



UvA-DARE (Digital Academic Repository)

Increasing the requirements to show antitrust harm in modernised effects-based analysis: an assessment of the impact on the efficiency of enforcement of Art 81 EC

Lankhorst, M.

Publication date
2010

[Link to publication](#)

Citation for published version (APA):

Lankhorst, M. (2010). *Increasing the requirements to show antitrust harm in modernised effects-based analysis: an assessment of the impact on the efficiency of enforcement of Art 81 EC*. [Thesis, fully internal, Universiteit van Amsterdam]. Amsterdam Center for Law & Economics.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Economic Analysis of the Effects-based Standard

3.1 Introduction

The Commission's decision practice applying the effects-based standard cannot properly be evaluated without a clear view of the function of this standard within the antitrust enforcement system as a whole. Its primary roles are to signal to firms what contracts they are permitted to sign and to determine what evidence each of the parties in litigation must bring forward in order to achieve the results they seek. This chapter presents an economic analysis of these core features of the legal standard.¹ The analysis has two layers.

First, in Section 3.3, we adopt a policymaker's perspective and consider the optimal calibration of the legal standard. It is natural for students of antitrust to assume that economic concepts determine the shape of substantive rules. What approach should be adopted under the antitrust laws towards a certain type of restrictive practice does not only depend on its potential to cause harm, however. To be sure, the costs that will be incurred in showing such an effect and the costs of errors that may be produced will also affect the optimal calibration of the legal standard. This is a key insight that allows us to structure the examination that is engaged in the following chapters, by focusing on these three variables: priors about

¹ That is not to say that it informs the reader about the content of the rule contained in Art 81 EC, that is, the types of economic behaviour it permits and prohibits, and why. A first impression of these matters was given in Chapter 2, Sections 2.3 and 2.4, and a more comprehensive discussion follows in Chapter 4. The present chapter concentrates on the calibration of the legal standard and the way in which it influences the behaviour of firms and litigants, and the effects this has on the performance of the system as a whole. With a view to the comparison of EU and US antitrust made in Chapter 4 it should be mentioned, also, that the analysis presented in Sections 3.2; 3.3.1; and 3.5 is purported to apply in equal measure to the functioning of the *rule of reason* in US antitrust.

the potential to cause harm, enforcement costs, and error costs in European effects-based analysis.

Next, in Section 3.4, in order to deepen our understanding of the error costs that may be generated in enforcement, we study how the legal standard influences the decisions of firms that consider concluding a restrictive contract with possible antitrust implications.² This leads us to examine the role of legal certainty. In the antitrust literature on *per se* analysis, it has been shown that uncertainty about whether an intended agreement will be caught may result in under-compliance (see e.g. Schinkel and Tuinstra, 2006). By contrast, it is argued here that the characteristics of vertical restraints and joint ventures – the types of agreement subjected to the effects-based standard – are such that uncertainty about the precise location of the legal standard may, on occasion, also result in over-compliance. This, too, is important for the analysis pursued below, since over-compliance must be addressed by different means than under-deterrence.

To prepare for these discussions, Section 3.2 provides an introduction to the general principles that underlie the enforcement system as a whole, the role played by the legal standard within this system, and some basic concepts of decision theory.

3.2 Efficient enforcement, the legal standard, and decision theory

The system as a whole

The operation of the enforcement mechanism is best explained in relation to the objective of Art 81 EC. This provision is meant to prevent firms from adopting restrictive agreements that impose a cost on society. Enforcement action – necessary to ensure that firms actually take this rule in the book into account when deciding on the nature of the agreements they conclude with other market participants – is also costly to society. The optimal system of enforcement, which corresponds best to the logic of Art 81 EC, minimizes the combined costs of non-compliance and enforcement.³

² In this and the following chapters the terminology used in Chapter 2 is adopted. There it was explained that not all agreements that restrict competition cause consumer harm. A restrictive agreement therefore constitutes a potential violation of Art 81 EC.

³ It may be surmised that in a system based on deterrence such as has been instituted by Regulation 1/2003 ([2003] OJ L1/1, see Chapter 2, Section 2.4.1, and Chapter 5, Section 5.3) the costs of non-compliance (which is taken here to include welfare losses due to over-compliance, see Section 3.4.1) are superior to the costs of enforcement by at least an order of magnitude.

