, para. 135). Single branding may be detrimental to welfare if the brand in question is a *must stock* brand that is essential for the retailer for doing business. In this context, “[c]ountervailing power is relevant, as powerful buyers will not easily allow themselves to be cut off from the supply of competing goods or services” (Guidelines on Vertical Restraints, para. 137).

Alternatively, “[u]nder the heading of **exclusive supply** [*emphasis added*] fall those restrictions that have as their main element that the supplier is obliged or induced to sell the contract products only or mainly to one buyer, in general or for a particular use. [...] The main competition risk of exclusive supply is anti-competitive foreclosure of other buyers. [...] If the buyer has no market power downstream, then no appreciable negative effects for consumers can be expected. Negative effects may arise when the market share of the buyer on the downstream supply market as well as the upstream purchase market exceeds 30%.” (Guidelines on Vertical Restraints, paras. 192-194).

Resale Price Maintenance

Resale price maintenance (RPM), i.e. “agreements or concerted practices having as their direct or indirect object the establishment of a fixed or minimum resale price or a fixed or minimum price level to be observed by the buyer” (para. 223), has played an important role in many parts of the above subsection. Therefore, we devote the following paragraphs to the description how resale price maintenance is considered in the EU's Guidelines on Vertical Restraints (paras. 223-229). In principle, RPM is regarded a hardcore restriction within Art. 101 I TFEU. “However, undertakings have the possibility to plead an **efficiency defence** [*emphasis added*] under Article 101(3) in an individual case. It is incumbent on the parties to substantiate that likely efficiencies result from including RPM in their agreement and demonstrate that all the conditions of Article 101(3) are fulfilled” (para. 223).

“RPM may restrict competition in a number of ways. Firstly, RPM may **facilitate collusion** [*emphasis added*] between suppliers by enhancing price transparency on the market, thereby making it easier to detect whether a supplier deviates from the collusive equilibrium by cutting its price. [...] Second, by **eliminating intra-brand price competition** [*emphasis added*], RPM may also facilitate collusion between the buyers, that is, at the distribution level. Strong or well organised distributors may be able to force or convince one or more suppliers to fix their resale

price above the competitive level and thereby help them to reach or stabilise a collusive equilibrium. [...] Third, RPM may more generally **soften competition** [*emphasis added*] between manufacturers and/or between retailers, in particular when manufacturers use the same distributors to distribute their products and RPM is applied by all or many of them. Fourth, the immediate effect of RPM will be that all or certain distributors are prevented from lowering their sales price for that particular brand. In other words, **the direct effect of RPM is a price increase** [*emphasis added*]. Fifth, RPM may lower the pressure on the margin of the manufacturer, in particular where the manufacturer has a **commitment problem** [*emphasis added*], that is, where it has an interest in lowering the price charged to subsequent distributors. In such a situation, the manufacturer may prefer to agree to RPM, so as to help it to commit not to lower the price for subsequent distributors and to reduce the pressure on its own margin. Sixth, RPM may be implemented by a manufacturer with market power to **foreclose smaller rivals** [*emphasis added*]. The increased margin that RPM may offer distributors, may entice the latter to favour the particular brand over rival brands when advising customers, even where such advice is not in the interest of these customers, or not to sell these rival brands at all. Lastly, RPM may **reduce dynamism and innovation** [*emphasis added*] at the distribution level" (para. 224).

However, RPM may also create efficiencies. "Most notably, where a manufacturer introduces a new product, RPM may be helpful during the introductory period of expanding demand to induce distributors to better take into account the manufacturer's interest to **promote the product** [*emphasis added*]. [...] Similarly, fixed resale prices, and not just maximum resale prices, may be necessary to organise in a **franchise system** [*emphasis added*] or similar distribution system applying a uniform distribution format a coordinated short term low price campaign (2 to 6 weeks in most cases) which will also benefit the consumers. In some situations, the extra margin provided by RPM may allow retailers to provide (additional) **pre-sales services** [*emphasis added*], in particular in case of experience or complex products. If enough customers take advantage from such services to make their choice but then purchase at a lower price with retailers that do not provide such services (and hence do not incur these costs), high-service retailers may reduce or eliminate these services that enhance the demand for the supplier's product. RPM may help to **prevent** such **free-riding** [*emphasis added*] at the distribution level. The parties will have to convincingly demonstrate that the RPM agreement can be expected to not only provide the means but also the incentive to overcome possible free riding between retailers on these services and that the pre-sales services overall benefit consumers as part of the demonstration that all the conditions of Article 101(3) are fulfilled" (para. 225).

"The possible competition risk of **maximum and recommended prices** [*emphasis added*] is that they will work as a focal point for the resellers and might be followed by most or all of them

and/or that maximum or recommended prices may **soften competition** [*emphasis added*] or **facilitate collusion** [*emphasis added*] between suppliers.

An important factor for assessing possible anti-competitive effects of maximum or recommended resale prices is the **market position of the supplie**r [*emphasis added*]. The stronger the market position of the supplier, the higher the risk that a maximum resale price or a recommended resale price leads to a more or less uniform application of that price level by the resellers, because they may use it as a focal point. They may find it difficult to deviate from what they perceive to be the preferred resale price proposed by such an important supplier on the market" (paras. 227-228).

Lessons Learned

After reading this section you should be able to answer the following questions.

1. What is meant by the term *sub-game perfect Nash equilibrium*?
2. Describe intuitively how a cartel can be stabilized when the timing of the break-down of a market is indeterminate.
3. Do the terms *collusion* and *cartel* mean the same? Explain.
4. How does a *grim trigger strategy* work?
5. Describe the effects of a competition authority on the stability of cartels.
6. Reply to the following statement: "Collusion is easier to sustain in markets for homogeneous products."
7. Reply to the following statement: "Cartels are more likely to break down in an upturn than in a downturn."
8. What types of agreements are prohibited by Art. 101 TFEU? Which conditions must be satisfied such that an agreement can nonetheless be allowed?
9. In what ways may the exchange of information affect market behavior and market outcomes?
10. Describe the main anti-competitive concerns and efficiency effects of production agreements.
11. List and describe some common vertical restraints.
12. Define the terms *inter-brand competition* and *intra-brand competition*.

13. What pro-competitive effects are likely to arise from vertical restraints?
14. What anti-competitive effects can be caused by vertical restraints.
15. What position does EU competition policy take towards resale price maintenance?

References

Bishop, S. and Walker, M. (2010). "The Economics of EC Competition Law." 3rd edition. Sweet and Maxwell: London

BKartA (2012). "Guidance on Substantive Merger Control." http://www.bundeskartellamt.de/wEnglisch/download/pdf/Merkblaetter/2012-03-29_Guidance_final_neu.pdf

European Commission (2010.a). "Commission Regulation (EU) No 1217/2010 of 14 December 2010 on the application of Article 101(3) of the Treaty on the Functioning of the European Union to certain categories of research and development agreements." Official Journal of the European Union. L 335/36. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:335:0036:0042:EN:PDF

European Commission (2011). "Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements." Official Journal of the European Union. C11/1. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2011:011:0001:0072:EN:PDF

Green, E.J. and Porter, R.H. (1984). "Noncooperative Collusion under Imperfect Price Information." Econometrica. Vol. 52 No. 1, pp. 87-100

Haltiwanger, J. and Harrington, J.E. (1991). "The impact of cyclical demand movements on collusive behavior." RAND Journal of Economics. Vol. 22 No. 1, pp. 89-106

Kühn, K.U. and Vives, X. (1994). "Information Exchanges Among Firms and their Impact on Competition." http://webprofesores.iese.edu/xvives/Docs/reports/Information%20Exchanges%20and%20their%20Impact%20on%20Competition.pdf

Levenstein M.C. and Suslow, V.Y. (2011). "Breaking Up Is Hard to Do: Determinants of Cartel Duration." The Journal of Law and Economics. Vol. 54 No. 2, pp. 455-492

Motta, M. (2004). "Competition Policy – Theory and Practice." Cambridge University Press: Cambridge

Paha, J. (2010). "Endogenous Cartel Formation with Heterogeneous Firms and Differentiated Products." MAGKS Discussion Paper No. 07-2010. http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/07-2010_paha.pdf

Pepall, L. and Richards, D. and Norman, G. (2008). "Industrial Organization – Contemporary Theory and Empirical Applications." 4th edition. Blackwell Publishing: Malden, MA

Ross, T.W. (1992). "Cartel Stability and Product Differentiation." International Journal of Industrial Organization. Vol. 10, pp. 1-13

Rotemberg, J.J. and Saloner, G. (1986). "A Supergame-Theoretic Model of Price Wars during Booms." The American Economic Review. Vol. 76 No. 3, pp. 390-407

I ART. 102 TFEU – ABUSE OF DOMINANCE

I.1 Introduction

Article 102 TFEU seeks to prevent firms from engaging in anti-competitive behavior:

> *"Any abuse by one or more undertakings of a dominant position within the internal market or in a substantial part of it shall be prohibited as incompatible with the internal market in so far as it may affect trade between Member States. Such abuse may, in particular, consist in:*
>
> (a) *directly or indirectly imposing unfair purchase or selling prices or other unfair trading conditions;*
>
> (b) *limiting production, markets or technical development to the prejudice of consumers;*
>
> (c) *applying dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage;*
>
> (d) *making the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts."*

The corresponding parts in the German competition law are §§ 19-21 GWB (*Gesetz gegen Wettbewerbsbeschränkungen*; see p. 282 below). In the following, we concentrate on the European policy against the abuse of a dominant position as outlined in the Commission's *Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings* (EU 2008; hereafter Article 102 Guidance).

The Article 102 Guidance (para. 1) establishes that "it is not in itself illegal for an undertaking to be in a dominant position [...]. However, the undertaking concerned has a special responsibility not to allow its conduct to impair genuine undistorted competition on the common market." In particular, the Commission applies a **consumer welfare standard** (also see EAGCP 2005: 2) by focusing "on those types of conduct that are most harmful to consumers. [...] The Commission, therefore, will direct its enforcement to ensuring that markets function properly and that consumers benefit from the efficiency and productivity which result from effective competition between undertakings" (EU 2008: para. 5). "The emphasis of the Commission's enforcement activity [...] is on safeguarding the competitive process in the internal market and ensuring that undertakings which hold a dominant position do not exclude their competitors by other means than

competing on the merits of the products or services they provide. In doing so the Commission is mindful that what really matters is protecting an effective competitive process and **not simply protecting competitors** [*emphasis added*]. This may well mean that competitors who deliver less to consumers in terms of price, choice, quality and innovation will leave the market" (EU 2008: para. 6). Insofar, the Commission recognizes that exit of, for example, productively inefficient firms may contribute to improving overall welfare in a market (see section C).

In the EU the assessment of an alleged abuse of a dominant position follows an effects-based approach rather than applying a **form-based approach** (*per se* rules) that focuses on the mere form of a particular business practice (for example, exclusive dealing, tying etc.). The **effects-based approach** (rule of reason) requires an assessment of both the pro- and anti-competitive effects of a practice in each specific case. Ideally, this approach does not affect the "amount" of market-intervention but establishes more effective intervention (EAGCP 2005: 3-4). In this context, the Economic Advisory Group on Competition Policy advises that competition "policy is based on the principle that competition itself is the best mechanism for avoiding inefficiencies, so the competition authority should not try to let its own intervention replace the role of competition in the market place" (EAGCP 2005: 10).

Investigating the alleged abuse of a dominant position requires a **two-step procedure**. First, the Commission must assess whether an undertaking is in a dominant position and what degree of market power it holds. Second, it must assess whether a particular type of behavior is abusive. These two matters are treated in greater detail in subsections I.2 and I.3 .

I.2 Existence of a Dominant Position

Para. 4 of the Article 102 Guidance states that a dominant position "may be held by one undertaking (single dominance) or by two or more undertakings (collective dominance)." The Guidance document only relates to single dominance. Both types of dominance are described below.

Single Dominance

"**Dominance** has been defined under Community law as a position of economic strength enjoyed by an undertaking, which enables it to **prevent effective competition** being maintained on a relevant market, by affording it the power to **behave** to an appreciable extent **independently of its competitors**, its **customers** and ultimately of **consumers**. [...] Dominance entails that [the] competitive constraints are not sufficiently effective and hence that the undertaking in question **enjoys substantial market power** [see section E] over a period of time. This means that the undertaking's decisions are largely insensitive to the actions and reactions of competitors, customers

and, ultimately, consumers. [...] In general, a dominant position derives from a combination of several factors which, taken separately, are not necessarily determinative [*emphases added*]" (EU 2008: para. 10). In particular, the Commission considers the existence of **market power** (which is equivalent to the existence of a dominant position) to result in the capability of firms to **profitably raise prices** above the competitive level for a significant period of time (EU 2008: para. 11). This direct effect of dominance sides an indirect effect, i.e. **exclusionary conduct** (foreclosure) by dominant firms which harms competitors and, by thus, results in higher prices to consumers. In this context, note that harm to competitors represents a necessary but not a sufficient condition for harm to consumers. This requires an more detailed, effects-based assessment of supposedly exclusionary conduct. (Bishop and Walker 2010: p. 234).

"The **assessment of dominance** [*emphasis added*] will take into account the competitive structure of the market, and in particular the following factors:

- [...] the market position of the dominant undertaking and its competitors [...],
- constraints imposed by the credible threat of future expansion by actual competitors or entry by potential competitors [...],
- constraints imposed by the bargaining strength of the undertaking's customers (countervailing buyer power)" (EU 2008: para. 12).

These factors have been described more thoroughly in section E above. The Commission uses **market shares** as a useful first indication in the analysis of the market structure and the firms' respective strength, also taking into account the evolution and volatility of market shares. "The [European Court of Justice] has held that a market share in excess of 50 per cent can be considered to be so large that except in exceptional circumstances such an undertaking could be presumed to be dominant. The Commission has stated that it takes the view that a dominant position can generally be taken to exist when a firm has a market share greater than 40 or 45 per cent, although it cannot be ruled out for undertakings with a lower market share" (Bishop and Walker 2010: p. 224). This focus on market shares emphasizes the importance of an adequate definition of the relevant market (see section F).

An undertaking can be deterred from increasing prices if **expansion or entry** is likely (i.e. it is sufficiently profitable), timely and sufficient (i.e. it is not only small-scale). The profitability of entry also crucially depends on the existence of entry barriers as described in section E.2 . A further factor that may deter firms from increasing prices, and thus needs needs to be assessed, is the existence of **countervailing buyer power** (see section E.2).

Collective Dominance

In principle, the provisions of Art. 102 TFEU also apply to firms that are collectively dominant, i.e. firms who collectively possess market power. "For collective dominance to exist requires that there exist a group of firms that: (a) do not face significant competitive constraints from other firms outside the group; and (b) these firms are able to adopt a mode of parallel behaviour that reduces the effectiveness of competition between themselves. Under these circumstances, it is possible that prices would be increased above the competitive level. There is therefore a close correspondence between the concepts of collective dominance and **tacit coordination** [*emphasis added*]" (Bishop and Walker 2010: p. 340). The main difficulty with collective dominance is that tacitly collusive behavior can hardly be distinguished from competitive behavior by analyzing market data.

I.3 Types of Abuses

"Once a firm has been held to enjoy a dominant position, the second stage of the investigation asks whether that firm's conduct represents an abuse of a dominant position. Article [102] does not prohibit firms from holding dominant positions, only from abusing such positions" (Bishop and Walker 2010: p. 225). The term abuse refers to a situation where a dominant firm behaves in a manner different from normal competitive behavior. As the Article 102 Guideance focuses on consumer welfare, an **abuse** can "be defined as a dominant firm adopting a particular mode of behaviour that significantly reduces consumer welfare relative to the alternative of the firm not adopting that mode of behaviour. Where it cannot be shown that the behaviour of a dominant firm adversely affects consumers, either immediately or in the longer term, such behaviour should be considered as constituting normal competitive behaviour" (Bishop and Walker 2010: p. 230).

Abusive conduct can be grouped into two broad types, i.e. **excessive pricing** (exploitative abuses) which directly harms consumers and **foreclosure** (exclusionary abuses). The latter can occur in different forms and harms consumers "indirectly by excluding or foreclosing competitors and as a result increasing that firm's ability to increase prices to consumers" (Bishop and Walker 2010: 230). Article 102 TFEU is primarily concerned with exclusionary abuses. A relevant issue in the assessment of an abuse is the question how one can discriminate between pro- and anti-competitive conduct.

Exploitative Abuses (Excessive Pricing)

"Excessive pricing occurs where a dominant firm exercises its market power by raising prices above the competitive level. Since a dominant firm is defined as a firm possessing significant market power, and market power is the ability to raise prices profitably above the competitive level, then

excessive pricing represents perhaps the most intuitive form of abuse; the dominant firm increases its price to consumers and in so doing increases it profits. However, there have been very few pure excessive pricing cases under Article [102]. Instead, excessive pricing has tended to be subsumed into aftermarket cases where higher prices are charged to consumers that are locked-in" (Bishop and Walker 2010: 237).

Proving excessive pricing is extremely difficult in practice as it requires to compare the observed level of prices to **counterfactual**, competitive values that are hard to construct. Moreover, one would need to prove the profitability of this price-rise. This is particularly difficult when firms produce at different marginal costs. The discussion in section E.2 shows that entry into a market occurs until the least efficient firm in the market makes zero economic profits with no firm being outside the market that could make positive profits. Hence, the price only equals the average total cost of this marginal firm but will be more than average total costs of the infra-marginal firms. Observing a firm charging a price above its costs does not indicate that this firm exercises market power but can simply be a result of its **productive efficiency**. Disentangling the effects of market power and productive efficiency is a difficult task that requires analyzing the industry as a whole rather than focusing on a single firm alone.

A further difficulty arises when production requires a large amount of **sunk costs** for, e.g., R&D but relatively small (marginal) costs of production. For example, this is the case in the pharmaceutical industry or in the software industry. In these cases, it is hard if not impossible to establish a meaningful (empirical) relationship between the costs of production and the market price. Moreover, in case of a fixed production capacity prices will be high (above average production costs) when demand is high and be low when demand is low, i.e. when **demand is cyclical**. Thus, observing prices above costs is rather a sign of the allocative efficiency of the market than the exercise of market power (Bishop and Walker 2010: p. 238).

Testing for excessive prices can basically be done by answering at least one of five questions.

1. Are prices in the relevant market higher than in another geographic market?
2. Are the prices of the dominant firm above the prices of its competitors?
3. Are prices higher than those of (fairly) similar products?
4. Are current prices higher than in previous periods?
5. Are profit rates above competitive levels?

Comparing prices across different regions can be instructive. However, one must make sure that the comparator-regions are sufficiently similar to the region of interest, e.g., in terms of

demand conditions and the production technology used. When demand differs across these regions, price differences can be a consequence of welfare enhancing price discrimination. When the production technology differs, the production costs are likely to differ causing a difference in prices. Moreover, one would like to make sure that prices in the comparator-region are free from effects like the abuse of a dominant position or other market-imperfections.

Comparing prices across firms usually is not an ideal approach. “If one firm charges higher prices than other firms, and yet consumers still buy the product, this indicates that consumers consider the product in question to be superior in some respect to the products of other firms. If this is the case, a simple price comparison is not comparing, in the eyes of consumers at least, apples with apples but rather apples with pears. Only if consumers are in some sense "locked in" to buying the higher-priced product might such a simple price comparison make sense” (Bishop and Walker 2010: p. 239). Moreover, by raising its own price the dominant firm gives its competitors the freedom to raise their prices, which then cannot be used as a benchmark for truly competitive prices.

Comparing prices across similar products is a difficult task. First, think of a comparison of prices of the allegedly dominant firm. For example, the national prices of a postal service provider may be compared to its own international prices. However, such a comparison is subject to the above objections that differences in demand conditions and production costs may cause price differences other than those caused by the abuse of a dominant position. Second, these effects become even more severe if one compares the prices of the allegedly dominant firm's products to those of similar products offered by other firms.

A comparison of prices across periods, i.e. prices before and during the period of the supposed abuse, is advantageous as many industry characteristics remain quite similar over time and, thus, reduce the difficulties associated with comparisons across regions, firms, or products. Even in this case one must pay close attention to changes in the demand and supply conditions that occur over time and affect prices. Moreover, one is required to identify the start date of the abuse beforehand and make sure that the transition from the supposedly competitive prices to the allegedly abusive prices is modeled correctly.

“A further approach to testing for excessive prices is to use **rates of profit** [*emphasis added*] to infer abuse. As noted [in sections A.2 and E.1], the theoretical link between high profits and the exercise of market power is not unambiguous for a number of reasons. In addition, there are often serious practical difficulties. Most attempts to measure profitability rely on accounting measures of profitability. These do not usually reflect the relevant economic concepts. Comparisons across firms, particularly across regions, can be biased by the use of different accounting conventions.

Profitability measures are usually based on rates of return on some measure of capital. Measuring capital appropriately is difficult where a significant proportion of the asset base is not made up of physical assets, but of intangible assets, such as human capital. If intangibles are not properly accounted for, it leads to estimates of supra-normal profits that in reality are merely a reflection of a mis-measured capital base. Finally, the question of what constitutes too high a rate of return is also not an easy question to answer. There is a tendency to relate accounting measures of profitability to market-based estimates of the required rate of return. But this does not constitute a like-for-like comparison and hence is not a valid approach" (Bishop and Walker 2010: p. 241).

"The difficulties of distinguishing between excessive prices and those implied by conditions of effective competition are reflected in the paucity of decisions in this area. It is noticeable that apart from United Brands there are few instances of the Commission dealing with pure excessive pricing cases outside of cases where firms possess statutory or what are often held to be de facto monopoly positions. Thus, for instance, the issue of excessive pricing arose in the context of telecommunications, where fixed-line operators often have statutory or de facto monopolies" (Bishop and Walker 2010: p. 241; see there for a range of examples).

"In some respects, the conclusion of this section might appear surprising. Even though excessive pricing would seem to be the most direct form of abuse (i.e. exercising substantial market power to raise prices), there are few pure excessive pricing cases in the case law. Both economic theory and practice indicate that testing for whether prices are excessive is extremely difficult. Given these difficulties, the observation of high prices relative to unit production costs should not in isolation be seen as being indicative of abusive behaviour. Rather, a detailed assessment of the market characteristics is required to determine whether the firm is subject to effective competitive constraints or not and whether its behaviour is likely to be harming consumers.

Finally it should be noted that a finding of excessive pricing logically implies that the Commission should insist that the firm in question lowers its price to a non-excessive level. This requires the Commission to have some idea of what this price level is, which implies detailed price regulation. Given that the Commission has said on numerous occasions that it does not wish to act as a price regulator, we expect pure excessive pricing cases to remain rare in the future" (Bishop and Walker 2010: p. 244).

Exclusionary Abuses

"The aim of the Commission's enforcement activity in relation to exclusionary conduct is to ensure that dominant undertakings do not impair effective competition by foreclosing their competitors in an anti-competitive way, thus having an adverse impact on consumer welfare, whether in the form

of higher price levels than would have otherwise prevailed or in some other form such as limiting quality or reducing consumer choice. In [the Art. 102 Guidance] the term **'anti-competitive foreclosure'** [*emphasis added*] is used to describe a situation where effective access of actual or potential competitors to supplies or markets is hampered or eliminated as a result of the conduct of the dominant undertaking whereby the dominant undertaking is likely to be in a position to profitably increase prices to the detriment of consumers" (EU 2008: para. 19).

Moreover, the EU can also be concerned with behavior where firms intentionally lower their price below its competitive level. In case of **price-based exclusionary conduct** a dominant firm may set prices below costs and by thus drive a competitor out of the market. If the excluded firm had previously posed a competitive constraint on the dominant undertaking the exclusion is likely to result in higher prices. In order to detect such behavior the Commission will engage in analyses of prices and costs. Setting prices below **average variable costs** "indicates that the dominant undertaking is sacrificing profits in the short term and that an equally efficient competitor cannot serve the targeted customers without incurring a loss" (EU 2008: para. 26). Setting prices below **average total costs** "indicates that the dominant undertaking is not recovering all the (attributable) fixed costs of producing the good or service in question and that an equally efficient competitor could be foreclosed from the market" (EU 2008: para. 26).

Three categories of exclusionary behavior can be distinguished (EAGCP 2005: 17): First, **exclusion within the same market**, where an incumbent forces the exit or prevents the entry of a competitor. Such conduct is described on pages 280-273 (predation). Second, **exclusion in an adjacent market** where the dominant firm excludes producers active in markets different but related to its main market (see pages 275-280 on tying and bundling). Third, **exclusion in a vertically related market**, where exclusion takes places in different stages of the production process (see pages 267-275 on exclusive dealing and pages 273-282 on refusal to supply and margin squeeze).

Similar to the assessment of infringements of Art. 101, a dominant firm may justify its conduct with recourse to **efficiency effects** that arise because of this conduct. "In this context, the dominant undertaking will generally be expected to demonstrate, with a sufficient degree of probability, and on the basis of verifiable evidence, that the following cumulative conditions are fulfilled:

- the efficiencies have been, or are likely to be, realised as a result of the conduct. [...]
- the conduct is indispensable to the realisation of those efficiencies: there must be no less anti-competitive alternatives to the conduct that are capable of producing the same efficiencies,

- the likely efficiencies brought about by the conduct outweigh any likely negative effects on competition and consumer welfare in the affected markets,
- the conduct does not eliminate effective competition, by removing all or most existing sources of actual or potential competition" (EU 2008: para. 30).

Exclusionary Abuses – Exclusive Dealing (Price Discrimination)

"A dominant undertaking may try to foreclose its competitors by hindering them from selling to customers through use of **exclusive purchasing** obligations or **rebates** [*emphases added*], together referred to as exclusive dealing" (EU 2008: para. 28). We describe these types of exclusionary abuses in turn.

"An exclusive purchasing obligation requires a customer on a particular market to **purchase exclusively or to a large extent only from the dominant undertaking**" (EU 2008: para. 33; *emphasis added*). As is shown on pp. 252-254 above, the dominant undertaking may have to compensate customers for accepting an exclusive purchasing contract which implies a positive effect on customers by the agreement on exclusivity. Therefore, the Commission focuses its attention on those cases where the customers are likely to be harmed by exclusive purchasing. In particular, this is the case if the exclusivity agreement forecloses other suppliers from the market or prevents the entry of new suppliers.

"**Conditional rebates** are rebates granted to customers to reward them for a particular form of purchasing behaviour. The usual nature of a conditional rebate is that the customer is given a rebate if its purchases over a defined reference period exceed a certain threshold, the rebate being granted either on all purchases (**retroactive rebates**) or only on those made in excess of those required to achieve the threshold (**incremental rebates**) [*emphases added*]. Conditional rebates are not an uncommon practice. Undertakings may offer such rebates in order to attract more demand, and as such they may stimulate demand and benefit consumers. However, such rebates – when granted by a dominant undertaking – can also have actual or potential foreclosure effects similar to exclusive purchasing obligations. Conditional rebates can have such effects without necessarily entailing a sacrifice for the dominant undertaking" (EU 2008: para. 37).

"The following stylised discount types can all be viewed as loyalty rebates.

- **An exclusivity discount**: The buyer obtains a discount only by purchasing all its needs from the supplier.
- **An individualised quantity discount**: Each buyer is offered a discount conditional on purchasing a given quantity within a particular reference period. In this case the targets may

differ for buyers of different sizes.

- **A growth discount**: The buyer receives a discount if its purchases in the current period exceed its purchases in the relevant past period by a given amount. A growth target does not necessarily imply that meeting the threshold target entails an increase in the supplier's market share; if the target growth threshold is set at a level lower than the growth in the overall market, then the threshold can be met without diverting share from competing suppliers.
- **A bundled discount**: The target relates to purchases across a range of products. In this case, it might be necessary for the buyer to purchase a certain amount of another product supplied by the firm in question in order to qualify for the rebate where the target amounts to a large portion of its needs of that product.

The term "loyalty rebate" therefore covers a wide range of discount schemes. What each of these different schemes have in common is that they provide incentives for customers to purchase more product or services from the firm offering the loyalty scheme. Put simply, all loyalty rebate schemes create incentives for customers to purchase more from the firm employing the loyalty rebate scheme, and they are therefore likely to result in share-shifting" (Bishop and Walker 2010: p. 257).

The economic essence of loyalty rebates is **price discrimination**, i.e. dissimilar prices are charged from different groups of customers where the difference in prices cannot be attributed to differences in the costs of production or supply. Economists distinguish between three types of price discrimination.

1. "**First degree price discrimination** [*emphasis added*] occurs when a firm is able to discriminate perfectly between its customers" (Bishop and Walker 2010: p. 251). First degree price discrimination is virtually never possible because firms can hardly measure the willingness to pay of each single customer. Even if they could this pricing strategy might result in arbitrage, i.e. customers with a low willingness to pay would buy the good cheaply and resell it to customers with a high(er) willingness to pay at prices lower than the one that the supplier would charge.
2. "**Second degree price discrimination** [*emphasis added*] occurs when certain selling practices are used to induce consumers to self-select themselves to reveal whether they have a high or low willingness to pay. The most common form of second degree price discrimination are volume discounts. Another common form is the use of two part tariffs where consumers pay a lump sum up front and then a per usage charge thereafter. This type of pricing is common in network markets such as telecoms, water, gas and electricity. As a consumer uses the service more (e.g. uses more electricity), the average price per unit

declines. Many rebates and discounts are based on volumes purchased and so are a form of second degree price discrimination" (Bishop and Walker 2010: p. 251).

3. "**Third degree price discrimination** [*emphasis added*] occurs when firms use information about their consumers (e.g. age or location) to price discriminate. Common examples are firms charging less to children or old age pensioners (e.g. buses and cinemas) or to those willing to travel at off-peak times (e.g. trains and buses)" (Bishop and Walker 2010: p. 251).

As first degree price discrimination is hardly ever observed, competition policy is concerned with second and third degree price discrimination.

In order to evaluate practices of exclusive dealing one must analyze the **welfare effects of price discrimination**, which can be ambiguous. In this context, note that in its assessment, the Commission will also consider claims of the dominant firm that rebate systems achieve cost or other **efficiencies** which are passed on to consumers. It can easily be shown that **third degree price discrimination** is welfare enhancing when it allows the firm "to supply a group of consumers that would not otherwise be supplied. For example, consider a firm that sells branded goods in a low income country at much lower prices than it sells those same goods in a high income country. It may be that if a firm has to charge a uniform price to two groups of consumers with different levels of willingness to pay, the group with a low willingness to pay would make no purchases" (Bishop and Walker 2010: p. 253).

To see this, consider a simple model of third degree price discrimination that is defined by three key features (Pepall et al. 2008: p. 91). First, prices are discriminated according to some observable characteristic (for example, age, location, education status). Second, the supplier can prevent arbitrage across the groups. Third, the supplier charges the same price to all consumers within a particular group. In our example, let the observable characteristic be location in a poorer (B) or in a richer (A) country with a different willingness to pay and demand functions

$$\begin{aligned} Q_A &= a - p_A \\ Q_B &= b - p_B \quad \text{with} \quad b < a \end{aligned} \qquad (162)$$

Moreover, let the supplier be a monopolist in both markets that produces at constant marginal costs c without fixed costs.

When the monopolist can price-discriminate it charges the profit maximizing price in each market.

$$\begin{aligned} p^*_A &= (a+c)/2 = c + (a-c)/2 \\ p^*_B &= (b+c)/2 = c + (b-c)/2 \end{aligned} \qquad (163)$$

This yields the following consumer surplus.

$$\begin{aligned} CS^*_A &= [(a-c)/2]^2 \\ CS^*_B &= [(b-c)/2]^2 \end{aligned} \tag{164}$$

When the monopolist must charge the same price in both regions, it charges a uniform price that maximizes its joint profits.

$$\begin{aligned} max\,\pi &= (p-c)\cdot[(a-p)+(b-p)] \\ p_u &= 0.5\cdot\left[\frac{a+c}{2}+\frac{b+c}{2}\right] \end{aligned} \tag{165}$$

This is the (weighted) average of the prices that the monopoly would charge in case of price discrimination. It is easy to check that banning price discrimination benefits the customers in the rich region A ($p_u<p^*_A$) and harms the consumers in the poor region B ($p_u>p^*_B$). Therefore, a prohibition of price discrimination can have **adverse distribution consequences**. For

$$c \geq (3b-a)/2 \tag{166}$$

the monopolist would set a price p_u at or above the maximum willingness to pay of the customers in region B so that the customers in B would not buy the product. As a consequence, the monopolist would decide to serve market A only and charge p^*_A there. This shows that the banning of price discrimination can have welfare deteriorating effects. Therefore, a *per se* prohibition of price discrimination does not appear justified. This conclusion also applies to second degree price discrimination.

In an illustrative example on train travel, Bishop and Walker (2010: p. 254) also consider the effects of **fixed costs** *F*. Peak train travel may cost more than off-peak train travel ($p^*_A>p^*_B$), meaning that peak time travellers (group A) contribute more to the recovery of the fixed costs of running trains than do off-peak travellers (group B). In the above example, the margin above marginal costs (i.e. (*a*-*c*)/2 and (*b*-*c*)/2) can be used to cover per-unit fixed costs ($F/(Q^*_A+Q^*_B)$). If train companies had to charge a uniform price throughout the day, off-peak travellers might well switch to alternative modes of transport or not travel at all. This would mean that off-peak travel no longer make any contribution to fixed costs and so peak time travellers would have to shoulder all the burden as per-unit fixed costs would rise to F/Q^*_A. This might even lead to higher prices for peak time travellers if average total costs $ATC=c+F/Q^*_A$ are above the profit-maximizing price p^*_A.

So far, we have only been concerned with showing that price discrimination and, thus, rebate schemes can have **pro-competitive efficiency effects**. We have not discussed the effect of price discrimination on competitors, yet. In particular, rebates may have **foreclosure effects** by enabling the dominant firm "to use the 'non contestable' portion of the demand of each customer (that is to say, the amount that would be purchased by the customer from the dominant undertaking in any event) as leverage to decrease the price to be paid for the 'contestable' portion of demand (that is to

say, the amount for which the customer may prefer and be able to find substitutes). In general terms, **retroactive rebates** [*emphasis added*] may foreclose the market significantly, as they may make it less attractive for customers to switch small amounts of demand to an alternative supplier, if this would lead to loss of the retroactive rebates. The potential foreclosing effect of retroactive rebates is in principle strongest on the last purchased unit of the product before the threshold is exceeded" (EU 2008: paras. 39 and 40).

To see this, consider the following example. A supplier produces at constant marginal costs c=1 and does not incur fixed costs. The firm sells its good at price p=2 as long as a buyer purchases less than a quantity of Q_d=10,000 units. If a customer purchases at least Q_d, the seller grants a **retroactive rebate** of 20% and charges a price of p_d=1.6 for the entire quantity sold. Figure 32 Shows that at Q_d the customer's expenditure drops from E_{d-}=9,999·2=19,998 to E_{d+}=10,000·1.6=16,000. However, we see that already for any quantity $Q_t \geq$16.000/2=8.000 competition would be eliminated as a customer would not switch suppliers. This is because it would get the quantity Q_d-Q_t "for free" once it purchases at least Q_d. This is the **loyalty-enhancing effect** (or suction effect) of such rebate schemes.

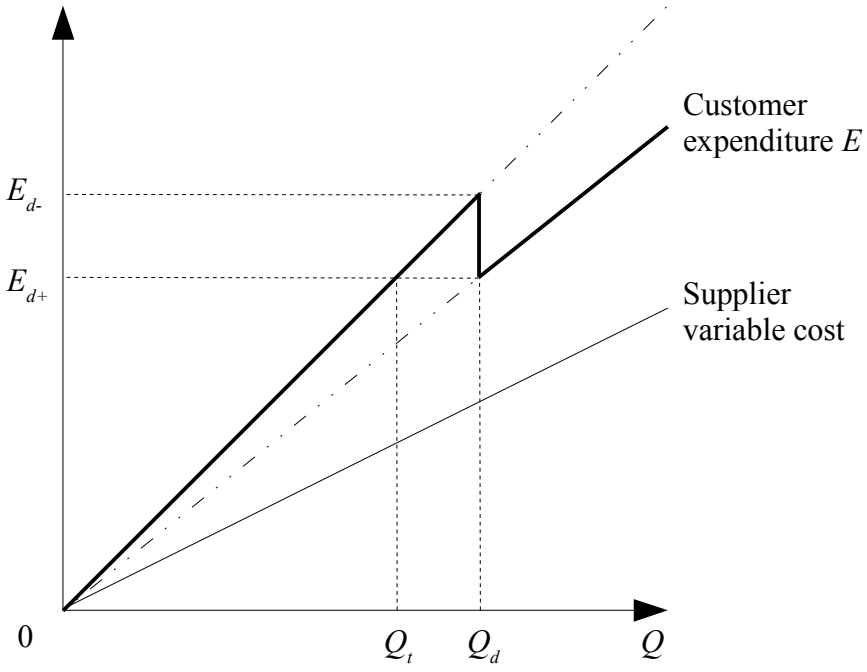


Figure 32: Incentive properties of loyalty rebate schemes

"The **Article [102] case law** [*emphasis added*] can fairly be characterised as being extremely hostile towards any dominant firm offering a loyalty rebate scheme. [...] In *Hoffmann-La Roche* the European Court of Justice (ECJ held that *Hoffmann-La Roche* had abused its dominant position both by entering into exclusive purchasing agreements with some of its customers and also by offering loyalty rebates. The ECJ distinguished standardised volume rebates and loyalty rebates

by stating that the former are discounts linked solely to the volume of purchases while the latter do not depend on quantities fixed objectively and applicable to all possible purchasers. The ECJ considered that, because they have the objective of increasing the dominant firm's share of a customer's purchasers rather than being related to the size of that purchase, loyalty rebate schemes can be considered to prevent customers from obtaining their supplies from competitors. The stance taken in this case towards loyalty rebates has led the Commission to argue that loyalty rebates are necessarily exclusionary when implemented by a firm held to be dominant unless the offered discounts reflect genuine cost savings associated with additional sales. The case law in this area has therefore developed with no regard being given to whether the competitors can match the offers or whether it is possible for such loyalty rebate schemes to foreclose a sufficient part of the market to reduce the competitive threat offered by competitors or whether consumer harm is likely" (Bishop and Walker 2010: pp. 260-261).

In analyzing whether rebates are likely to result in foreclosure – respectively hindering entry or expansion of firms that are as efficient as the dominant firm – the Commission has to estimate the **effective price** that a competitor would have to set such that a relevant part of customers would switch to the competitor. In the above example, suppose that customer C has already bought 8.000 units from company A that uses the outlined rebate scheme. Buying one additional unit from its current supplier A can be done for free as long as customer C in total buys at least 10,000 units. Therefore, an alternative supplier B must at maximum charge an effective price of 0 to sell one unit to C. If C has already bought 8,001 units, B must charge an effective price for the 8,002nd unit of at maximum -2. Hence, it must compensate customer C for having paid 2 monetary units to A for the 8,001st unit while now "purchasing" from B. This compensation must be the larger the more units C has already bought from A. However, if C had not bought any unit from supplier A so far (and wanted to buy at least 10,000 units), competitor B would only have to supply its good at an effective price of 1.6 monetary units. The effective price will c.p. be the lower the more customers have already bought a significant quantity from supplier A, respectively if they would always buy some specific quantity (i.e. an **assured base**) from supplier A.

As there is typically a variety of customers that have already bought different amounts of the good from supplier A, the Commission is concerned with estimating the effective price of B such that a relevant part of customers would switch to competitor B. The lower the estimated effective price is compared to the average price of the dominant firm, the stronger is the **loyalty-enhancing effect** of rebates. "However, as long as the effective price remains consistently above the [average total costs *ATC*] of the dominant undertaking, this would normally allow an equally efficient competitor to compete profitably notwithstanding the rebate. In those circumstances the rebate is normally not capable of foreclosing in an anti-competitive way. Where the effective price is below

[marginal costs *MC*], as a general rule the rebate scheme is capable of foreclosing even equally efficient competitors. Where the effective price is between [*MC*] and [*ATC*], the Commission will investigate whether other factors point to the conclusion that entry or expansion even by equally efficient competitors is likely to be affected. In this context, the Commission will investigate whether and to what extent competitors have realistic and effective counterstrategies at their disposal" (EU 2008: para. 43 and 44).

Rebate schemes can also have **pro-competitive effects** that can be categorized as follows (Bishop and Walker 2010: p. 264).

- **Providing incentives for customers to supply complementary services** – "Customers can add substantial value to a supplier's products or services by providing additional complementary services. These complementary services include promoting the product in store, providing detailed product information to customers (sometimes including a demonstration of how to use the products or services), keeping an appropriate stock of product so that at any time final consumers are able to purchase the product they require, or simply putting more effort into selling the products or services of the supplier. These services create benefits for the supplier while their costs mainly accrue to the supplier's customers" (Bishop and Walker 2010: p. 264). Hence, rebates lower the customer's costs and contribute to supply such complementary services.
- **Inducing customers to lower prices to end consumers** – Lower wholesale prices can reduce the problem of double marginalization.
- **Efficient fixed cost recovery** – Rebates contribute to selling a higher quantity of a good. Therefore, per-unit fixed costs decrease which may lead to lower prices (see above).

"In many important applications it is not sufficient to evaluate the welfare effects of price discrimination by comparing prices, production, and consumer surplus to those associated with uniform pricing within the framework of a given market structure [as in the above examples]. The option of price discrimination may affect entry (or exit) decisions and long-run investment decisions in ways that may significantly affect the long-run performance of the industry" (EAGCP 2005: 33).