In considering the practical implications of this principle it is important to appreciate that enforcement is a process that occurs on two levels.⁴ Enforcement action is directed at penalising firms who have signed and implemented harmful agreements (offenders). From society's point of view, however, it is the effect of these actions on the decision making process of firms who may consider signing such agreements (potential offenders) that is of primary importance. As suggested by Becker (1968), given the costs involved, it is preferable to encourage as many potential offenders as possible to comply at the earlier stage, with a minimum of enforcement effort.

Enforcement systems based on deterrence, such as the one currently in place in Europe,⁵ are designed to achieve this objective by the following means. Agreements that reduce consumer welfare will generally produce extra profits for firms. The aim is to cancel out this rise in expected profits, by adding to the cost side of the equation. To this end the threat is created that enforcers challenge harmful agreements in proceedings which lead to the imposition of sanctions if an infringement is found. Compliance, and the costs of non-compliance, therefore, depend on potential offenders' expectations as to the likelihood of intervention by enforcers, on their estimation of their own chances as defendants, and on the severity of the sanctions they expect would be imposed.

Notice that if costs and benefits in the form of sanctions and profits were the only determinants of behaviour, enforcement could easily be made to produce optimal results. Sanctions would be set at the highest of levels and imposed on all firms found to conclude restrictive agreements. Consequently, no restrictive agreements would be

⁴ This applies to systems based on *ex post* intervention, such as instituted in Europe by Regulation 1/2003 (supra, footnote 3).

⁵ The enforcement system that was in place before the entry into force of Regulation 1/2003 (supra, footnote 3), did not primarily depend on selective *ex post* intervention to induce firms to comply with Art 81 EC. Regulation 17/62 ([1962] OJ 13/204) instituted a system of *ex ante* screening. In view of the importance that will be attributed to the design of this enforcement regime in subsequent chapters, it is worth explaining the circumstances under which it made sense to create a system that depended on (costly) interventions in a very large number of cases. This underscores the importance of legal certainty. As is reported in Gerber (1998), at the outset of European competition policy, the Commission faced a difficult situation. The provisions in the Treaty employ broad language that leaves room for interpretation. The Commission and the Court would need powers to assess a large number of cases if they were to be able to exert sufficient influence over the future shape of competition policy to ensure that it would serve the goal of common market integration. In addition, firms could be expected to be hostile to competition policy. In several member states cartelisation had been part of industrial policy before the war. The system therefore also served to ensure a minimum level of compliance. In sum, at the time policymakers estimated that a high level of intervention was necessary to ensure a sufficient case load to speedily develop a body of European antitrust law that would be respected by firms.

concluded, no litigation would result, and enforcement would consume few, if any, resources.

The fact that in reality transgressions do occur and that costs are incurred in dealing with them implies that we must widen the scope of our analysis. It is essential to recognise that potential offenders may not know whether the agreement they intend to sign will infringe the law; whether, if so, it would be challenged; and, finally, exactly what negative consequences would follow. Potential enforcers and defendants, in turn, will be uncertain as to the exact nature of the evidence that their opposing party can produce, and the probable interpretation of facts or the law by the authority or court that reviews their case. These uncertainties have a decisive bearing on the decisions that each of the actors takes and, therefore, on the efficiency of enforcement.⁶ For example, the lower potential offenders estimate the probability of being challenged, the more attractive it will be to infringe Art 81 EC.

The legal standard

The legal standard is the test that is applied to see whether an agreement is prohibited under Art 81 EC and it is an integral and crucial part of the enforcement system. It regulates the investigation of restraints in the course of legal proceedings by signalling to parties in litigation what arguments and associated categories of evidence must be presented in order for their claims to succeed, and thus, determines to a considerable degree the amount of enforcement costs that the parties will bear. More importantly, the legal standard serves to inform potential offenders – at an earlier stage – as to whether their intended agreement is of the type that may suffer the consequences of antitrust enforcement.