Exclusionary Abuses – Refusal to Supply and Margin Squeeze

As has been argued in sections G.3 and H.4 , competition problems may arise in vertical relationships when the dominant undertaking competes on the downstream market with a buyer whom it **refuses to supply**. "The concept of refusal to supply covers a broad range of practices, such as a refusal to supply products to existing or new customers, refusal to license intellectual

property rights [...], or refusal to grant access to an essential facility or a network [see section E.2]" (EU 2008: para. 78). The main competitive concerns regarding a refusal to supply are a reduction of firms' incentives to invest or innovate, which harms competition and ultimately consumers.

Instead of refusing to supply a dominant firm may also engage in **margin squeeze**. This term describes a conduct where a dominant firm charges "a price for the product on the upstream market which, compared to the price it charges on the downstream market, does not allow even an equally efficient competitor to trade profitably in the downstream market on a lasting basis" (EU 2008: para. 80). This case has been illustrated on pp. 179-182 where a vertically integrated monopolist faces an equally efficient downstream competitor. It is shown that by charging different prices to its downstream branch and the independent competitor, the latter can be driven out of the market. This benefits the vertically integrated firm and harms consumers.

The Commission considers these practices as an enforcement priority if all the following circumstances are present (EU 2008: para. 81).

(a) The refusal relates to a product or service that is objectively necessary to be able to compete effectively on a downstream market (**objective necessity of the input**). "[A]n input is indispensable where there is no actual or potential substitute on which competitors in the downstream market could rely so as to counter – at least in the long-term – the negative consequences of the refusal. In this regard, the Commission will normally make an assessment of whether competitors could effectively duplicate the input produced by the dominant undertaking in the foreseeable future" (EU 2008: para. 83).

(b) The refusal is likely to lead to the **elimination of effective competition** on the downstream market. "The likelihood of effective competition being eliminated is generally greater the higher the market share of the dominant undertaking in the downstream market. The less capacity-constrained the dominant undertaking is relative to competitors in the downstream market, the closer the substitutability between the dominant undertaking's output and that of its competitors in the downstream market, the greater the proportion of competitors in the downstream market that are affected, and the more likely it is that the demand that could be served by the foreclosed competitors would be diverted away from them to the advantage of the dominant undertaking." (EU 2008: para. 85).

(c) The refusal is likely to lead to **consumer harm**. In this context, "the Commission will examine whether, for consumers, the likely negative consequences of the refusal to supply in the relevant market outweigh over time the negative consequences of imposing an obligation to supply" (EU 2008: para. 86).

These points shall be illustrated for the analysis of an **essential facility** which can be defined

as a facility or infrastructures – i.e. an input purchased in an upstream market – that is indispensable for providing services to customers in the downstream market. The existence of an essential facility raises competitive concerns when its (monopolistic) owner uses its market power to prevent competition downstream where it also provides services, i.e. it forcloses efficient competitors or prevents entry by other efficient firms. Similarly, essential facilities or bottlenecks can exist in the downstream market, and the owner uses the essential facility to leverage market power from this downstream market into the upstream market. Following the above framework, one should consider five economic conditions that an asset should satisfy before it may be considered an essential facility (Bishop and Walker 2010: pp. 327-329).

1. It must be **impossible** or at least uneconomic for any other firm **to replicate the asset**, i.e. it is not possible for another firm to develop a similar facility that could supply the product or service at a low enough cost to supply the downstream market at a competitive price. Otherwise, the other firm should carry out its own investment in such an asset.
2. There should be **no alternative means of entering the relevant market** at a reasonable cost. If there are alternative methods of entering a market, then mandating access to an asset is not necessary for the protection of competition.
3. There must be **spare capacity** on the asset in question. If there is no spare capacity on the asset in question, then mandating access does not increase the level of competition in the market.
4. There must be a **lack of effective competition downstream** and a reasonable expectation that mandating access will significantly improve the level of downstream competition.
5. The owner of the asset must compete in the **same relevant market** as the entrant wishes to compete in. This is more a reality check than an additional condition. If the new entrant is not going to compete directly with the asset owner, then the asset owner has no incentive to refuse supply if the new entrant is prepared to offer a reasonable price for access.

Exclusionary Abuses – Tying and Bundling

"**'Tying'** [*emphasis added*] usually refers to situations where customers that purchase one product [A] (the tying product) are required also to purchase another product [B] from the dominant undertaking (the tied product). [I.e. while B could also be bought separately, A can only be bought together with B.] Tying can take place on a technical or contractual basis. **'Bundling'** [*emphasis added*] usually refers to the way products are offered and priced by the dominant undertaking. In the case of **pure bundling** [*emphasis added*] the products are only sold jointly in fixed proportions. In the case of **mixed bundling** [*emphasis added*], often referred to as a multi-product rebate, the

products are also made available separately, but the sum of the prices when sold separately is higher than the bundled price" (EU 2008: para. 48). Note that selling the bundle of products at a lower price than the sum of prices to be paid when selling the products separately is also some sort of rebate (see above).

"Tying and bundling are common practices intended to provide customers with better products or offerings in more cost effective ways. However, an undertaking which is dominant in one product market (or more) of a tie or bundle (referred to as the tying market) can **harm consumers** through tying or bundling **by foreclosing the market** [*emphasis added*] for the other products that are part of the tie or bundle (referred to as the tied market) and, indirectly, the tying market" (EU 2008: para. 49). Hence, market power in one market may be leveraged into the other market. "The key condition needed for these actions to cause competitive harm, is that the linkage must place some rivals at a competitive disadvantage so that they cannot compete effectively in the adjacent market or so that they might be deterred from competing in the home market" (EAGCP 2005: 23).

Tying and bundling may also have pro-competitive efficiency effects and is even employed by non-dominant firms. "The problem is to identify cases where tying is anticompetitive, that is, profitable for the firm that implements the practice, while inducing exclusion and hurting consumers" (EAGCP 2005: 39). The "Commission will normally take action under Article [102] where an undertaking is **dominant in the tying market** [*emphasis added*] and where, in addition, the following conditions are fulfilled: (i) the tying and tied products are distinct products, and (ii) the tying practice is likely to lead to anti-competitive foreclosure" (EU 2008: para. 50).

"Whether the products will be considered by the Commission to be distinct depends on customer demand. **Two products are distinct** [*emphasis added*] if, in the absence of tying or bundling, a substantial number of customers would purchase or would have purchased the tying product without also buying the tied product from the same supplier, thereby allowing **stand-alone production** [*emphasis added*] for both the tying and the tied product" (EU 2008: para. 51). However, there should be some **demand complementarity** between products. If the products are completely distinct from customers' point of view and a firm engages in mixed bundling, no customer would buy the bundle so that bundling would not have an effect on competition. However, if the firm engaged in pure bundling customers would have to buy both products although wanting only one of them. This is likely to reduce demand for the bundle causing bundling to be unprofitable for the firm.

"Tying or bundling may lead to **anti-competitive effects** [*emphasis added*] in the tied market, the tying market, or both at the same time. [...] The risk of anti-competitive foreclosure is

expected to be greater where the dominant undertaking makes its tying or bundling strategy a lasting one, for example through **technical tying** [*emphasis added*] which is costly to reverse. Technical tying also reduces the opportunities for resale of individual components" (EU 2008: paras. 52 and 53). For a more detailed treatment of foreclosure, refer to sections G.3 and H.4 .

"A **multi-product rebate** [*emphasis added*] may be anti-competitive on the tied or the tying market if it is so large that equally efficient competitors offering only some of the components cannot compete against the discounted bundle" (EU 2008: para. 59).

The Commission may consider the following types of **efficiencies**:

1. The Commission may consider whether such practices **reduce transaction costs for customers**, who would otherwise be forced to buy the components separately, and enable substantial savings on packaging and distribution costs for suppliers. This point also applies to complex products where one component has to work together with another component. Selling these products together ensures that the the products are compatible to each other and that this reduces customers' transaction costs of finding compatible products.

2. It may also examine whether combining two independent products into a new, single product might **enhance the ability to bring such a product to the market** to the benefit of consumers. To see this, suppose a monopolist produces good A at marginal costs c_A=2 while demand for A is q_A=18-p_A. It is easy to show that the monopolist would choose an optimal price p_A*=10. Moreover, assume the monopolist also produces a good B at marginal costs c_B=4 with demand for B being q_B=16-p_B. The monopolist would choose the optimal price p_B*=10. In addition to these customers who would only buy good A or good B but not both goods together, there are customers who would always want to buy good A and good B together. Let the demand of these customers be given by q_{A+B}=24-p_{A+B}. If these customers have to pay p_{A+B}'=20, the firm makes profits π_{A+B}'=56 in this market with consumer surplus being CS_{A+B}'=8. It can be shown that by offering a bundle A+B, which is produced at marginal costs c_{A+B}=6, the monopolist would optimally set p_{A+B}*=15. This results in profits π_{A+B}*=81 and consumer surplus CS_{A+B}*=40.5. As a consequence, bundling may raise total welfare without causing any harm.

3. The Commission may also consider whether tying and bundling practices allow the supplier to **pass on efficiencies** arising from its production or purchase of large quantities of the tied product. For example, suppose a tying good (e.g. a printer) can only be used with a specific good (e.g. ink cartridges of the same producer), and the latter's production is subject to economies of scale. Then tying contributes to exploiting these economies of scale which can be passed on to consumers via lower prices.

Aftermarkets

"In many industries, the initial purchase decision of consumers can have long-run effects on their future choices. This occurs where consumers purchase durable products that also require the purchase of some comp1ementary products, at least some of which are purchased at a later date than the purchase of the durable product. There are numerous industries in which this is the case" (Bishop and Walker 2010: p. 245).

Industry	**Primary Product**	**Secondary Product**
Video games	Games console	Video game
Cars	Car	Spare car parts
Computers	Hardware	Hardware and software maintenance services
Printers	Printer	Toner cartridges
DVDs	DVD player	DVDs
Mobile telephony	Mobile phone/network	Mobile telephony calls

Source: Bishop and Walker (2010: p. 245)

"The peculiar competitive feature of these types of industry lies in the competitive interaction between the **primary durable product** [*emphasis added*] and the secondary or "aftermarket" for associated **complementary products or services** [*emphasis added*]. Often due to technical differences between the durable primary products, the choice of complementary products compatible with a particular primary product is limited. This implies that once the primary product has been purchased, consumer choice is confined to those aftermarket products or services compatible with that primary product. For example, the owner of a Ford motor car needs to purchase spare parts which are compatible with that type of car. [...] In other words, consumers are to a greater or lesser extent locked in to certain aftermarket suppliers."

In this context, recall from section E.2 that **switching costs** may arise from the need for compatibility with existing equipment. There it is also shown that the **lock-in effect** that is created by such switching costs is a source of market power. This may cause excessively high prices in the aftermarket. "Whether the price of secondary products or services can be raised above the effectively competitive level will depend on the extent to which **consumers will be deterred** [*emphasis added*] either from purchasing secondary services or in making the initial primary purchasing decision. A firm selling in both the primary and secondary markets faces a trade-off when it sets its aftermarket price. A higher price will allow it to earn more profits on aftermarket sales to consumers who have already purchased the equipment. On the other hand, a higher

aftermarket price will also reduce sales of the equipment, because potential buyers will take into account this higher expected cost of purchasing the associated aftermarket products. If high prices in the aftermarket deter a sufficient number of consumers from making primary product purchases, then anti-competitive pricing can be deterred. However, where the strength of this competitive constraint is weak – perhaps due to consumers having **incomplete information** [*emphasis added*] about future costs in the aftermarket – anti-competitive practices in the aftermarket could be profitable" (Bishop and Walker 2010: p. 246). Whether firms may exercise market power in aftermarkets depends on the following characteristics of an industry:

1. **The ratio of "locked-in" consumers to new purchasers** – The extent of market power is likely to be the smaller the higher the fraction of new consumers in the entire consumer base. This is e.g. the case in new and/or growing markets. In such markets the competition for the primary market may be intense (see the model on p. 111) causing low prices at least in this market segment. However, the competition for new consumers may also limit the extent to which market power can be exercised in the aftermarket.

2. **The ability to price discriminate in favor of new purchasers** – "The ability to price discriminate in the secondary market in favour of new consumers facilitates anti-competitive pricing in aftermarket products, since low competitive prices could be charged to new consumers to encourage them to enter the market and higher prices charged to existing consumers once they face costs of switching to other primary products. However, the circumstances in which such price discrimination can take place are likely to be fairly limited. As casual empirical observation usually shows, the price of secondary products or services is often the same to potential new consumers as to consumers who have already committed themselves to a particular system through the purchase of a primary product" (Bishop and Walker 2010: p. 247).

3. **The magnitude of switching costs** – Firms' ability to exercise market power is c.p. the higher the larger the switching costs of the customers. However, switching costs can vary over time or across customers and, thus, imply different degrees to which excessive prices can be set. For example, the switching costs of a customer who has just bought a new car may be assumed to be higher than those of the owner of an older car which is going to be replaced in the near future.

4. **The quality of information available to marginal consumers** – The extent to which market power can be exercised in aftermarkets also depends on the question to what extent customers consider the price of the secondary product when making their purchase decision of the primary product. For example, a business customer who prints a lot and, thus, has to

buy toner cartridges relatively often is quite likely to collect information on the prices of cartridges when buying a printer. In contrast, a private customer who prints more infrequently might put less attention to the prices of cartridges and, hence, might be charged higher prices. This point should also be seen with the above argument on price discrimination. I.e. it is easier to exercise market power when it is possible to charge different prices from private and business customers.

5. **The number of markets in which the selling firm competes** – If a firm is active in many markets and charges excessive prices in at least one of them, it may develop a reputation for charging excessive prices in the other markets, as well. This "bad reputation" may constrain the market power that the firm can exercise at all.

6. **Strength of system competition** - "It is often not possible to judge the degree of competition in the secondary market without also taking account of the degree of competition in the primary market. Consumers typically want to purchase the bundle of services provided by the combination of the primary product and the secondary product and so it is the total cost of these that matters. In such a framework, it is not uncommon for manufacturers to earn a relatively small (or even negative) margin on the primary product and a relatively larger margin on the secondary product (e.g. mobile phone handsets and calls). It has been argued that this can be a highly efficient way for firms to operate. For instance, the lower price of the primary product may make it easier to introduce new products to the market as consumers are more likely to be willing to "take a risk" when prices are lower" (Bishop and Walker 2010: p. 249).

"The Commission has investigated a number of industries with this aftermarket feature. [...] In Hugin/Liptons, the Commission found that Hugin, despite having only a 12 per cent share of the European cash register market, enjoyed a dominant position in respect of spare parts for its cash registers. The spare parts for Hugin cash registers were not interchangeable with those of other cash registers. Hugin cash registers could not therefore be maintained, repaired or rebuilt without Hugin parts. The European Court of Justice confirmed that Hugin enjoyed a dominant position in relation to those who owned a Hugin cash register" (Bishop and Walker 2010: p. 246).

Exclusionary Abuses – Predation

Predation refers to a situation where a firm deliberately engages in a behavior where it incurs losses or sacrifices profits in the short term so as to foreclose one or more of its actual or potential competitors. This behavior is aimed at strengthening the predatory firm's market power by inducing firms to exit the market or deterring new firms from entering, which harms consumers. Therefore,

one should note the following points (Bishop and Walker 2010: p. 291).

1. The predating firm should enjoy **substantial market power**. Without dominance the exclusion of other firms does not sufficiently weaken competition such as to recoup the losses caused by this behavior. I.e. predation pays off for the dominant firm if it can reasonably expect that "its market power after the predatory conduct comes to an end to be greater than it would have been had the undertaking not engaged in that conduct in the first place, that is to say, if the undertaking is likely to be in a position to benefit from the sacrifice" (EU 2008: para. 70). Here, it is not sufficient that market power somewhat increases but is raised so strongly that the long-run increase in profits exceeds the short-run sacrifice of profits.

2. Entry by firms into a market is generally accommodated by the incumbent(s) setting lower prices. Therefore, it needs to be demonstrated that this pricing conduct has involved a **sacrifice in short-run profits**. In particular, **pricing below marginal costs** will (in most cases) be viewed by the Commission as a clear indication of predation (Areeda-Turner test: comparing observed prices with a cost benchmark). Defined more widely, predation also refers to a situation where a firm engages in a strategy that causes it to make net revenues lower than could have been expected from a reasonable alternative conduct. In this case, only **pricing below average total costs** is capable of foreclosing competitors from the market that are as efficient as the dominant undertaking. In determining the appropriate cost benchmark it is the costs of the dominant firm that are relevant. This is because driving inefficient firms out of the market may contribute to overall productive efficiency.

3. The predatory conduct should result in the **foreclosure of** (as efficient) **competitors**. This requires that the predator access to greater financial resources ('deep pockets') than its prey. This assumption is difficult to satisfy especially when the prey is more efficient than the predating firm because banks could lend money to the prey and wait until the predator runs out of its resources and must leave the market itself.

4. After competitors have been forced out of the market or have been prevented from entering market conditions must be such that these firms do not (re-)enter when the predating firm raises its price to **recoup the short-run losses**. In other words, a credible predatory strategy requires there to be **barriers to entry**. Note that the more costly it is to foreclose rivals the less likely it is that long-run (excessive) profits outweigh the short-run loss. Also note that the recoupment of short-run losses is easier when the predator enjoys a high degree of market power after the foreclosure (see point 1 above).

5. In general, predatory behavior is unlikely to create **efficiencies**.

The question now is how predation works. With regard to the above **deep pocket** argument, capital market imperfections and asymmetric information may prevent banks to provide unlimited funds to the prey. This is e.g. the case when it is uncertain that the prey is indeed the more efficient firm. With limited funds the basic reasoning applies, i.e. a predating firm can foreclose its competitors if its funds are greater than those of its prey.

Low, predatory prices may also be used as **signaling devices**. When firms offer differentiated products (see the model in section G.2) they may set dissimilar prices. In particular, if demand for all product varieties is the same but firms' marginal costs of production are not identical, high-cost firms would prefer to set higher prices than low-cost firms. The latter would set lower prices and win a higher share of the market. However, if firms cannot (perfectly) observe their rivals' cost but their prices they may use prices as signaling devices for their productive efficiency. As a consequence, a high-cost incumbent might choose low prices to send the (misleading) signal to potential entrants of being an efficient and competitive low-cost firm. This may prevent the low cost firms from entering the market and leave the high-cost incumbent in a comfortable position.

Predatory behavior in one market may also be used as a threat in other markets. To see this, suppose a dominant firm is active in a variety of different local markets. Once a competitor enters in one of these markets the incumbent engages in predatory behavior even if this is associated with losses in this market. However, by doing so the incumbent builds up a **reputation** of being a tough competitor which prevents rivals from entering into the other regional markets where the incumbent may still enjoy its dominant position. “A possible example of aggressive responses to entry designed to create a reputation for predatory behaviour is provided by General Foods in the United States in the 1970s. General Foods was the producer of Maxwell House Coffee. At the time General Foods had a market share in the eastern states of about 45 per cent. When Folger's, a brand based in the western states, tried to enter various eastern markets General Foods responded with sharp price decreases in those markets where Folger's had entered. This strategy successfully discouraged Folger's from further entry in the eastern states” (Bishop and Walker 2010: p. 299).

§§ 19-21 GWB

§ 19 Missbrauch einer marktbeherrschenden Stellung

(1) Die missbräuchliche Ausnutzung einer marktbeherrschenden Stellung durch ein oder mehrere Unternehmen ist verboten.

(2) Ein Unternehmen ist marktbeherrschend, soweit es als Anbieter oder Nachfrager einer bestimmten Art von Waren oder gewerblichen Leistungen auf dem sachlich und

räumlich relevanten Markt

1. *ohne Wettbewerber ist oder keinem wesentlichen Wettbewerb ausgesetzt ist oder*

2. *eine im Verhältnis zu seinen Wettbewerbern überragende Marktstellung hat; hierbei sind insbesondere sein Marktanteil, seine Finanzkraft, sein Zugang zu den Beschaffungs- oder Absatzmärkten, Verflechtungen mit anderen Unternehmen, rechtliche oder tatsächliche Schranken für den Marktzutritt anderer Unternehmen, der tatsächliche oder potentielle Wettbewerb durch innerhalb oder außerhalb des Geltungsbereichs dieses Gesetzes ansässige Unternehmen, die Fähigkeit, sein Angebot oder seine Nachfrage auf andere Waren oder gewerbliche Leistungen umzustellen, sowie die Möglichkeit der Marktgegenseite, auf andere Unternehmen auszuweichen, zu berücksichtigen.*

Zwei oder mehr Unternehmen sind marktbeherrschend, soweit zwischen ihnen für eine bestimmte Art von Waren oder gewerblichen Leistungen ein wesentlicher Wettbewerb nicht besteht und soweit sie in ihrer Gesamtheit die Voraussetzungen des Satzes 1 erfüllen. Der räumlich relevante Markt im Sinne dieses Gesetzes kann weiter sein als der Geltungsbereich dieses Gesetzes.

(3) Es wird vermutet, dass ein Unternehmen marktbeherrschend ist, wenn es einen Marktanteil von mindestens einem Drittel hat. Eine Gesamtheit von Unternehmen gilt als marktbeherrschend, wenn sie

1. *aus drei oder weniger Unternehmen besteht, die zusammen einen Marktanteil von 50 vom Hundert erreichen, oder*

2. *aus fünf oder weniger Unternehmen besteht, die zusammen einen Marktanteil von zwei Dritteln erreichen,*

es sei denn, die Unternehmen weisen nach, dass die Wettbewerbsbedingungen zwischen ihnen wesentlichen Wettbewerb erwarten lassen oder die Gesamtheit der Unternehmen im Verhältnis zu den übrigen Wettbewerbern keine überragende Marktstellung hat.

(4) Ein Missbrauch liegt insbesondere vor, wenn ein marktbeherrschendes Unternehmen als Anbieter oder Nachfrager einer bestimmten Art von Waren oder gewerblichen Leistungen

1. *die Wettbewerbsmöglichkeiten anderer Unternehmen in einer für den Wettbewerb auf dem Markt erheblichen Weise ohne sachlich gerechtfertigten*

Grund beeinträchtigt;

2. *Entgelte oder sonstige Geschäftsbedingungen fordert, die von denjenigen abweichen, die sich bei wirksamem Wettbewerb mit hoher Wahrscheinlichkeit ergeben würden; hierbei sind insbesondere die Verhaltensweisen von Unternehmen auf vergleichbaren Märkten mit wirksamem Wettbewerb zu berücksichtigen;*

3. *ungünstigere Entgelte oder sonstige Geschäftsbedingungen fordert, als sie das marktbeherrschende Unternehmen selbst auf vergleichbaren Märkten von gleichartigen Abnehmern fordert, es sei denn, dass der Unterschied sachlich gerechtfertigt ist;*

4. *sich weigert, einem anderen Unternehmen gegen angemessenes Entgelt Zugang zu den eigenen Netzen oder anderen Infrastruktureinrichtungen zu gewähren, wenn es dem anderen Unternehmen aus rechtlichen oder tatsächlichen Gründen ohne die Mitbenutzung nicht möglich ist, auf dem vor- oder nachgelagerten Markt als Wettbewerber des marktbeherrschenden Unternehmens tätig zu werden; dies gilt nicht, wenn das marktbeherrschende Unternehmen nachweist, dass die Mitbenutzung aus betriebsbedingten oder sonstigen Gründen nicht möglich oder nicht zumutbar ist.*

§ 20 Diskriminierungsverbot, Verbot unbilliger Behinderung

(1) Marktbeherrschende Unternehmen, [...], dürfen ein anderes Unternehmen in einem Geschäftsverkehr, [...], weder unmittelbar noch mittelbar unbillig behindern oder gegenüber gleichartigen Unternehmen ohne sachlich gerechtfertigten Grund unmittelbar oder mittelbar unterschiedlich behandeln.

(2) Absatz 1 gilt auch für Unternehmen und Vereinigungen von Unternehmen, soweit von ihnen kleine oder mittlere Unternehmen als Anbieter oder Nachfrager einer bestimmten Art von Waren oder gewerblichen Leistungen in der Weise abhängig sind, dass ausreichende und zumutbare Möglichkeiten, auf andere Unternehmen auszuweichen, nicht bestehen. [...]

(3) Marktbeherrschende Unternehmen und Vereinigungen von Unternehmen im Sinne des Absatzes 1 dürfen ihre Marktstellung nicht dazu ausnutzen, andere Unternehmen im Geschäftsverkehr dazu aufzufordern oder zu veranlassen, ihnen ohne sachlich gerechtfertigten Grund Vorteile zu gewähren. Satz 1 gilt auch für Unternehmen und Vereinigungen von Unternehmen im Verhältnis zu den von ihnen

abhängigen Unternehmen.

(4) Unternehmen mit gegenüber kleinen und mittleren Wettbewerbern überlegener Marktmacht dürfen ihre Marktmacht nicht dazu ausnutzen, solche Wettbewerber unmittelbar oder mittelbar unbillig zu behindern. Eine unbillige Behinderung im Sinne des Satzes 1 liegt insbesondere vor, wenn ein Unternehmen

1. *Lebensmittel im Sinne des § 2 Abs. 2 des Lebensmittel- und Futtermittelgesetzbuches unter Einstandspreis oder*
2. *andere Waren oder gewerbliche Leistungen nicht nur gelegentlich unter Einstandspreis oder*
3. *von kleinen oder mittleren Unternehmen, mit denen es auf dem nachgelagerten Markt beim Vertrieb von Waren oder gewerblichen Leistungen im Wettbewerb steht, für deren Lieferung einen höheren Preis fordert, als es selbst auf diesem Markt*

anbietet, es sei denn, dies ist jeweils sachlich gerechtfertigt. Das Anbieten von Lebensmitteln unter Einstandspreis ist sachlich gerechtfertigt, wenn es geeignet ist, den Verderb oder die drohende Unverkäuflichkeit der Waren beim Händler durch rechtzeitigen Verkauf zu verhindern sowie in vergleichbar schwerwiegenden Fällen. Werden Lebensmittel an gemeinnützige Einrichtungen zur Verwendung im Rahmen ihrer Aufgaben abgegeben, liegt keine unbillige Behinderung vor.

(5) Ergibt sich auf Grund bestimmter Tatsachen nach allgemeiner Erfahrung der Anschein, dass ein Unternehmen seine Marktmacht im Sinne des Absatzes 4 ausgenutzt hat, so obliegt es diesem Unternehmen, den Anschein zu widerlegen [...].

(6) [...]

§ 21 Boykottverbot, Verbot sonstigen wettbewerbsbeschränkenden Verhaltens

(1) Unternehmen und Vereinigungen von Unternehmen dürfen nicht ein anderes Unternehmen oder Vereinigungen von Unternehmen in der Absicht, bestimmte Unternehmen unbillig zu beeinträchtigen, zu Liefersperren oder Bezugssperren auffordern.

(2) Unternehmen und Vereinigungen von Unternehmen dürfen anderen Unternehmen keine Nachteile androhen oder zufügen und keine Vorteile versprechen oder gewähren, um sie zu einem Verhalten zu veranlassen, das nach diesem Gesetz oder nach einer auf Grund dieses Gesetzes ergangenen Verfügung der Kartellbehörde

nicht zum Gegenstand einer vertraglichen Bindung gemacht werden darf.

(3) [...]

(4) [...]

Lessons Learned

After reading this section you should be able to answer the following questions.

1. Give a brief overview on the relevant laws concerning the abuse of a dominant position in Europe and Germany.
2. What two steps are required in the analysis of the alleged abuse of a dominant position?
3. What is meant by the terms *dominance*, *single dominance*, and *collective dominance*?
4. Which two main types of abuses do you know?
5. Explain why granting rebates can have anti-competitive consequences.
6. Does price discrimination always lower welfare? Explain your answer.
7. Why would a company tie or bundle its products?
8. Give some examples of primary and secondary markets (aftermarkets).
9. Explain why a margin squeeze may result in the foreclosure of competitors.

References

Bishop, S. and Walker, M. (2010). "The Economics of EC Competition Law." 3rd edition. Sweet and Maxwell: London

Economic Advisory Group on Competition Policy EAGCP (2005). "Report by the EAGCP – An Economic Approach to Article 82." http://ec.europa.eu/dgs/competition/economist/eagcp_july_21_05.pdf

European Commission (2008). "Communication from the Commission — Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings." Official Journal C 45/7. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52009XC0224%2801%29:EN:NOT

Motta, M. (2004). "Competition Policy – Theory and Practice." Cambridge University Press: Cambridge

J TWO-SIDED MARKETS

J.1 Introduction

A **two-sided market** is composed of one or more platforms that brings together two kinds of agents for a transaction, with prices being set for both sides. Examples of two-sided markets are shown in Table 5.

Market	**Platform**	**Side 1**	**Side 2**
Media	TV, newspaper	Advertisers	Readers
Credit cards	Visa, Mastercard, American Express	Retailers	Consumers
Video games, operating sytems	Xbox, Windows	Application developers	Users
Employment	Monster.de	Employers	Applicants
Dating	Elitepartner.de, parship.de, nightclubs	Women	Men
Online auctions, shopping malls	Ebay	Sellers	Buyers

Table 5: Two-sided markets

The defining characteristic of two-sided markets is that one group's benefit from joining a platform depends on the size of the other group that joins the platform. For example, a man's pleasure of going to a nightclub might rise with the number of women in this club, i.e. the two groups exercise **positive externalities** on each other. These externalities determine the prices charged from each group. In the above example, men might be charged the full entrance fee while women only have to pay a reduced fee. Alternatively, women might receive free drinks as an extra bonus. In other words, the positive externality that a high ratio of femal clubgoers provides to men allows the nightclub to exercise market power on men. Armstrong (2006: 669) shows that "positive cross-group externalities act to intensify competition and reduce platform profit. [...] To be able to compete effectively on one side of the market, a platform needs to perform well on the other side (and vice versa). This creates a downward pressure on the prices offered to both sides compared to the case where no externalities exist." These effects apply at least as long as the positive externalities do not contribute to lowering the number of competitors, leading to the extreme of a monopolist platform.

The importance of externalities shows that the analysis of two-sided markets is related to the analysis of network effects (see section E.2). The main disctinction criterion for two-sided markets is the existence of a market intermediary (or **platform**) whose pricing choices are analyzed.

Two-sided markets need to be distinguished from **one-sided markets** that are analyzed in the above sections. The difference can be illustrated by the example of Amazon (Rysman 2009: 126). Amazon acts on a one-sided market when selling new books. It buys the book at a wholesale price from the publisher and sells the book at a retail price to readers. This market is one-sided because there are no externalities between the publishers and the readers. Amazon acts on a two-sided market when it provides a platform for producers (or wholesalers) to sell their products to endconsumers. In this case the consumer pays a retail price to the wholesaler, who pays a fee to Amazon for being allowed to use this platform. In principle, it would also be possible for Amazon to charge the buyers for purchasing products over its platform. This example illustrates another aspect that is relevant in the context of two-sided markets. Amazon must be concerned of the **quality** of sellers on its platform. If the quality of their products and services is below standard this may have negative impacts on the perceived quality of the platform and the other vendors.

The **economics of two-sided markets** deals with exactly this question: "[R]esearch in two-sided markets explores choices by market intermediaries, particularly pricing, when there is some kind of interdependence or externality between groups of agents that the intermediary serves" (Rysman 2009: 126). The above example illustrates that it can be difficult to distinguish two-sided and one-sided markets as many one-sided markets have at least some two-sided characteristics. The relevant distinction criterion is how important two-sided issues are in determining market outcomes such as prices, welfare and profits.

In subsection J.2 , we show that **pricing** in two-sided markets does not only depend on demand and costs at either side of the market but also on how the participation on one side of the market affects participation on the other side. For example, a newspaper may be sold to (relatively price-elastic) readers below costs if this raises sales and, thus, makes advertising in this newspaper more valuable. This allows to charge a higher price from (relatively price-inelastic) advertising customers. These effects may even be more pronounced when various platforms compete for the agents on both market sides, for example, when newspapers compete for readers and advertisiers. "In a one-sided market, we can characterize the price–cost mark-up in terms of elasticity of demand and the marginal cost. But in a two-sided market, pricing decisions will also include the elasticity of the response on the other side and the mark-up charged to the other side" (Rysman 2009: 129).

The pricing effects also depend on whether agents single-home or multi-home. One speaks of **single-homing** when an agent uses only one platform. For example, many PC-users install only a

single operating system on all of their computers, or readers subscribe to only one newspaper. **Multi-homing** is when an agent uses two or more platforms. For example, a viewer may watch different TV channels, or firms advertise in a variety of newspapers. A market may have a single-homing side (e.g., all consumers have a single credit card) and a multi-homing side (e.g., all retailers accept various types of credit cards).

Such a situation is shown in the below model with **competitive bottlenecks**. This is important "because the intermediary can be viewed as a monopolist over access to members that do not use other intermediaries. Hence, firms compete aggressively on the side that uses a single network in order to charge monopoly prices to the other side that is trying to reach them (Armstrong, 2006). As a result, competition between platforms can have large price effects on the side of the market that uses a single platform and little or no effect on the side that uses multiple platforms. This result might explain why payment card pricing has increasingly favored consumers over time rather than merchants (for instance, with the rise of rewards programs), since consumers and not merchants typically use a single network and competition among card networks has become more important relative to competition between card networks and cash."

Further pricing issues arise when firms can price discriminate or when the dynamics of pricing is important. For example, nightclubs often price-discriminate by charging men a higher entrance fee than women are provide women with free drinks. A ban of price discrimination would make men better off while women would have to pay more. Regarding the dynamics of pricing one might think of penetration pricing where an intermediary lowers price early in the product life cycle and raising it after having established a customer base. Note that these issues are not unique to two-sided markets and are discussed in the context of one-sided markets in sections E.2 and I.3 . A further issue in this context is standardization (also see section H.3), i.e. is it in the interest of a platform to be compatible with other platforms? For example, a bank can open its ATM (automatic teller machines) to customers of other banks or not.

J.2 Economic Analysis

In the following, we present some elementary models for analyzing the economic properties of two-sided markets (Armstrong 2006). We start with the case of a single platform that the individuals in two groups of agents may join or not. We continue with the analysis of two platforms who compete for the individuals in two groups. These agents are assumed to single home.

Monopoly Platform

A market might be served by a monopoly platform. For example, the distance between two

shopping malls might be sufficiently large such that (some) customers always shop at the one or the other mall. However, in many cases the assumption of a monopolist platform is rarely satisfied.

In this setting, we start with describing the **demand model**, i.e. the features of the two groups of agents (indexed 1 and 2). The utility of an agent is defined by equation (167).

$$\begin{aligned} u_1 &= \alpha_1 n_2 - p_1 \\ u_2 &= \alpha_2 n_1 - p_2 \end{aligned} \tag{167}$$

The utility (u_1) of a group 1 agent depends on the number of group 2 agents on this platform (n_2). It rises if group 2 agents exercise a positive externality on group 1 agents ($\alpha_1>0$) and falls in case of a negative externality ($\alpha_1<0$). Utility u_1 falls with the uniform price p_1 charged from the agent for joining the platform. The number of agents who join a platform is determined by increasing functions ϕ_1 and ϕ_2 of utility.

$$\begin{aligned} n_1 &= \phi_1(u_1) \\ n_2 &= \phi_2(u_2) \end{aligned} \tag{168}$$

The platform incurs a per-agent **cost** of f_1 or f_2 for serving group 1 or 2. The platform makes a **profit** π of the following form.

$$\begin{aligned} \pi &= n_1 \cdot [p_1 - f_1] + n_2 [p_2 - f_2] \\ &= \phi_1(u_1) \cdot [\alpha_1 \phi_2(u_2) - u_1 - f_1] + \phi_2(u_2) \cdot [\alpha_2 \phi_1(u_1) - u_2 - f_2] \end{aligned} \tag{169}$$

Given that the **consumer surplus** of group i=1, 2 is

$$\nu_i(u_i) = n_i \cdot u_i \tag{170}$$

the condition

$$\frac{\partial \nu_i(u_i)}{\partial u_i} = n_1 = \phi_1(u_1) \tag{171}$$

applies. **Welfare** is defined by

$$\omega = \pi(u_1, u_2) + \nu_1(u_1) + \nu_2(u_2) \quad . \tag{172}$$

Maximizing profit function (169) with respect to u_1 and u_2 yields the following **profit-maximizing prices**.

$$\begin{aligned} p_1 &= f_1 - \alpha_2 n_2 + \frac{\phi_1(u_1)}{\phi_1'(u_1)} \\ p_2 &= f_2 - \alpha_1 n_1 + \frac{\phi_2(u_2)}{\phi_2'(u_2)} \end{aligned} \tag{173}$$

"Thus, the profit-maximizing price for group 1, say, is equal to the cost of providing service (f_1), adjusted downward by the external benefit of group 2 ($\alpha_2 n_2$), and adjusted upward by a factor

related to the elasticity of the group's participation" (Armstrong 2006: 672). In case of a positive externality ($\alpha_2>0$) it is possible that group 1 is being offered a subsidized service ($p_1<f_1$). The same is true if the price elasticity of demand of the customers is high, which translates into $\phi_1'<0$. "The analysis [... also applies] to software markets in which one type of software is required to create files in a certain format and another type is required to read such files. Often, the reading software is supplied for free, while the writing software needs to be paid for" (Armstrong 2006: 673). Note that the **welfare-maximizing** prices are given by

$$\begin{aligned} p_1 &= f_1-\alpha_2 n_2 \\ p_2 &= f_2-\alpha_1 n_1 \end{aligned} \quad . \tag{174}$$

Two-Sided Single Homing

In this model, we assume that there are two platforms (indexed by $i=A, B$), e.g. nightclubs, which are located at the ends of an imaginary road with distance 1. There are two groups of customers (indexed 1 and 2), e.g. men and women, who live along this road where they are distributed uniformly. Agents receive a gross utility from joining platform i that is defined by

$$\begin{aligned} u_1^i &= \alpha_1 n_2^i - p_1^i \\ u_2^i &= \alpha_2 n_1^i - p_2^i \end{aligned} \quad , \tag{175}$$

if platform i attracts n_1^i and n_2^i members of the two groups. The members of these groups incur transport costs t_1 and t_2 when joining either of these platforms, i.e. when going to one of the nightclubs. Hence, the services offered by the platforms are differentiated by the agents' distance from the platform and their transport costs. **Product differentiation** is modelled as a **Hotelling model**.

To find the **fraction of group 1 agents** who join platform A at location 0, we determine the location n_1^i of the indifferent consumer whose utility net of transport costs from joining platform A is the same than that of joining platform B at location 1.

$$\begin{aligned} u_1^A-(n_1^A-0)\cdot t_1 &= u_1^B-(1-n_1^A)\cdot t_1 \\ n_1^A &= 0.5+\frac{u_1^A-u_1^B}{2t_1} \end{aligned} \tag{176}$$

The number of group 1 customers on platform B is determined by $n_1^B=1-n_1^A$. Analogous expressions can be found for agents of type 2. Plugging (175) in (176) and solving for n_1^A and n_1^B yields

$$\begin{aligned} n_1^A &= 0.5+0.5\cdot\frac{\alpha_1(p_2^B-p_2^A)+t_2(p_1^A-p_1^B)}{t_1t_2-\alpha_1\alpha_2} \\ n_2^A &= 0.5+0.5\cdot\frac{\alpha_2(p_1^B-p_1^A)+t_1(p_2^A-p_2^B)}{t_1t_2-\alpha_1\alpha_2} \end{aligned} \tag{177}$$

As the **market share of platform *A*** in customers of group 1 and 2. Analogous expressions can be found for platform *B*.

For positive externalities ($\alpha_1,\alpha_2>0$) the market share of type 1 agents on platform *A* falls in the price charged from type 2 agents. A high price p_2^A implies that there are few type 2 agents on platform *A* which is undesirable form the viewpoint of the type 1 agents. n_1^A rises in the price that platform *B* charges from agents of type 2. The positive impact of the price difference p_1^A-p_1^B on n_1^A is the stronger the higher the transport costs of group 2 (t_2) are. Note that condition (178) must be satisfied for the market shares of both firms to be positive.

$$4t_1t_2>(\alpha_1+\alpha_2)^2 \tag{178}$$

This condition is negative when the network effects (α_1,α_2) are large compared to brand preferences (t_1,t_2). If this were the case, there can be equilibria where only one platform serves the entire market.

If platforms can differentiate from each other they may coexist. For example, magazines differentiate in many dimensions and, hence, coexist. Moreover, if the goods of the seller group are sufficiently differentiated (for example movies), they are more willing to coordinate on the same standard (for example Blu Ray) to distribute this product (see section E.2).

Given that every platform incurs per agent costs of f_1 (f_2) for serving group 1 (2), it can be shown that the first-order conditions for equilibrium prices are

$$\begin{aligned} p_1 &= f_1+t_1-\frac{\alpha_2}{t_2}(\alpha_1+p_2-f_2) \\ p_2 &= f_2+t_2-\frac{\alpha_1}{t_1}(\alpha_2+p_1-f_1) \end{aligned} \tag{179}$$

Expressions (179) can be interpreted in the following manner. “First, note that in a Hotelling model without cross-group externalities, the equilibrium price for group 1 would be $p_1=f_1+t_1$. In this two-sided setting, the price is adjusted downward by the factor $(\alpha_2/t_2)(\alpha_1+p_2-f_2)$. This adjustment factor can be decomposed into two parts. The term $(\alpha_1+p_2-f_2)$ represents the external benefit to a platform when an additional group-2 agent is attracted. To see this, note first that the platform makes profit (p_2-f_2) from an extra group-2 agent. Second, α_1 measures the extra revenue the platform can extract from its group-1 agents (without losing market share) when it has an extra group-2 agent. [...] Finally, [...] a platform attracts α_2/t_2 extra group-2 agents when it has an extra group-1 agent. In

sum, the adjustment factor $(\alpha_2/t_2)(\alpha_1+p_2-f_2)$ measures the external benefit to the platform from attracting an extra group-1 agent; in other words, it measures the opportunity cost of raising the group-1 price by enough to cause one group-1 agent to leave. [...] Thus, a platform will target one group more aggressively than the other if that group is (i) on the more competitive side of the market and/or (ii) causes larger benefits to the other group than vice versa" (Armstrong 2006: 674-675).

When the nightclubs can **price discriminate**, the equilibrium prices p_1 and p_2 can be determined from (179) as follows.

$$\begin{aligned} p_1 &= f_1+t_1-\alpha_2 \\ p_2 &= f_2+t_2-\alpha_1 \end{aligned} \tag{180}$$

When the nightclubs must charge a uniform price $p=p_1=p_2$ because, for example, sex discrimination laws prevent differential pricing by nightclubs, the platforms set an optimal price of

$$p = f + 2\frac{t_1 t_2 - \alpha_1 \alpha_2}{t_1+t_2+\alpha_1+\alpha_2} \quad . \tag{181}$$

(This specification assumes $f=f_1=f_2$ because with very different costs price discrimination is more likely.) "One can show that this uniform price [(181)] lies between the discriminatory prices in [(180)]. Therefore, the consequence of a ban on price discrimination is that one group is made better off (the group that has the higher price with discriminatory pricing) while the other group is made worse off" (Armstrong 2006: 676).

In some situations the platforms can set **two-part tariffs**. For example, the publisher of a newspaper might charge its advertising customers a fixed fee for every advertisement plus a variable fee for every newspaper sold. This greater freedom in setting prices gives rise to multiple equilibria of the game which generate different profits for the firms.

Competitive Bottlenecks

In this setting, group 1 single-homes while group 2 multi-homes. Armstrong (2006: 678) provides the following **examples**: "most people might read a single newspaper (perhaps due to time constraints), but advertisers might place ads in all relevant newspapers; consumers might choose to visit a single shopping mall (perhaps because of transport costs), but the same retailer might choose to open a branch in several malls; consumers might visit a single supermarket, but suppliers might wish to place their products on the shelves of several supermarkets; or a travel agent might use just one computerized airline reservation system, while airlines are forced to deal with all such platforms in order to gain access to each travel agent's customers."

As group 2 agents multi-home and do not make an "either-or" decision between platforms, a group-2 agent's decision to join one platform does not depend on whether she chooses to join the rival platform. Hence, there is no competition between platforms to attract group-2 customers. The **utility of a single-homing group-1 agent** can be affected by the participation of multi-homing group-2 agents in various ways. First, the presence of a great variety of shops (group 2) in a mall may be expected to raise the utility of a shopper (group 1). Second, a great number of advertisements (group 2) in a newspaper can be considered a nuisance by the reader (group 1). Third, a customer, who enters a mall in order to buy a certain product in a certain shop, might not be affected by the presence of other shops at all.

Armstrong (2006: 679) assumes that the platform chooses the set of prices that **maximizes the joint interest** of the platform (profits) and its single-homing group-1 agents (utility). The interests of group 2 are ignored. "In general, this implies that there is a market failure, and there is a suboptimal number of group-2 agents on each platform for a given distribution of group-1 agents. [...] The excessive prices faced by the multi-homing side do not necessarily result in excess profits for platforms, since platforms might be forced by competitive pressure to transfer their monopoly revenues to the single-homing agents. Rather, the market failure is a suboptimal balance of prices to the two sides of the market" (Armstrong 2006: 679-680).

The outcomes in such markets also depend on the way in which charges are levied from group-2 agents. For example, a newspaper may charge advertisers (i) a **variable fee** on a per-reader basis or (ii) a **lump-sum** fee for placing an ad. In Armstrong's (2006) model with a variable fee, the number of ads does not change if a platform attracts more readers. Moreover, readers are tipped by a low price, which sometimes is even lower than costs. The price is even lower when competition for readers is intense or the advertising revenue is large. In case of a lump-sum fee, having more readers attracts more advertisers. When readers like seeing ads, the price charged from them is lower than in the case of a variable advertising fee. This is because a high number of readers attracts advertisers and raises advertising revenues. When readers dislike ads, the price charged from them is higher than in the case of a variable advertising fee. When readers do not care about ads at all, there is no difference between the regimes of lump-sum and per-reader charges for advertising.

J.3 Two-Sided Markets and Competition Policy

The above discussion shows that the competitive analysis of two-sided markets can be complex because three types of agents and their optimization rationales must be considered. This difficulty is particularly severe in the case where some group of agents decides to single-home while the other multi-homes (scenario of competitive bottlenecks). Further complexities arise when one allows the

platforms to charge different prices from agents in the same group or to apply different strategies to them. For example, a "shopping mall might charge a higher rent to a retailer with the promise that it will not let a competing retailer into the same mall" (Armstrong 2006: 686). If a mall prefers to use either such a strategy or one where it allows competing retailers in the mall also depends on its ability to charge shoppers. If it can make revenue by charging shoppers an entry fee (or a parking fee) it is more prone to allow competition among retailers. This lowers retail prices and attracts shoppers. However, if the mall cannot generate revenues with shoppers it might by more likely to charge retailers for not allowing their competitors in the mall.

These aspects have implications for the analysis of such markets in **competition policy**. Rysman (2009: 128) names the **payment card industry** as one example where this is relevant. When a consumer uses her credit card to buy a good she may pay a fee to her credit card company, for example, Visa or Mastercard. The seller of the good pays a fee in order to receive money by means of electronic payment. Moreover, the seller's bank pays an **interchange fee** to the buyer's bank in order to receive the money. This interchange fee is reflected in the fee that the merchant has to pay. Visa and Mastercard used to be organized as associations of banks that provide a clearinghouse for transactions using their cards. These banks would have had the opportunity to collude on the interchange fee in order to raise the fee that has to be paid by the seller. "The interchange fee presumably explains why banks can offer generous rewards program to consumers for using their credit cards" (Rysman 2009: 128).

The existence of two-sided markets and their pricing implications also affect **merger simulations** (see section G). For example, if a merger creates cost reductions on one side of the market, this has implications for pricing on the other side. This also poses problems when calculating counterfactual prices as may be required in **abuse of dominance** cases or when determining the damage caused by a **cartel**. Rysman (2009: 140) predicts "that two-sided analysis will grow in importance in areas where pricing analysis is important".

Lessons Learned

After reading this section you should be able to answer the following questions.

1. What is a two-sided market. What is its defining characteristic?
2. How can two-sided markets be distinguished from one-sided markets?
3. What is meant by the terms single-homing and multi-homing?
4. The agents of group A exert larger externalities on the agants of group B than the other way round? Will the platform charge higher prices from group A or group B?
5. Name some examples where two-sided markets play a role in competition policy.

References

Armstrong, M. (2006). “Competition in two-sided markets.” RAND Journal of Economics. Vol. 37 No. 3, pp. 668-691

Rysman, M. (2009). “The Economics of Two-Sided Markets.” Journal of Economic Perspectives. Vol. 23 No. 3, pp. 125-143

K ART. 107 TFEU – STATE AID

European state aid control "is an essential component of competition policy and a necessary **safeguard for effective competition and free trade** [*emphasis added*]. [...] State aid rules, first and utmost, ensure a level playing field for European companies and **avoid** that Member States engage in **wasteful subsidy races** [*emphases added*], which are non-sustainable for individual Member States and detrimental to the EU as a whole. Secondly, the Commission encourages Member States and regions to prioritize action to strengthen the competitiveness of their economy as well as increase social and regional cohesion. [...] State aid control thus contributes to avoiding a wasteful use of public resources, for which, in the end, the taxpayers would have to pay the bill" (EU 2008: p. 5). Apart from DG Agriculture, DG Fisheries and DG Energy and Transport, DG Competition enforces the rules for aid in all other sectors.

K.1 State Aid and the Law

"State aid control is unique to the EU – no other competition regime in the world has anything quite like it" (Niels et al. 2011: p. 382). Therefore, in the following we will provide an overview on this field. In this section, we start with a definition of state aid before describing the authorization and notification procedures and concluding with some key figures on state aid.

Definition and Types of State Aid

The European policy towards state aid is laid down in **Art. 107 TFEU**. Art. 107 (1) TFEU establishes the general provision that state aid is incompatible with the internal market.

1. *Save as otherwise provided in the Treaties, any aid granted by a Member State or through State resources in any form whatsoever which distorts or threatens to distort competition by favouring certain undertakings or the production of certain goods shall, in so far as it affects trade between Member States, be incompatible with the internal market.*

For a measure to be characterized as state aid it must in particular satisfy the four below **criteria** (EU 2008: p. 6). State aid that satisfies these criteria will be considered incompatible with the common market. The basic concern (in a perfectly working market) is that state aid is used to distort competition and grant some firms an advantage at the cost of other firms.

(a) State aid rules cover only measures involving a **transfer of state resources** but must not be granted by the state itself. It may also be granted by a private or public intermediate body appointed by the State. Financial transfers that constitute aid can take many forms: not just grants or interest rate rebates, but also loan guarantees, accelerated depreciation allowances,

capital injections, tax exemptions etc. Therefore, the concept of state aid is broader than that of a subsidy.

The criterion of state aid being a transfer of state resources also highlights the difference between state aid control and the ban of cartels, merger control, or abuses of dominance. While the latter relate to the behaviour of companies, state aid control is designed to prevent European Member States "from using government funds to circumvent competition rules by gaining an unfair advantage with the support of the domestic industry" (Haucap and Schwalbe 2011: 14).

"Another key difference between aid and other areas of competition policy is that aid equals state transfers that flow from the taxpayers to the companies. In an economic analysis regarding the effects of aid, it is just as important to always consider the effects of aid payments on the financing side (that is, the so-called "shadow costs of taxation") as the effects on the expenditure side. Here, it is crucial to examine which allocative distortions are caused by the tax financing of aid, or how this tax revenue could have been utilized differently" (Haucap and Schwalbe 2011: 15).

(b) The aid should constitute an **economic advantage** that the undertaking would not have received in the normal course of business. The qualification of a measure as state aid follows the **market economy investor principle** (MEIP). A measure may only be qualified as state aid if the funding from the state occurs on more favorable terms than those of a private investor. There is no advantage to the firm if it could have obtained, for example, the funds on the same terms from the private sector. The question is whether the public investor acted equivalently to a market economy investor.

(c) State aid must be **selective** and, thus, affects the balance between certain firms and their competitors. Selectivity is what differentiates state aid from general measures that apply to all firms in a Member State equally.

(d) Aid must have a potential **effect on competition and trade** between Member States. The Commission has taken the view that small amounts of aid (*de minimis* aid) do not have a potential effect on competition and trade between Member States. It therefore considers that such aid falls outside the scope of Article 107(1) of the Treaty.

"The underlying economics of determining competitive distortion are similar to those applied to Article 101 [TFEU ...]. In Article 101 cases, as with state aid, balancing positive and negative effects can be required. The following factors are relevant to the assessment of distortion under the state aid rules:

- Is there strong selectivity (i.e., the aid is targeted at individual companies or specific industries)?
- What is the market position of the company?
- What price does the aided company charge?
- Is there much cross-border trade in the markets affected?
- Is the particular form of aid prone to lead to distortions" (Niels et al. 2008: p. 403)? This question relates to whether the aid is given as a variable subsidy, which affects a firm's variable costs of production, or as a lump-sum subsidy, which affects the firm's fixed costs. Variable subsidies are more likely to distort the firm's pricing behavior than fixed subsidies.
- What was the procedure for selecting beneficiaries? Did all firms have the same chance to apply for state aid or is there strong selectivity?

As markets do not always work perfectly, state aid – or more general: government intervention – can be a tool to correct **market failures** such as externalities (e.g. provide incentives to innovate or provide disincentives to pollute the environment by means of production) and to achieve goals of common interest. The role of economics lies in explaining how and when markets fail and is critical to a sensible policy on state aid (Niels et al. 2011: p. 381; see subsection K.2). Moreover, a state may buy **services of general economic interest** (SGEI), i.e. services which otherwise would not be provided by the market (see subsection K.3 below). As an example one might name the provision of public service broadcasting, i.e. TV programming with a 'higher purpose'.

Art. 107 (2) and (3) TFEU establish under what conditions some types of state aid can be declared compatible with the internal market.

2. *The following shall be compatible with the internal market:*

   (a) *aid having a social character, granted to individual consumers, provided that such aid is granted without discrimination related to the origin of the products concerned;*

   (b) *aid to make good the damage caused by natural disasters or exceptional occurrences;*

   (c) *aid granted to the economy of certain areas of the Federal Republic of Germany affected by the division of Germany, in so far as such aid is required in order to compensate for the economic disadvantages caused by that division. Five years after the entry into force of the Treaty of Lisbon, the Council, acting on a proposal from the Commission, may adopt a decision repealing this point.*

3. *The following may be considered to be compatible with the internal market:*

   (a) *aid to promote the economic development of areas where the standard of living is abnormally low or where there is serious underemployment, and of the regions referred to in Article 349, in view of their structural, economic and social situation;*

   (b) *aid to promote the execution of an important project of common European interest or to remedy a serious disturbance in the economy of a Member State;*

   (c) *aid to facilitate the development of certain economic activities or of certain economic areas, where such aid does not adversely affect trading conditions to an extent contrary to the common interest;*

   (d) *aid to promote culture and heritage conservation where such aid does not affect trading conditions and competition in the Union to an extent that is contrary to the common interest;*

   (e) *such other categories of aid as may be specified by decision of the Council on a proposal from the Commission.*

"For the majority of State aid cases, the most relevant **exemption** [*emphasis added*] clauses are those of Article" 107(3)(a) and 107(3)(c) of the Treaty (EU 2008: p. 7) which set out the rules for three main categories of aid (a-c).

(a) **Regional aid**: Art. 107(3)(a) applies to regions where the per-capita GDP is lower than 75% of the EU-25 average. This legal rule contributes to leveling out differences between states across the EU. In addition, Art. 107(3)(c) gives Member States the possibility to assist regions which are disadvantaged compared to the national average" (EU 2008: p. 8). Such regional aid with non-economic objectives (i.e., a regional balance) and ultimately with fairness and distributive political objectives.

(b) **"Other cross-industry (or: horizontal) rules**: "Cross-industry or "horizontal" rules set out the Commission's position on particular categories of aid which are aimed at tackling fundamental economic problems which may arise in any industry and region. To date, the Commission has adopted "frameworks", "guidelines" or "block exemption regulations" setting out the criteria that are to be applied to the following categories of aid" (EU 2008: p. 8):

   - Aid for climate change and for other environmental protection;
   - Aid for research and development and innovation;
   - Aid for the rescue and restructuring of firms in difficulty;

- Aid for small and medium-sized enterprises;
- Aid to employment;
- Training aid;
- Aid for risk capital; and
- Aid for services of general economic interest.

(c) **Sectoral rules**: The Commission has also adopted industry-specific rules defining its approach to State aid in particular industries. These industries are considered sensitive "due to many years of economic problems and are therefore excluded from the ban on state aid" (Haucap and Schwalbe 2011: 4). They include **general sectors** such as electricity, postal services, and shipbuilding. State aid control in these general sectors is carried out by the Directorate-General for Competition. The Directorate-General for Maritime Affairs and Fisheries and the Directorate-General for Agriculture and Rural development are concerned with state aid rules in the areas of **fishery and agriculture**. The Directorate-General for Energy and Transport is concerned with specific state aid rules for the **transport** and the coal sector.

(d) **Specific aid instruments**: "For the use of specific aid instruments such as guarantees, fiscal aid, capital injections, or for the calculation of the aid content of a measure, guidance has been provided through the publication of a number of notices" (EU 2008: p. 10).

It may be noted that the exempted types of state aid are not necessarily targeted at correcting market failures, i.e. promoting efficiency. The exceptions in Art. 107 (3) TFEU include normative reasons for granting state aid, i.e. they are motivated largely by distribution policy (Haucap and Schwalbe 2011: 5, 15).

Notification and Authorization Procedures

"Community supervision of State aid is based on a system of ex ante authorisation. Under this system, Member States are required to inform ('**ex ante notification**') the Commission of any plan to grant or alter State aid and they are not allowed to put such aid into effect before it has been authorised by the Commission ('**standstill-principle**'). [...] Any aid, which is granted in absence of Commission approval, is automatically classified as '**unlawful aid**'" (EU 2008: p. 13).

The **ex ante notification** requires Member State to submit a summary description of the aid measure within 20 days following the implementation of the measure. No notification is necessary if an aid measure satisfies the conditions laid down in the ***de minimis* regulation** or in the **General Block Exemption Regulation**.

"The Commission has two months within which to [do a **standard assessment** of] the proposed aid. The two-month period runs from the date that the Commission has received all the information it needs to assess the case and the notification can be considered as complete. This examination will normally be concluded either by a "decision not to raise objections" or by a "decision to initiate Article [108](2) proceedings"[.] If the Commission decides not to raise any objection, the aid measure concerned can be implemented" (EU 2008: p. 14).

Upon its decision to initiate Article 108(2) proceedings the Commission performs a **detailed assessment** of the aid measure and adopts a final decision within 18 months. This decision "may be either positive (aid can be implemented), negative (aid can not be implemented) or positive, but subject to stated conditions (aid can be implemented if certain conditions are met)" (EU 2008: p. 14).

Such **remedies** should improve the design of an aid measure in order to reduce its adverse impact on competition and trade. The EU considers five main types of remedies: First, capacity reductions that shall limit the aided firm's market share and may be implemented, second, in the form of divestments of assets. Third, with regard to the firms the Commission also recognizes behavioral commitments. Fourth, with regard to states the Commission recognizes measures of market liberalization which reduces domestic firms' market power by lowering entry barriers. A fifth possibility is the open licensing of intellectual property rights or standards which improves competition.

Key Figures

In 2009, **total state aid** in the EU amounted to 427bn EUR or 3.6% of the GDP. Excluding measures related to the economic crisis, state aid amounts to 73bn EUR or 0.6% of the GDP.[92] In 2009, state aid is distributed across sectors as follows:

	crisis measures excl.	crisis measures incl.
Manufacturing Sectors	71%	13%
Financial Services	2%	82%
Other Services	1%	0%
Agriculture	16%	3%
Fisheries	0%	0%
Coal	4%	1%
Other Non Manufacturing Sectors	2%	0%
Transport (Excl. Railways)	4%	1%

If state aid on agriculture, fisheries, and transport is excluded the above state aid measure drops from 73bn EUR to 58bn EUR or 0.5% of the GDP.

K.2 The Economics of State Aid

Incentives for Granting State Aid

On the one hand, state aid may be used as a tool for correcting market failure. On the other hand, the granting of state aid may be politically motivated.

Market failure describes situations where the market forces alone lead to (productively, allocatively, or dynamically) inefficient outcomes. Some classic reasons for market failure are: (1) external effects, (2) public goods, (3) imperfect competition, (4) information problems, and (5) adjustment or coordination deficiencies.

"Positive and negative **external effects** [*emphasis added*] describe a situation in which the activity (e.g., production or consumption) of an economic agent affects the utility [...], profit or production possibilities of other economic agents without these effects being considered in the market-based pricing system. External effects are thus a matter of damaging or favouring otherwise uninvolved third parties. External effects are a direct result of ill-defined or definable and

92 All data are provided by DG-Comp at http://ec.europa.eu/competition/state_aid/studies_reports/studies_reports.html.

enforceable rights of disposal so that there is no compensation for the damage or favouritism. The missing assignment of rights of disposal leads both in the case of negative and positive external effects to a price mechanism that cannot ensure efficient market results by itself. [...] State intervention must, in the case of serious external effects, be aimed at eliminating the divergence between microeconomic and effective costs and revenues for the whole society with measures to internalize the external effects" (Haucap and Schwalbe: 7).

A **public good** is defined by the characteristics of *non-rivalry in consumption* and a *lack of applicability of the exclusion principle*. There is non-rivalry in consumption if additional demand for the same good does not cause further supply costs, i.e. the marginal cost of an additional user is zero. If potential consumers can be excluded from the consumption of the good a private supplier of the good must charge a price above marginal costs to cover total costs of production. This market outcome is allocatively inefficient because consumers are excluded from buying the good whose willingness to pay is above the marginal costs of producing the good (but below the price charged by the firm). Additionally, in case of a pure public good a consumer cannot be excluded from consuming public goods such as national defense or police services. The latter examples are indeed exempted from state aid control because they are legally defined as the exercise of public power which is considered a non-economic activity (EU 2012.a: para. 16). A "pure public good is not supplied – or at least not in an efficient scale – without government intervention. The supply of a public good does not necessarily need to occur by the government itself. It can also be supplied by private enterprises if they can cover their costs through public funding and are publicly assigned the supply. Such funding could come from state aid [...]. However, if a proper temporary tender regarding the private supply of a public good takes place, and the company with the lowest costs for the supply is awarded the job, public grants by the European Courts do not qualify as state aid in this context" (Schwalbe 2011: 8; also see subsection K.3 below).

Market failure may also be caused by the existence of supply-side economies of scale (see section C.1) and demand-side economies of scale such as positive network effects (see sections E.2 and J). Both types of effects may lead to **imperfect competition** in the form of concentrated industry structures where firms have the potential to exercise market power. "In general, natural monopolies or economies of scale do not require the need for state aid. In many cases, they even require government price controls to protect consumers against exploitation and to protect potential competitors against interferences in upstream and downstream markets. These cases occur when the monopolistic supplier is protected by high barriers to entry against potential competitors and he can therefore raise his price permanently above the competitive level, i.e., even in later phases of the market. In the sector of broadband Internet coverage, however, a state subsidy for second-service providers may be allowed in the market in order to prevent a consolidation of power in local

markets due to economies of scale" (Haucap and Schwalbe 2011: 9).

Information is a good that is often characterized by non-rivalry in consumption. In section H.4 we have shown that it may be welfare enhancing to allow for, e.g., vertical restraints like exclusive supply to ensure that information goods like (after) sales services are provided to a greater extent. A similar reasoning applies to the provision of state aid, that can also be justified by the asymmetric distribution of information. "Asymmetric information exists when one of the relevant actors in the market may be better informed or can more easily (cheaply) inform himself than the other side. The asymmetric distribution of information can create problems of moral hazard and adverse selection" (Haucap and Schwalbe 2011: 9). This may be relevant with respect to the raising of capital which may be more difficult for small and medium-sized enterprises. This is because banks have greater difficulties to assess the risk of such enterprises. "In order to compensate for these competitive disadvantages, the public sector often grants concessional credits to small and medium-sized businesses. Due to the selective nature of this preferential treatment, these credits have state aid character" (Haucap and Schwalbe 2011: 10).

"**Adjustment defects** [*emphasis added*] are situations in which market equilibrium does not exist or a new equilibrium or one at the desired speed does not come into effect due to unfavourable supply and demand situations - especially due to a lack of flexibility on the part of the market players. [...] In the EU and its Member States, state aid is often used as an instrument of sectoral structural policy. Its objective is to either accelerate the structural change or make socially acceptable the problems caused by the structural change from the agricultural (primary) and goods-producing (secondary) sector towards the service sector (tertiary sector). Sectoral structural policy is needed because of the previously mentioned lacks of flexibility. Under certain conditions, adjustment aid (or restructuring aid under European law) can at best be economically justified as an economic policy instrument of sectoral structural policy. Adjustment aid is granted to companies with the aim to simplify the process of adjustment to the prevailing economic conditions. [...] In principle, adjustment aid is intended to help people help themselves. It should only be paid until the necessary adjustment to modified structural framework conditions has completely been implemented" (Haucap and Schwalbe 2011: 10).

"No market failure exists, however, when markets deliver efficient, though not politically desired, results. Thus, widespread access to broadband Internet; high-quality, nationwide postal services on weekdays; an extensive public transportation system (ÖPNV) with busses or trains; or a politically correct media offered on the radio, television and Internet might be politically desirable but often economically inefficient. The market's ineffectiveness to produce the desired political results is in economic theory not referred to as market failure" (Haucap and Schwalbe 2011: 5).

Such **politically motivated reasons** may be classified as (1) Regional, distributive, employment and industrial political purposes or as (2) merit goods and services of general interest.

Conservation subsidies (rescue aid) are granted to preserve economic, cultural or rural structures. “They are granted, for example, in the form of a compensatory allowance in the agricultural and mining sector. Conservation aid is intended to maintain the income of workers in the affected sector by structural change (e.g., coal mining) at a certain socially desired level (distributive political objectives) and to avoid excessive unemployment in the affected regions (employment political objectives)” (Haucap and Schwalbe 2011: 11). **Adjustment and conservation aid** must be seen critical because such aid measures lack accuracy and hinder the necessary structural adjustment process. “Through targeted support of workers in old industries, such as in the form of training and retraining measures, the objectives were achieved more efficiently and sustainably without the negative side effects” (Haucap and Schwalbe 2011: 11).

“The focus of **regional policy** [*emphasis added*] is on the distribution of the production potential and the infrastructural development of the spaces within an economy. The objective of regional policy measures is to create equal living conditions in a region. Against this background, establishment aid was granted in structurally weak areas, which are characterized by high unemployment, in order to attract companies from promising sectors and to increase labour demand. In addition, policymakers hope for further positive effects such as agglomeration advantages following the establishment of companies. Aid motivated by regional politics, however, contains a substantial forecast risk and may therefore miss the intended effect.

In recent years, targeted support of large enterprises (for example, through selective tax breaks) was evident on a national level. The policymakers' intention was, among other considerations, to strengthen the international competitiveness of these national champions above all. According to the theory of the so-called strategic foreign trade policy, an active industrial policy can, in the presence of significant economies of scale and scope, cause the respective domestic companies to generate medium- to long-term profits abroad, which benefit one's own economy.

Most economists are very critical of such support for “national champions” for several reasons. They doubt that the subsidization of one's industry leads to advantages for the aid-granting state, as the benefits generated by the (possible) profits abroad with the help of support are usually smaller than the costs for the concentration of market power on the domestic market. Even the supporters of the strategic foreign trade policy theory assume that situations in which all states grant aid lead to a poor outcome for everyone (prisoner's dilemma) and to inefficient aid races (rat races)” (Haucap and Schwalbe 2011: 11).

Aid is also granted to ensure the supply of a supposedly socially desired amount of goods.

The idea is that such **merit goods** are supplied by the market to an extent that is deemed to low by policymarkers. This is because consumers are assumed to be ignorant about what is good for them, e.g. high-quality TV programs or healthy food, or they know what is good for them but succumb to other temptations. It is also assumed that political decisionmakers have a better knowledge about what is good for those individuals. Hence, state aid is used to supply the merit goods at a reduced price in order to stimulate consumption of these goods. This leads to potential problems: First, the policymaker must identify merit goods and determine the degree of interference with the individual preferences of the citizens. Second, the policymaker typically does not know how its measure affects the behavior of the consumers, which requires a trial-and-error method to fine-tune the measure. The administrative term for such merit goods is **services of general interest** with one example being public transport services. "An efficient provision of goods of public general interest can be achieved by competitively compliant tendering processes. These need to be structured in a way in which the grant goes to the company that supplies the respective good in the desired quality and volume at the lowest cost" (Haucap and Schwalbe 2011: 13).[93]

The Balancing Test

The granting of state aid causes economic costs. "Since they are paid from tax revenues, they first of all represent a removal of income that is distributed to privileged branches or companies. In addition, bureaucratic costs and transaction costs on the part of the company (e.g., for aid consultation, application and reporting duties) are consequences of granting aid. In particular, the protection of a stagnant sector by the means of conservation aid takes away further funding from an economy. Moreover, aid causes undesirable side effects such as price distortions, which may lead to additional state-support payments" (Haucap and Schwalbe 2011: 13). Therefore, it is necessary to weigh the pros and cons of state aid in specific cases.

As in the case of merger control, not every state aid case undergoes a specific in-depth assessment (EU 2008: p. 11). The assessment rather proceeds in four steps (also see the notification procedures above).

1. The least distortive cases are not considered to be state aid (***de minimis* aid**). This is the case when aid to an enterprise is below the threshold of 200,000 EUR over a period of three fiscal years subject to some further criteria (EU 2008: p. 58).
2. Cases for which it is possible to design an easily applicable ex ante set of compatibility criteria (**general block exemption regulation**) are exempted from the notification

93 More information on this topic is provided in subsection K.3 .

requirement.[94]

If a measure neither meets the requirements of the *de minimis* regulation nor one of the block exemption regulations, it is subject to notification of Art. 108 (3) TFEU and is examined by the EU Commission.

3. Most cases undergo a **standard assessment** that is based on predefined conditions which aim to ensure the proportionality and necessity of the aid and its distortive effects.
4. The **detailed assessment** is applied for the potentially most distortive cases where the Commission verifies the economic rationale of the aid on a case by case basis.

This assessment procedure shows that in state aid control a combination of per se rules and individual case analyses is implemented. Hence, the first two steps apply a form-based approach while the last two steps implement an effects-based approach.

Economic analysis plays a role in the two central steps in the assessment of state aid, i.e. the **qualification of state aid** under Art. 107(1) (see above) and the **compatibility assessment of aid measures** under Art. 107(2) and (3), which is described in the following. One should note that an economics-based approach towards state aid is not always used at present, as it is in merger control or abuse of dominance (Niels et al. 2011: p. 382).

"The assessment of aid compatibility is essentially a balancing of the positive effects of aid (in terms of contributing to the achievement of a well-defined objective of common interest) and its negative effects (namely the resulting distortion of competition and trade) (the "**balancing test**" [*emphasis added*])" (EU 2008: p. 11). In this context, it has to be noted that "the same type of aid can have different effects in different situations, and different forms of aid can lead to the same result. An analytical approach that focuses solely on the form of aid will therefore, in many cases, provide incorrect results" (Haucap and Schwalbe 2011: 18). Therefore, the balancing test evaluates state aid with regard to its effects (effects-based approach) not its form. In order to be declared compatible, aid must be necessary and proportionate to achieve a particular objective of common interest. The balancing test involves three steps.

1. The welfare standard in state aid is not clearly defined.[95] However, as a first step in the analysis one should determine whether aid is "aimed at a **well-defined objective of a common interest** [*emphasis added*] (such as growth, employment, cohesion, environmental protection, etc). This can include both **efficiency** and **equity** objectives. The efficiency

94 For a more detailed description of the general block exemption regulation see Haucap and Schwalbe (2011: 38).

95 Haucap and Schwalbe (2011: 20-22) discuss several possible welfare standards. Interestingly, they argue that even the total welfare standard (see section D) would be a concept that is too narrow. This is because total welfare in the affected market does not take into account the effect of state aid on the financing side, i.e. the taxpayers.

objective aims at correcting a market failure (e.g. externalities, imperfect information, coordination problems). The equity objectives can include, for example, the employment of disabled workers, or encouraging firms to set up factories in disadvantaged regions. In some cases aid can also be authorised in order to promote the transition to better functioning markets" (EU 2008: p. 11).

2. As a second step, one should determine whether aid measures are well designed to deliver the objective of common interest. The analysis follows three crucial questions (EU 2008: 12).

   (a) Is state aid an **appropriate policy instrument**, i.e. are the advantages of the aid measure established and demonstrated?

   (b) Does the aid measure have an **incentive effect**? The aid should change the behavior of the beneficiary who then engages in an activity that he would not have carried out at all or only to a limited extent.

   (c) Is the aid measure **proportionate** to the problem tackled? This question addresses whether the same change in behaviour could have been achieved with less aid.

   These questions shall be answered by comparing the aided project with a **counterfactual scenario** where no aid was given. Answering the three above questions can be difficult because the quantification of market failure is problematice, i.e. the assessment of welfare loss that is caused by market failure. "Only when the extent of market failure is known can the amount of aid be determined that should be used at most for the elimination of market failure" (Haucap and Schwalbe 2011: 20).

3. Third and finally, to come to a conclusion one must **balance the positive and the negative effects** of an aid measure and the magnitude of these effects (justification level). The negative effects can be divided into four categories:

   (a) "**Productive efficiency** could be influenced on a market if an inefficient company or an inefficient economic sector receives aid and is thus artificially kept alive. Aid will in this case lead to output that is not produced at minimal cost or that is not desired by the consumers. [... Moreover, if] aid subsidizes the variable production costs of a company, then this usually translates into a direct effect on prices and therefore on competitors and consumers" (Haucap and Schwalbe 2011: 23-24).

   (b) "Negative effects of aid on **allocative efficiency** can occur in the form of the creation or strengthening of market power or a market-dominating position" (Haucap and Schwalbe 2011: 23). Such negative effects are more likely to occur when the beneficiary of state

aid holds a high market share. These effects are particularly strong when market shares are quite asymmetrically distributed and/or the industry is fairly concentrated. Negative effects may also concern upstream or downstream markets.

(c) "State aid, such as in the form of establishment aid, can also help to influence the location decisions of companies and thus affect **spatial allocation of economic activities**. This in turn has consequences for interstate commerce, as the flow of goods between Member States is altered" (Haucap and Schwalbe 2011: 23).

(d) State aid may have negative effect on dynamic efficiency.

These "negative effects are primarily distortive effects on competition and trade. They may include prevention of exit and keeping inefficient firms afloat, crowding out of private investment, disrupting dynamic incentives, costs of State aid linked to fiscal spending, etc." (EU 2008: p. 12). An aid measure should be considered compatible with the internal market, i.e. aid should be granted, if its overall balance of positive and negative effects is positive.

To analyze the positive and negative effects of a state aid measure, one must **define the relevant market**. This is important because a certain amount of aid may be assumed to have a stronger impact in a small market than in a large market. Market definition can be performed by a SSNIP-test methodology (see section F). However, in the context of state aid it needs to be considered that effects can spread across markets and economies. Therefore, geographic market definition is of great importance in state aid cases. Moreover, problems like a reverse cellophane fallacy complicate the task of market definition: "Due to the artificially low prices, consumers are not willing to consider alternative products, which they would have thoroughly accepted as attractive substitutes at a higher, competitive analogue price. In this case, there is a risk to define the relevant product market too narrowly, as important substitutes are not included in the examination" (Haucap and Schwalbe 2011: 23).

The balancing test must not be considered a precise numerical task. However, for weighing the effect of the distortion of competition, which results from the aid measure, against the economic benefit of ameliorating a market failure both effects should be quantified in terms of welfare measures. This assessment is similar to the analysis of efficiency gains according to Art. 101(3) TFEU.

K.3 Services of General Economic Interest

Art. 14 TFEU specifies that the Union and the Member States shall take care that services of general economic interest operate on the basis of principles and conditions, particularly economic and financial conditions, which enable them to fulfil their missions. However, Art. 106(2) TFEU

specifies that *undertakings entrusted with the operation of services of general economic interest or having the character of a revenue-producing monopoly shall be subject to the rules contained in the Treaties, in particular to the rules on competition, in so far as the application of such rules does not obstruct the performance, in law or in fact, of the particular tasks assigned to them.* In other words, the European Union and the Member States have to ensure that the providers of SGEI, irrespective of being public or private undertakings, receive just as much state aid to provide SGEI. **State aid for SGEI shall neither be too high, and thus distortive of competition, nor too low.** Art. 108(3) TFEU specifies that any plans to grant or alter state aid (granted for the provision of SGEI) have to be notified to the European Commission. This enables the Commission to decide whether the aid measure in the current form is compatible or incompatible with the internal market. In the latter case, it decides whether the member State shall abolish or alter such aid.

In this section, we present an introduction to state aid control in the European Union concerning services of general economic interest (SGEI) and present the relevant laws. Moreover, we comment on the *Altmark*-criteria that determine under which conditions state aid for SGEI may be declared compatible with the internal market.

Introduction to SGEI

The European Commission (2011.c) provides the below answers to the following frequently asked questions on services of general economic interest (SGEI):

1. What are services of general economic interest (SGEI)?

> "SGEI are economic activities that would not be produced by market forces alone or at least not in the form of an affordable service available indiscriminately to all. SGEI are carried out in the public interest under conditions defined by the State, who imposes a public service obligation on one or more providers. SGEI range from large commercial activities, such as postal services, energy supply, telecommunications or public transport, to social services, such as care for the elderly and disabled" (EC 2011.c).
>
> The Member States are free to define a service to be of general economic interest. The Commission (EC 2012.b: para. 8) notes that there is a "wide margin of discretion in the definition of services that could be clasified as being services of general economic interest. Thus the Commission's task is to ensure that there is no manifest error as regards the definition of services of general economic interest."

2. Why is public service compensation needed for SGEI?

> "The provision of an SGEI under the conditions imposed by the state may not generate a (sufficient) benefit for the provider. Public service compensation might be needed to offset the additional costs stemming from a provider's public service obligations" (EC 2011.c).
>
> "In the absence of specific Union rules, Member States are generally free to determine how their SGEIs should be organised and financed" (EC 2012.a: para. 2). However, "the Commission must ensure that public funding granted for the provision of such services does not unduly distort competition in the Internal Market" (EC 2011.a). In this context, only "advantages granted directly or indirectly through State resources can constitute State aid within the meaning of Article 107 of the Treaty" (EC 2012.a: para. 31).

3. What aim does the Commission pursue when controlling compensation for SGEI?

> "Public funding granted to companies carrying out economic activities may provide an economic advantage to the beneficiaries that their competitors do not have. This may give rise to **distortions of competition** [*emphasis added*] in the internal market. State aid control of SGEI ensures that public service compensation does not exceed what is necessary for the provision of the public service, that companies have incentives to innovate and compete, and that the internal market is not fragmented" (EC 2011.c).
>
> "In order to be caught by Article 107 of the Treaty, public service compensation must affect or threaten to **affect trade between Member States** [*emphasis added*]. Such an effect generally presupposes the existence of a market open to competition. [...] In fact, where economic activity has been opened up to competition, the decision to provide the SGEI by methods other than through a public procurement procedure that ensures the least cost to the community may lead to distortions in the form of preventing entry by competitors or making easier the expansion of the beneficiary in other markets. Distortions may also occur in the input markets. Aid granted to an undertaking operating on a non-liberalised market may affect trade if the recipient undertaking is also active on liberalised markets" (EC 2012.a: para. 37).

The EU Legislation on SGEI

New rules governing the provision of SGEI in the European Union have been implemented in the years 2011 and 2012. In particular, the Commission (EC 2012.a) has released a **Communication** "on the application of the European Union State aid rules to compensation granted for the provision of services of general economic interest" that clarifies key concepts related to state aid for SGEIs.

A **Commission Decision** 2012/21/EU (EC 2012.b) specifies the conditions under which certain types of state aid in the form of public service compensation is compatible with the Treaty for the Functioning of the European Union (TFEU): This Decision enters into force on 31 January 2012 and "sets out the conditions under which State aid in the form of public service compensation granted to certain undertakings entrusted with the operation of services of general economic interest is compatible with the internal market and exempt from the requirement of notification laid down in Article 108(3) of the Treaty."

A complementary objective is pursued by the Communication (EC 2012.c) from the Commission concerning the European Union **framework** for State aid in the form of public service compensation. "The Commission will apply the provisions of this Communication from 31 January 2012" (EC 2012.c: para. 67). "The principles set out in this Communication apply to public service compensation only in so far as it constitutes State aid not covered by Decision 2012/21/EU. Such compensation is subject to the prior notification requirement under Article 108(3) of the Treaty. This Communication spells out the conditions under which such State aid can be found compatible with the internal market pursuant to Article 106(2) of the Treaty" (EC 2012.c: para. 7).

The *de minimis* **Regulation** (EC 2012.d) establishes that a compensation is deemed no aid if "the total amount of aid granted for the provision of services of general economic interest received by the beneficiary undertaking does not exceed EUR 500 000 over any period of three fiscal years" (EC 2012.d: para. 4, Art. 2). This threshold is more lenient than the "general *de minimis* ceiling of EUR 200 000 per beneficiary over a period of three fiscal years" (EC 2012.d: para. 2). This is because state aid for the provision of SGEI to some extent constitutes a compensation for additional costs linked to the provision of SGEI. "Moreover, many activities qualifying as the provision of services of general economic interest have a limited territorial scope" (EC 2012.d: para. 3). "Compensation of this magnitude is deemed unproblematic because it is too low to have any impact on trade and competition" (EC 2012.e). The *de minimis* Regulation "shall enter into force on the third day following that of its publication in the Official Journal of the European Union" (EC 2012.d: Art. 5). The regulation was published on 26 April 2012.

Based on Article 107(1) TFEU, the State aid rules generally only apply where the recipient,

i.e. the provider of the SGEI, is an **undertaking**. The Court of Justice has consistently defined undertakings as entities engaged in an economic activity, regardless of their legal status and the way in which they are financed. The most relevant criterion for defining an undertaking is whether it carries out an **economic activity**, which is defined as any activity consisting in offering goods and services on a market (EC 2012.a: paras. 8, 9, 11).

Article 107 TFEU does not apply, and thus the legislation on SGEI is not applicable, where the state acts by **exercising public power** or where authorities emanating from the state act as public authorities. Examples for such cases are the army or the police, air navigation safety and control, maritime traffic control and safety, anti-pollution surveillance, and the organization, financing and enforcement of prison sentences (EC 2012.a: para. 16).

Special rules apply to hospitals and **undertakings in charge of social services**, such as providers of services meeting "social needs as regards health and long-term care, childcare, access to and reintegration into the labour market, social housing and the care and social inclusion of vulnerable groups" (EC 2012.b: para. 11). State aid granted to such undertakings in charge of social services is exempted from the requirement to notify the Commission about granting state aid to such undertakings. "Previously only hospitals and social housing were exempted" (EC 2011.a).

Whether **social security schemes** are exempted depends on their nature as either schemes based on the principles of solidarity or economic schemes (EC 2012.a: para. 17). Similar criteria apply to **health care systems** whose structure can differ greatly across EU Member States. "Case-law of the Union has established that public **education** [*emphasis added*] organised within the national educational system funded and supervised by the State may be considered as a non-economic activity. [...] These principles can cover public educational services such as vocational training, private and public primary schools and kindergartens, secondary teaching activities in universities and the provision of education in universities" (EC 2012.a: paras. 26-27).

The Altmark-criteria

State aid granted for the provision of SGEI is considered compatible with the internal market if the four cumulative ***Altmark*-criteria** are met (EC 2012.b: para. 4). These criteria shall ensure that the provider of SGEI "do not enjoy a real financial advantage and the measure thus does not have the effect of putting them in a more favourable competitive position than the undertakings competing with them" (EC 2012.a: para. 43).

1. The recipient undertaking must actually have **public service obligations** to discharge, and the obligations must be clearly defined. **Defining a SGEI** is not always undisputable because the "concept of service of general economic interest is an evolving notion that

depends, among other things, on the needs of citizens, technological and market developments and social and political preferences in the Member State concerned" (EC 2012.a: para. 45). The provider of a SGEI "must have been specifically entrusted by the Member State with the operation of a particular service of general economic interest" (EC 2012.b: para. 13). "Generally speaking, the entrustment of a **'particular public service task'** [*emphasis added*] implies the supply of services which, if it were considering its own commercial interest, an undertaking would not assume or would not assume to the same extent or under the same conditions" (EC 2012.a: para. 47).

2. The **parameters on the basis of which the compensation is calculated** must be established in advance in an objective and transparent manner in order to ensure that they do not confer an economic advantage that could favour the recipient undertaking over competing undertakings (EC 2012.a: para. 54). It is necessary to specify the precise content, duration and territory concerned by the public service obligation. This includes a description of the compensation mechanism, the parameters for determining the compensation and avoding and recovering any possible overcompensation (EC 2012.b: para. 14, Art. 4). "Where the amount of **overcompensation** [*emphasis added*] does not exceed 10% of the amount of the average annual compensation, such overcompensation may be carried forward to the next period and deducted from the amount of compensation payable in respect of that period" (EC 2012.b: Art. 6(2)).

3. In order to avoid unjustified distortions of competition, the **compensation** should not exceed what is necessary to **cover the net costs** incurred by the undertaking in operating the service, **including a reasonable profit** (EC 2012.b: para. 15, Art. 5). Any compensation that exceeds this level constitutes incompatible state aid that should be repaid to the Member State.

   The **net cost** should be "calculated as the difference between the cost incurred in operating the service of general economic interest and the revenue earned from the service of general economic interest or, alternatively, as the difference between the net cost of operating with the public service obligation and the net cost or profit operating without the public service obligation" (EC 2012.b: para. 17). More detailed provisions for the calculation of costs and revenues are provided in Art. 5 of the decision (EC 2012.b).

   Reasonable profit "should be determined as a rate of return on capital that takes into account the degree of risk, or absence of risk, incurred. The rate of return on capital should be defined as the internal rate of return that the undertaking obtains on its invested capital over the duration of the period of entrustment" (EC 2012.b: para. 18, Art. 5(5)). The

Member states should also be able to introduce **incentive criteria** relating, in particular, to the quality of service provided and gains in productive efficiency. If the undertaking does not meet the objectives the compensation is to be reduced, whereas if the undertaking exceeds the objectives, the compensation may be increased (EC 2012.b: para. 22, Art. 5(6)).

4. Where the undertaking that is to discharge public service obligations, in a specific case, is not chosen (**entrustment act, selection of provider**) pursuant to a **public procurement procedure** which would allow for the selection of the tenderer capable of providing those services at the least cost to the community. "Where a generally accepted market remuneration exists for a given service, that market remuneration provides the best benchmark for the compensation in the absence of a tender. Where no such market remuneration exists, the amount of compensation must be determined on the basis of an analysis of the costs that a typical undertaking, well run and adequately provided with material means so as to be able to meet the necessary public service requirements, would have incurred in discharging those obligations, taking into account the relevant receipts and a reasonable profit for discharging those obligations" (EC 2012.a: paras. 69-70).

Article 3 of the Commission Decision on SGEI (EC 2012.b) specifies that state aid "in the form of public service compensation that meets the conditions laid down in this Decision shall be compatible with the internal market and shall be exempt from the prior notification obligation provided for in Article 108(3) of the Treaty".

"State aid falling outside the scope of Decision 2012/21/EU may be declared compatible with Article 106(2) of the Treaty if it is necessary for the operation of the service of general economic interest concerned and does not affect the development of trade to such an extent as to be contrary to the interests of the Union" (EC 2012.c: para. 11). The conditions that such state aid must meet to be declared compatible with the internal market are similar to the above ones laid out in Decision 2012/21/EU.

K.4 Regional State Aid

Article 107(3)(a) TFEU specifies that the "following may be considered to be compatible with the internal market: aid to promote the economic development of areas where the standard of living is abnormally low or where there is serious underemployment". "This kind of State aid is known as national regional aid. [Under the Guidelines applicable for the period 2006-2013 **national**] **regional aid consists of** [*emphasis added*]

- [*bullet points added*] aid for investment granted to large companies, or in certain limited circumstances,

- operating aid, which in both cases are targeted on specific regions in order to redress regional disparities. Increased levels of
- investment aid granted to small and medium-sized enterprises located within the disadvantaged regions over and above what is allowed in other areas are also considered as regional aid" (EC 2006.b, para. 1)

Moreover, **Art. 107(3)(c) TFEU** considers as compatible with the common market "aid to facilitate the development of certain economic activities or of certain economic areas, where such aid does not adversely affect trading conditions to an extent contrary to the common interest"

"By addressing the handicaps of the disadvantaged regions, national regional aid promotes the **economic, social and territorial cohesion** [*emphasis added*] of Member States and the Community as a whole. National regional investment aid is designed to assist the development of the most disadvantaged regions by **supporting investment and job creation** [*emphasis added*] in a sustainable context. It promotes the **expansion, rationalisation, modernisation and diversification** [*emphases added*] of the economic activities of undertakings located in the less-favoured regions, in particular by encouraging firms to set up new establishments there" (EC 2006.a, para. 3).

“Regional aid can only play an effective role if it is used sparingly and proportionately and is concentrated on the most disadvantaged regions of the European Union. In particular, the permissible aid ceilings should reflect the relative seriousness of the problems affecting the development of the regions concerned. Furthermore, the advantages of the aid in terms of the development of a less-favoured region must outweigh the resulting distortions of competition” (EC 2013.b, para. 3). “The **primary objective of State aid control** [*emphasis added*] in the field of regional aid is to allow aid for regional development while ensuring a level playing field between Member States, in particular by preventing subsidy races that may occur when they try to attract or retain businesses in disadvantages areas of the EU, and to limit the effects of regional aid on trade and competition to the minimum necessary” (EC 2013.b, para. 4).

At the time of writing this chapter (March 2013) the *Guidelines on Regional State Aid for 2014-2020* (EC 2013.b) are under discussion. One objective of this revision is to improve the consistency among the guidelines and block exemptions for different types of aid within the State Aid Modernisation (SAM)[96] initiative that was set out in May 2012. The new Guidelines shall "facilitate the treatment of ‘good aid’ (well-designed, targeted at identified market failures and objective of common interests, proportionate and least distortive) and prevent the granting of ‘bad aid’ (which distorts competition, frustrates innovation, delays necessary adjustments, fragments the

96 http://ec.europa.eu/competition/state_aid/modernisation/index_en.html

internal market)" (EC 2013.a: 1). The new Guidelines apply to the following types of aid:

1. “Regional investment aid schemes targeted at specific sectors of economic activity;
2. Individual aid (including ad hoc aid) above the notification threshold: Between €15 million and €37.5 million depending on the region;
3. Investment aid potentially linked to a closure of a similar or same activity in the EEA;
4. Certain regional operating aid schemes, namely: (i) aid to reduce certain specific difficulties faced by SMEs in [areas falling under Art. 107(3)(a) TFEU], (ii) aid to compensate for certain additional costs (other than transport costs) in the outermost regions, (iii) aid to prevent or reduce depopulation in areas with a very low population density” (EU 2013.a: 2).

Demarcation of Regions and Regional Aid Maps

The *Guidelines on National Regional Aid for 2007-2013* (EC 2006.b, para. 12) specify that "the total population coverage of assisted regions in the Community must be substantially less than that of unassisted regions. "Accordingly, the Commission "has decided to fix the limit for the overall population coverage to 42 % of the population of the current [in 2006] Community of 25 Member States" (EC 2006.b, para. 13). The same provision is made in the Draft Guidelines for the period 2014-2020 (EC 2013.b, paras. 132-134).

The criteria of *serious underemployment* and a *standard of living that is abnormally low*, which are laid down Art. 107(3)(a) TFEU, are considered to be satisfied when a region has a per capita gross domestic product (GDP) of less than 75% of the EU-25 average (EC 2006.b, paras. 15-16) respectively 75% of the EU-27 average in the new draft Guidelines (EC 2013.b, para. 137). Art. 107(3)(c) TFEU is wider in scope than Art. 107(3)(a) and "gives the Commission power to authorize aid intended to further the economic development of areas of a Member State which are disadvantaged in relation to the national average" as long as such regional aid forms "part of a well-defined regional policy of the Member State" (EC 2006.b, paras. 21-22). Paragraph 30 of the *Guidelines on National Regional Aid for 2007-2013* lists regions that "may be eligible for selection by the Member States concerned for the award of regional investment aid pursuant to the derogation under Article" 107(3)(c) TFEU.

"The regions of a Member State eligible for regional investment aid under the derogations and the ceilings on the intensity of aid for initial investment approved for each region together form a Member State's regional aid map" (EC 2006.b, para. 96). Aid granted to regions listed in the regional aid maps are exempted from the notification obligation under Art. 108(3) TFEU (EC 2006.b, paras. 96 and 98) as is shown below.

Forms of Regional Aid

Regional aid may be granted in the following forms:

1. Regional investment aid,
2. Operating aid, and
3. Aid for newly created small enterprises

Regional investment aid is aid awarded for initial investments into material or immaterial assets for setting up or extending new establishments, for product innovations, and for process innovations (EC 2006.b, paras. 33-34). "Regional investment aid is calculated either in reference to material and immaterial investment costs resulting from the initial investment project or to (estimated) wage costs for jobs directly created by the investment project" (EC 2006.b, para. 36). "The form of the aid is variable. It may, for example, take the form of grants, low-interest loans or interest rebates, state guarantees, the purchase of a share-holding or an alternative provision of capital on favourable terms, exemptions or reductions in taxes, social security or other compulsory charges, or the supply of land, goods or services at favourable prices" (EC 2006.b, para. 37).

"The **level of the aid** [*emphasis added*] is defined in terms of intensity compared with reference costs. [...] The aid intensity [...] is the discounted value of the aid expressed as a percentage of the discounted value of the eligible costs. [... The aid] is calculated at the moment of notification. [... The] eligible investment costs are discounted to their value at the moment of the granting of the aid. [...] The interest rate to be used for discounting purposes and to calculate the aid amount in a soft loan is the reference rate applicable at the time of grant" (EC 2006.b, para. 41). These provisions do not change under the new draft Guidelines which specify that **gross grant equivalent** (GGE) “means the discounted value of the aid expressed as a percentage of the discounted value of the eligible costs. The GGE is calculated at the time of award of the aid on the basis of the reference rate applicable on that date” (EC 2013.b, para. 16).

Regional investment aid may be compatible with the common market if it satisfies certain **conditions**. The "beneficiary must provide a financial contribution of at least 25 % of the eligible costs, either through its own resources or by external financing, in a form which is free of any public support" (EC 2006.b, para. 39). In paragraphs 42-48 the *Guidelines* explicitly specify maximum aid intensities for certain types of aid. The "admissible aid intensities are from the outset less high in regions qualifying for exemption under Article 87(3)(c) than in those qualifying under Article 87(3)(a)" (EC 2006.b, para. 42). These ceilings may be increased by up to 20% for aid granted to small- and medium-sized enterprises (EC 2006.b, para. 49). In "order to ensure that the investment makes a real and sustained contribution to regional development, aid must be made

conditional [...] on the maintenance of the investment in question in the region concerned for a minimum period of at least five years after its completion" (EC 2006.b, para. 40).

In order to limit distortions of competition, the maximum admissible aid intensities are stricter for **large investment projects** (EC 2006.b, para. 67). These are defined as investment projects with an eligible expenditure above EUR 50 million (EC 2006.b, paras. 60-62). They may be required to be notified individually to the Commission if they exceed certain notification thresholds (millions of EUR of aid granted) that rise with the aid intensity (EC 2006.b, paras. 64). Criteria for the in-depth assessment of regional aid to large investment projects are laid down in a Communication from the Commission (EC 2009) as is shown further below.

Operating aid "means aid to reduce an undertaking's current expenditure that is not related to an initial investment. This includes costs categories such as personnel costs, materials, contracted services, communications, energy, maintenance, rent, administration, etc., but excludes depreciation charges and the costs of financing if these have been included in the eligible costs when granting regional investment aid. Operating aid may be based on actual costs but may be granted in the form of periodic instalments to cover expected costs (periodic lump sum payments)" (EC 2013.b, para. 16). "Regional aid aimed at reducing a firm's current expenses (operating aid) is normally prohibited. Exceptionally, however, such aid may be granted in regions eligible under the derogation in Article 87(3)(a) provided that (i) it is justified in terms of its contribution to regional development and its nature and (ii) its level is proportional to the handicaps it seeks to alleviate" (EC 2006.b, para. 76). Operating aid "is intended to overcome delays and bottlenecks in regional development [... and, thus,] should always be temporary and reduced over time" (EC 2006.b, 79).

Aid for newly created small enterprises may "be granted in addition to regional investment aid, in order to provide incentives to support business start-ups and the early stage development of small enterprises in the assisted areas" (EC 2006.b, para. 84). Here, the idea is to spur the economic development of the assisted regions by promoting entrepreneurial activity in the form of business start-ups. "The Commission will [...] approve aid schemes which provide aid of up to a total of EUR 2 million per enterprise for small enterprises with their economic activity in regions eligible for the derogation in Article [107](3)(a), and up to EUR 1 million per enterprise for small enterprises with their economic activity in regions eligible for the derogation in Article [107] (3)(c). Annual amounts of aid awarded for newly created small enterprises must not exceed 33 % of the above mentioned total amounts of aid per enterprise" (EC 2006.b, para. 86). Such aid bears the risk "risk of distortions of competition, including the risk of crowding-out existing enterprises" (EC 2006.b, para. 85).

In its new draft guidelines, the Commission takes on a more restrictive view on **aid to be**

given to large companies. "Large companies tend to be less affected than small and medium enterprises (SMEs) by regional handicaps for investing or maintaining economic activity in a less developed area. Firstly, large companies can more easily obtain capital and credit on global markets and are less constrained by the more limited offer of financial services in a particular disadvantaged region. Investments by large undertakings can produce economies of scale that reduce location-specific initial costs and, in many respects, are not tied to the region in which the investment takes place. Secondly, large companies making investments usually possess a considerable bargaining power vis-à-vis the authorities, which might lead to aid being awarded without due justification. Finally, large companies are more likely to be significant players on the market concerned and, consequently, the investment for which the aid is awarded may modify the conditions of competition in that market. As a result, the incentive effect and proportionality of such aid might not be ensured, leading to significant distortions of competition" (EC 2013.b, para. 9).

Therefore, the new draft guidelines specify that in "view of potentially high distortive effects, **regional aid cannot be awarded to large undertakings** [*emphasis added*] for their investments in areas eligible for regional aid under Article 107(3)(c) of the TFEU. Furthermore, large undertakings cannot receive operating aid even in regions eligible for aid under Article 107(3)(a) of the TFEU. Large companies may, however, receive operating aid in outermost regions as defined under Article 349 of the TFEU and in sparsely populated areas as defined in paragraph 144(b) of these guidelines" (EC 2013.b, para. 10).

In addition, regional "aid may also not be awarded to firms in difficulties within the meaning of the Community guidelines on State aid for rescuing and restructuring firms in difficulty. Regional aid to the coal, steel and synthetic fibres sectors shall not be considered to be compatible with the internal market. Operating aid may not be awarded to undertakings [in the financial and insurance industry] or to undertakings that perform intra-group activities" (EC 2013.b, paras. 11-13).

The new draft "guidelines shall apply to regional aid in all sectors of economic activity, apart from fisheries and aquaculture, agricultural and transport sector which are subject to special rules laid down by specific legal instruments, which might derogate partially or totally from these guidelines. Broadband network infrastructures, energy-related and environmental infrastructures and RDI infrastructures may benefit from regional investment aid if, in addition to the conditions of [the new draft Guidelines for] regional aid [...], they comply with" sector specific conditions (EC 2013.b, para. 15).

(Exemptions from) Notification

A compulsory state aid notification form is provided by the Commission Regulation No. 1627/2006.

"**Any aid that meets all the requirements of [Regulation 1628/2006] should be exempted from the notification requirement**" [*emphasis added*] (EC 2006.a, para. 7). This applies to aid for initial investments if

(a) "the aid is granted in regions eligible for regional aid, as determined in the approved regional aid map for the Member State concerned for the period 2007 to 2013; and

(b) the aid intensity in present gross grant equivalent does not exceed the regional aid ceiling which is in force at the time the aid is granted for the region in which the investment takes place, as determined in the approved regional aid map for the Member State concerned for the period 2007 to 2013" (EC 2006.a, Art. 4.

"In order to ensure transparency and effective monitoring, this [Regulation 1628/2006] should apply only to regional investment aid schemes which are transparent. [...] Regional aid schemes which are not transparent should always be notified to the Commission. Notifications of non-transparent regional aid schemes will be assessed by the Commission in particular in the light of the criteria set out in the Guidelines on national regional aid for 2007 to 2013" (EC 2006.a, para. 5). "In order to determine whether or not aid is compatible with the common market pursuant to [Regulation 1628/2006], it is necessary to take into consideration the aid intensity and thus the aid amount expressed as a grant equivalent. The calculation of the grant equivalent of aid payable in several instalments requires the use of market interest rates prevailing at the time of grant" (EC 2006.a, para. 4).

The exemption from notification does not apply

- to certain sectors in which special rules apply such as the coal and steel industry, the synthetic fibres and shipbuilding sectors, fisheries and aquaculture, as well as the primary production of agricultural products (EC 2006.a, Art. 1).
- if aid amounts granted to a single undertaking exceed a certain threshold (EC 2006.a, Art. 1).
- to export-related activities (EC 2006.a, para. 18).
- under certain conditions to regional aid awarded in favor of large investment projects (EC 2006.a, Art. 7).

These rules are different under the **new draft Guidelines for 2014-2020**. Under these rules the following types of aid are subject to notification pursuant to Article 108 (3) TFEU.

1. **Investment aid schemes targeted at specific sectors of economic activity**: “As a general rule, investment aid should be awarded under multi-sectoral schemes which should form an integral part of a regional development strategy with clearly defined objectives. Investment

aid schemes targeted at specific sectors of economic activity must be notified" (EC 2013.b, para. 19).

2. **Regional operating aid schemes**: Regional operating aid is normally prohibited. Exceptionally, however, such aid may be awarded under Article 107(3)(a) TFEU if it is justified in terms of its contribution to regional development and that its level is proportional to the difficulties or handicaps its seeks to alleviate. Because of the higher risk of distortion to competition and trade, the following regional operating aid schemes must be notified:
   a) Aid to reduce certain specific difficulties faced by SMEs under Article 107(3)(a) TFEU
   b) Aid to compensate for certain additional costs (other than transport costs) in the outermost regions
   c) Aid to prevent or reduce depopulation in areas with a very low population density (EC 2013.b, paras. 26-27).
3. Two categories of individual investment aid
   a) **Individual investment aid above the notification threshold**: "A Member State must notify aid for an investment project, if the aid awarded from all sources exceeds the maximum allowable aid amount" (EC 2013.b, paras. 21-22) specified in the following table.

Aid intensity	20%	25%	30%	35%	50%
Notification threshold	EUR 15m	EUR 18.75m	EUR 22.5m	EUR 26.25m	EUR 37.5m

   b) **Investment aid (potentially) linked to a closure of the same or a similar activity in the EEA**: "The Commission considers that aid that lead to the relocation of existing capacity within the EEA might have seriously distortive effects on competition and trade between Member States. In particular, aid given in one region may put a premature end to existing economic activity in the other region and may expose the economy of the latter region to significant adjustment costs due to frictions in the labour market and in the redeployment of invested capital.

      Therefore, Member States must notify investment aid for any project by large enterprises or SMEs where the beneficiary has closed down a similar productive activity in the EEA in the two years preceding the award of aid or where the beneficiary plans to close down such an activity in the two years after the investment is completed" (EC

2013.b, paras. 24-25).

Under the new Guidelines for 2014-2020 and the General Block Exemption Regulation the following types of will be **block exempted** and will no longer require notification (EU 2013.a: 2):

- Ad hoc aid below the notification threshold
- Aid for newly created small enterprises
- Certain types of operating aid for outermost regions and sparsely populated areas

Compatibility Assessment and the Balancing Test

The compatibility assessment is undertaken with respect to the standards set in the *Guidelines on National Regional Aid for 2007-2013* (EC 2006.b). These apply "to regional aid granted in every sector of the economy apart from the fisheries sector and the coal industry which are subject to special rules laid down by specific legal instruments" (EC 2006.b, para. 8). Moreover, the guidelines do not apply to the production of some agricultural products and may not fully apply to some other sectors that are subject to specific rules. "As a general rule, regional aid should be granted under a multi-sectoral aid scheme which forms an integral part of a regional development strategy with clearly defined objectives" (EC 2006.b, para. 10).

The compatibility assessment of regional state aid is made with recourse to the **balancing test** as is explained in the new draft Guidelines (EC 2013.b).

1. The **first step of the balancing test** specifies that the aid shall contribute to a well-defined **objective of common interest**. For “regional aid the main objective is to address in particular equity considerations, namely furthering economic cohesion in the EU by helping to reduce the gap between the development levels of the various regions of the EU” (EC 2013.b, para. 28). “This would be the case in particular for measures implemented in accordance with regional development strategies defined in the context of the European Regional Development Fund (ERDF), the European Social Fund, the Cohesion Fund, the European Agricultural Fund for Rural Development or the European Maritime and Fisheries Fund [...]. For aid schemes outside an operational programme financed from the cohesion policy funds, Member States should demonstrate that the measure is consistent and contributes to the development strategy of the area concerned. [...] To ensure that the investment makes a real and sustained contribution to the development of the area concerned, the investment must be maintained in the area concerned for at least five years, or three years for SMEs, after its completion” (EC 2013.b, paras. 34-38).

   To demonstrate the regional contribution of individual investment aid, Member

States may use direct and indirect indicators. These include the direct and indirect jobs created by the investment as well as the commitment to enter into training activities that improve the skills of the workforce. External economies of scale (clustering effect), knowledge spillovers, and co-operation with local higher education institutions may also be considered positively. Moreover, the "duration of the investment and possible future follow-on investments are an indication of a durable engagement of a company in the region" (EC 2013.b, para. 43).

In addition to contributing to a well-defined objective of common interest, aid may be granted if the market fails or it fails to deliver an **equity objective**. "State aid should be targeted towards situations where aid can bring about a material improvement that the market cannot deliver itself. This holds especially in a context of scarce public resources" (EC 2013.b, para. 48). This "element of the compatibility test is fulfilled if the regions concerned are included in approved regional aid maps" (EC 2013.b, para. 28).

2. The **second step of the balancing test** contains the following elements:

   a) One needs to establish that a certain type of aid is an **appropriate policy instrument**, i.e. the advantages of the aid measure are established and demonstrated. Other measures such as infrastructure development, enhancing the quality of education and training, or improvements in the business environment can also be used to address the policy goals. An "aid measure will not be considered compatible if other less distortive [with respect to trade and competition] policy instruments or aid instruments allow reaching the same positive contribution to regional development" (EC 2013.b, para. 28).

   b) It is important to ensure that regional aid produces a **real incentive effect** as compared to a counterfactual situation without the aid. "An incentive effect occurs when the aid changes the behaviour of an undertaking in such a way that is engages in additional activity contributing to the development of an area which it would not carry out without the aid or it would carry out in a restricted or different manner or location. The aid must not subsidise the costs of an activity that an undertaking would anyhow incur and must not compensate for the normal business risk of an economic activity" (EC 2013.b, para. 57). To provide evidence of the incentive effect, the Member States may rely on official board documents, risk assessments, financial reports, internal business plans, expert opinions, as well as demand, cost, or financial forecasts (EC 2013.b, para. 69). "The existence of an incentive effect can be proven in two possible scenarios:

      1. The aid gives an incentive to adopt a positive investment decision because an investment that would otherwise not be sufficiently profitable for the beneficiary can

take place in the area concerned (scenario 1, investment decision) or

2. The aid gives an incentive to opt to locate a planned investment in the relevant area rather than elsewhere because it compensates for the net handicaps and costs linked to a location in the area concerned (scenario 2, location decision)" (EC 2013.b, para. 58).

c) The aid measure must be **proportionate to the problem** tackled. In this context, regional "aid can only play an effective role if it is used sparingly and proportionately and is concentrated on the most disadvantaged regions of the European Union. In particular the permissible aid ceilings should reflect the relative seriousness of the problems affecting the development of the regions concerned" (EC 2006.b, para. 5).

The amount of **aid shall be limited to the minimum** needed to induce additional investment or activity in the area concerned. For aid granted to SMEs this criterion is deemed to be fulfilled if the aid intensity remains under a predefined maximum. "The maximum aid intensities are modulated in function of three criteria: 1) the socio-economic situation of the area concerned, [...] 2) the size of the beneficiary [...], and 3) the size of the investment project" (EC 2013.b, para. 77).

The maximum aid intensities are higher in regions covered by Art. 107(3)(a) TFEU than in regions covered by Art. 107(3)(c) TFEU (EC 2013.b, para. 156). Specific values are listed in paras. 157-163 of the draft Guidelines.

d) Any **undue negative effects shall be avoided**. "The Commission identifies two main potential distortions of competition and trade caused by regional aid. These are product market distortions (which lead mainly to allocative inefficiencies) and location effects (which may lead to both allocative inefficiencies and distributional concerns)" (EC 2013.b, para. 100).

"One potentially harmful effect of State aid is that it prevents the market mechanism to deliver efficient outcomes through the reward of the most efficient producers and the pressure on the least inefficient to improve, restructure or exit the market. A substantial capacity expansion induced by State aid in an underperforming market [...] might in particular unduly distort competition, as the creation or maintenance of overcapacity could lead to a squeeze on profit margins, a reduction of competitors' investments or even their exit from the market. This might lead to a situation where competitors that would otherwise be able to stay on are forced out of the market as a consequence of State aid. It may also prevent firms from entering the market and it may weaken incentives for competitors to innovate. This results in inefficient market structures which are also harmful to consumers in the long run. Further, the

presence of aid may make (potential) beneficiaries complacent or more risk seeking. The long run effect on the overall performance of the sector is likely to be negative.

Aid may also have distortive effects in terms of **increasing or maintaining substantial market power** [*emphasis added*] on the part of the beneficiary. Even where aid does not strengthen substantial market power directly, it may do so indirectly, by discouraging the expansion of existing competitors or inducing their exit or discouraging the entry of new competitors.

Apart from distortions on the product markets, regional aid by nature also **affects the location of economic activity** [*emphasis added*]. Where one area attracts an investment due to the aid, another area loses out on that opportunity. These negative effects in the areas adversely affected by aid may be felt through lost economic activity and lost jobs including those at the level of subcontractors. It may also be felt in a loss of positive externalities (e.g. clustering effect, knowledge spillovers, education and training, etc.)" (EC 2013.b, paras. 101-103).

"In order to identify and assess the potential distortions of competition and trade, Member States should provide evidence permitting the Commission to (i) identify the product markets concerned (i.e. products affected by the change in behaviour of the aid beneficiary) and (ii) identify the competitors and customers/consumers affected.

The Commission will use various criteria to assess these potential distortions, such as market structure of the product concerned, performance of the market (declining or growing market), selection process of the aid beneficiary, entry and exit barriers, product differentiation." (EC 2013.b, paras. 112-113).

"The Commission distinguishes two main sources of potential negative effects on product markets: (i) cases of significant capacity expansion which leads to or deteriorates an existing situation of overcapacity, especially in a declining market and (ii) cases where the aid beneficiary holds substantial market power. [...] Where the market in question is growing, there is normally less reason to be concerned that the aid will negatively affect dynamic incentives or will unduly impede exit or entry. More concern is warranted when markets are in decline" (EC 2013.b, para. 115-118).

The evaluation of possible negative effects "shall be carried out by an expert independent from the state aid granting authority on the basis of a common methodology and shall be made public" (EC 2013.b, para. 130).

e) The aid shall be awarded in a **transparent way**. It "must be ensured that the Member

States, economic operators, the interested public and the Commission have easy access to all relevant acts and pertinent information about the aid awarded thereunder" (EC 2013.b, para. 28).

The draft Guidelines (EC 2013.a: 4) specify that the aid must not be considered with the internal market if any of the criteria from steps 1 or 2 are not satisfied. The third step, i.e. the balancing assessment, shall be performed only when the other criteria are met.

3. In the **third step of the balancing test**, the positive and negative effects shall be balanced, i.e. "the advantages of the aid in terms of the development of a less-favoured region must outweigh the resulting distortions of competition. The weight given to the advantages of the aid is likely to vary according to the derogation applied, so that a greater distortion of competition can be accepted in the case of the most disadvantaged regions covered by Article 87(3)(a) than in those covered by Article 87(3)(c)" (EC 2006.b, para. 5).

Specific rules have been set for the analysis of **regional aid to large investment projects**. With regard to the effects of such aid there is an increased risk that trade will be affected by large investment projects associated with the risk of a distortive effect on competitors in other regions. "State aid for these projects would lead to perverse effects such as inefficient location choices, higher distortion of competition and, since aid is a costly transfer from taxpayers in favour of aid recipients, net welfare losses, i.e. the cost of the aid exceeds the benefits to consumers and producers" (EC 2009, para. 3). To take account of these possible distortions the "maximum aid intensities are graduated between 10 % and 50 % of eligible costs [...]. These graduated aid intensities reflect, in essence, the balancing exercise which the Commission must perform between, on the one hand, the positive effects that regional investment aid can have, in particular in terms of promoting cohesion through attracting investment to disadvantaged areas, and, on the other hand, limiting the potential negative effects" (EC 2009, para. 2).

"Despite the automatic scaling-down [caused by the definition of maximum aid intensities], certain large amounts of regional aid for large investment projects could still have significant effects on trade, and may lead to substantive distortions of competition" (EC 2009, para. 6). Therefore, the Commission performs a balancing test, i.e. in "line with the State Aid Action Plan, the Commission will carry out an overall evaluation of the aid based on a balance of its positive and negative effects in order to determine whether, as a whole, the aid measure can be approved" (EC 2009, para. 8).

1. **Objective of the aid**: Regional aid has an objective of common interest which mainly reflects equity considerations but may also address efficiency issues of market failure (EC 2009, paras. 11-13). By stimulating investment and job creation, regional investment aid addresses equity considerations and contributes to furthering economic cohesion by

reducing the gap between the development levels of the various regions in the EU. However, regional aid may also address issues of market failure such as imperfect information, co-ordination problems, difficulties for the beneficiary to appropriate investments in public goods or externalities from investments. The "positive effects of the aid can be both direct (e.g. direct jobs created) and indirect (e.g. local innovation)" (EC 2009, para. 14).

2. Positive effects and design of the aid measure

a) **Appropriateness of the aid instrument**: A selective aid instrument is considered appropriate when the Member States considered the use of general measures – such as infrastructure development, enhancing the quality of education and training, or improvements in the general business environment – and found established advantages of the selective aid instrument (EC 2009, paras. 17-18).

b) **Incentive effect**: "Analysing the incentive effect of the aid measure is one of the most important elements in the in-depth assessment of regional aid to large investment projects" (EC 2009, para. 19). "The objective of this detailed assessment is to determine whether the aid actually contributes to changing the behaviour of the beneficiary, so that it undertakes (additional) investment in the assisted region concerned" (EC 2009, para. 21). The incentive effect can be shown by means of comparison to a counterfactual in two possible scenarios:

1. "The aid gives an incentive to adopt a positive investment decision because an investment that would otherwise not be profitable for the company at any location can take place in the assisted region.
2. The aid gives an incentive to opt to locate a planned investment in the relevant region rather than elsewhere because it compensates for the net handicaps and costs linked to a location in the assisted region" (EC 2009, para. 22).

As evidence of the incentive effect, the Member States may, for example, rely on risk assessments, financial reports, internal business plans, expert opinions, as well as demand, cost, and financial forecasts.

c) **Proportionality of the aid**: "For the regional aid to be proportional, the amount and intensity of the aid must be limited to the minimum needed for the investment to take place in the assisted region" (EC 2009, para. 29). "In scenario 1, for an investment incentive, the aid will generally be considered proportionate if, because of the aid, the return on investment is in line with the normal rate of return applied by the company in other investment projects" (EC 2009, para. 32). "In scenario 2, for a location incentive,

the aid will generally be considered proportionate if it equals the difference between the net costs for the beneficiary company to invest in the assisted region and the net costs to invest in the alternative region(s)" (EC 2009, para. 33).

3. **Negative effects of the aid**: Thre are three potential negative effects arising from the aid. These are (i) high market shares respectively the creation of market power, (ii) overcapacity respectively the maintenance of inefficient market structures, and (iii) negative effects on trade. "To assess market shares and potential overcapacity in a market in structural decline, the Commission needs to define the relevant product market and geographic market" (EC 2009, para. 37).

   a) **Crowding-out of private investment**:

      1. **Market power**: Investment aid to just one beneficiary in a concentrated market may lead to a distortion of competition as the competitors of the beneficiary may react by reducing their own investment expenditures (crowding out) or even exit the market. This distortive effect is to the detriment of consumers while it is particularly strong when a dominant market player is subsidised. "Therefore, the Commission wants to limit State aid to companies with market power [... and takes] account of the market shares and other related factors before and after the investment" (EC 2009, paras. 42 and 44).

      2. **Creating or maintaining inefficient market structures**: In effective competition, inefficient firms are forced to exit the market which ensures an efficient use of scarce resources. This mechanism might be impeded by state aid. If the aid is selectively given to only some firms other, possibly more efficient firms, might either be driven out of the market or be prevented from enetring the market. These effects are connected to the creation or cementation of overcapacity by the beneficiaries. Therefore, capacity created by state aid funding is considered distortive of competition especially when the market declines (EC 2009, paras. 45-49).

   b) **Negative effects on trade**: The "geographical specificity of regional aid distinguishes it from other forms of horizontal aid. It is a particular characteristic of regional aid that it is intended to influence the choice made by investors about where to locate investment projects. When regional aid is off-setting the additional costs stemming from the regional handicaps and supports additional investment in assisted areas, it is contributing not only to the development of the region, but also to cohesion and ultimately benefits the whole Community" (EC 2009, para. 50). However, aid must not be given to attract investments outside these disadvantaged areas.

4. **Balancing the effects of the aid**: The "Commission will balance the positive effects of the regional investment aid to a large investment project with its negative effects. Careful consideration will be given to the overall effects of the aid on cohesion within the Community" (EC 2009, para. 52). This is particularly important when the aid measure has as its effect that an investment planned to be made in one region is made in another region. The "Commission considers that attracting an investment to a poorer region [...] is more beneficial for cohesion within the Community than if the same investment is located in a more advantaged region" (EC 2009, para. 53). This is clearly no Pareto-improvement such that the negative effects in the richer region should be given special consideration in the balancing assessment.

The new draft Guidelines propose to limit the possibility of **investment aid to large enterprises** under Art. 107(3)(a) TFEU. It is argued that there "is a strong body of evidence suggesting that regional investment aid is more effective and efficient when geared towards [small- and medium-sized enterprises (SMEs) ...]. According to this evidence, compared with SMEs, large enterprises would more often have made the investment concerned even without financial support in assisted areas, rendering such support both ineffective and costly with high distortive effects for the internal market" (EC 2013.a: 4).

"Available empirical evidence suggests that the lack of incentive effect for large companies could be attributed in part to the observation that access to finance is more often a problem for SMEs than it is for large enterprises. From this perspective, financial support to SMEs can be expected to make more of a difference than financial support to large enterprises. Second, large enterprises typically have more leverage (bargaining power) vis-à-vis public authorities, as they are relatively more important to the region than individual SMEs. The efficiency of financial support given to large enterprises, as measured e.g. in terms of cost per job created, may be adversely affected as a result" (EC 2013.a: 5). The Commission relies this point of view on studies that there is a lack of the "incentive effect for large firms: large enterprises are using the money for projects they would have carried out anyhow. [... The regional aid would not be" the determining factor to invest or to locate these investments in assisted areas. [... The] main drivers for locating the investments in an assisted area [would be] the high economies of scale and agglomeration effects, the location costs were lower than in any other location. Therefore, there [would be] little or nothing to compensate for with regional investment aid. [... However, it would be useful to recall] that aid to large enterprises can still be granted [under Art. 107(3)(c) TFEU] for aid geared towards specific objectives (e.g. broadband, RDI, energy and environment, etc.)" (EC 2013.b: 5).

Lessons Learned

After reading this section you should be able to answer the following questions.

1. What are the main objectives of European state aid control?
2. What criteria must be satisfied by a measure for being classified as state aid?
3. Under what conditions can some types of state aid be declared compatible with the internal market?
4. Describe the notification and authorization procedure of state aid.
5. Describe the three steps of the balancing test.

References

Danwitz, T.v. (2009). “Dienste von allgemeinem wirtschaftlichen Interesse in der europäischen Wettbewerbsordnung.” in *Die Daseinsvorsorge im Spannungsfeld von europäischem Wettbewerb und Gemeinwohl*. Krautscheid, A. (ed.). Verlag für Sozialwissenschaften: Wiesbaden

European Commission (2006.a). "COMMISSION REGULATION (EC) No 1628/2006 of 24 October 2006 on the application of Articles 87 and 88 of the Treaty to national regional investment aid." Official Journal of the European Union, L 302/29

European Commission (2006.b). "GUIDELINES ON NATIONAL REGIONAL AID FOR 2007-2013." Official Journal of the European Union, C 54/13

European Commission (2006.c). "COMMISSION REGULATION (EC) No 1627/2006 of 24 October 2006 amending Regulation (EC) No 794/2004 as regards the standard forms for notification of aid." Official Journal of the European Union, L 302/10

European Commission (2008). “Vademecum Community Law on State Aid.” http://ec.europa.eu/competition/state_aid/studies_reports/vademecum_on_rules_09_2008_en.pdf

European Commission (2009). "Communication from the Commission concerning the criteria for an in-depth assessment of regional aid to large investment projects." Official Journal of the European Union, C 223/3

European Commission (2011.a). “State aid: Commission adopts new rules on services of general economic interest (SGEI).” Press release IP/11/1571. http://europa.eu/rapid/pressReleasesAction.do?

reference=IP/11/1571&format=HTML&aged=0&language=EN&guiLanguage=en

European Commission (2011.b). “Commission adopts Quality Framework for Services of General Interest in the EU.” Press Release IP/11/1579. http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/1579&format=HTML&aged=0&language=EN&guiLanguage=en

European Commission (2011.c). “State aid: Commission adopts new package on State aid rules for services of general economic interest (SGEI) – frequently asked questions.” MEMO/11/929. http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/11/929&format=HTML&aged=0&language=EN&guiLanguage=en

European Commission (2012.a). “Communication from the Commission on the application of the European Union State aid rules to compensation granted for the provision of services of general economic interest.” Official Journal of the European Union. C 8/4. 11 January 2012. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2012:008:0004:0014:EN:PDF

European Commission (2012.b). “Commission Decision of 20 December 2011 on the application of Article 106(2) of the Treaty on the Functioning of the European Union to State aid in the form of public service compensation granted to certain undertakings entrusted with the operation of services of general economic interest.” Official Journal of the European Union. L 7/3. 11 January 2012. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:007:0003:0010:EN:PDF

European Commission (2012.c). “Communication from the Commission. European Union framework for State aid in the form of public service compensation.” Official Journal of the European Union. C 8/15. 11 January 2012. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2012:008:0015:0022:EN:PDF

European Commission (2012.d). “Commission Regulation (EU) No 360/2012 of 25 April 2012 on the application of Articles 107 and 108 of the Treaty on the Functioning of the European Union to de minimis aid granted to undertakings providing services of general economic interest.” Official Journal of the European Union. L 114/8. 26 April 2012. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:114:0008:0013:EN:PDF

European Commission (2012.e). “State aid: Commission adopts de minimis Regulation for services of general economic interest (SGEI).” Press Release IP/12/402.

http://europa.eu/rapid/pressReleasesAction.do?reference=IP/12/402&format=HTML&aged=0&language=EN&guiLanguage=en

European Commission (2013.a). "Explanatory note on the paper of the services of DG Competition containing draft regional aid guidelines 2014-2020." http://ec.europa.eu/competition/consultations/2013_regional_aid_guidelines/explanatory_note_en.pdf, accessed on March 01, 2013

European Commission (2013.b). "Paper of the services of DG Competition containing draft guidelines on regional State aid for 2014-2020." http://ec.europa.eu/competition/consultations/2013_regional_aid_guidelines/paper_en.pdf, accessed on February 27, 2013

Haucap, J. and Schwalbe, U. (2011). “Economic principles of state aid control.” DICE discussion paper, No. 17. http://hdl.handle.net/10419/45613

Niels, G., Jenkins, H., Kavanagh, J. (2011). “Economics for Competition Lawyers.” Oxford University Press: Oxford