Note, however, that in the context of effects-based analysis this test should not be thought of as a detailed and precise rule expressed in the Treaty itself. Art 81 EC only describes the most basic features of the investigation to which restraints must be subjected. For the prohibition to apply, an agreement must suppress the process of

⁶ It is not suggested that uncertainty is the only factor explaining transgressions from EC competition law. (In this regard, see the discussion below, in Section 3.4.1, where it is explained that firms' incentives to cross the line, to stay in the clear, or to take too much caution, are determined by the interplay between the expected costs and benefits of the different forms of agreement between which they can choose.) The simple point that is made here is that we may expect that no transgressions would occur if enforcement could be structured in such a way as to give firms 100% certainty about (1) the dividing line between permissible and impermissible behaviour, (2) being caught in case of transgression, and (3) being punished in such a way that no net profits can be made.

rivalry between firms and, consequently, do more harm than good to consumers. Clearly, these abstract notions do not provide all firms with sufficient information to determine the legality of their contracts. Such certainty is only provided *ex post*, should the agreement be examined by the Commission or a court. *Ex ante*, firms may obtain additional guidance by looking at the relevant case law of the European Courts, the Commission's decision practice, or guidelines. In this way they can form an image of the likely implications for the agreements they wish to implement. Yet the specific circumstances of prior cases may be different from the situation that other firms face and guidelines may remain too abstract.⁷ Inevitably, therefore, they have to make their decisions to invest in distribution networks or in joint production facilities without absolute certainty about whether their agreement will produce the type of effect that will escape the ban of Art 81 EC.

And with a view to our analysis of the calibration of the legal standard in the next section it is important to realise that policymakers face similar difficulties when drawing up antitrust rules. They cannot easily foresee the welfare implications of every possible form of restriction under all possible market conditions. As is evidenced by the broad language of Art 81 EC, they must therefore decide how much to invest in gathering the necessary information to draw up clear and specific rules *ex ante* and how much of the detail should be left over to courts to fill in *ex post*.

Decision theory

These facts suggests that it is important to examine (1) how decision makers – whether they are firms deciding on contracts, parties determining their litigation strategy, authorities and courts trying facts, or policymakers crafting rules – that are faced with imperfect information may determine which decision is optimal and (2) under what circumstances they should make investments to increase the amount of

⁷ Moreover, as realists we are forced to admit that the law is not a watertight system. The facts pointed at by older cases and guidelines are not necessarily the only determinants of the outcome of proceedings. Other factors – a differently composed panel of judges or a different economic or political tide – may well play a role. On the other hand, this should not lead us to the conclusion that the legal standard in effects-based analysis is wholly indeterminate. The mode of reasoning employed in decisions and case law suggests that, as a general matter, precedents and (in the case of the Commission) guidelines do limit decision making and it makes sense, therefore, for firms to orientate themselves in the way described above on the legality of their contracts, even if the outcome will not be full proof. In this regard, see Maxeiner (2007) and the discussion on legal certainty below in Section 3.4.

information available.⁸ To understand these issues better, consider the position of a firm that wants to invest in its distribution network. Suppose that it faces a choice between two different forms of contract that it may sign with its resellers. The firm has no doubts that the antitrust authorities will not oppose the contract of type A. Type B, on the other hand, imposes considerable restrictions on the commercial freedom of certain market participants. It is expected that this alternative will produce €100 in extra profits. But the firm estimates that there is a probability of 0.5 that the agreement will be caught and terminated, in which event a fine of €90 is imposed and the profits are foregone. If no more information is available, the firm should compare the expected values of these alternatives to determine whether it should rely on contract A or B to structure its relations with resellers. Assuming that our firm is a rational profit maximizing agent,⁹ it will prefer B over its neutral alternative A (zero gains and zero losses), since it has a positive expected net benefit of €5 ($0.5 \times €100 - 0.5 \times €90$).¹⁰

Although this decision is perfectly rational, the agreement may subsequently still be caught and terminated. In other words, a rational *ex ante* decision may produce error costs *ex post*. Suppose that the firm can remove all uncertainty, for example, by notifying the agreement to the antitrust authorities as Regulation 17/62 once

⁸ These questions are dealt with by the branch of economics that is commonly referred to as decision theory. For a contribution that applies decision theory to decision making in antitrust see Beckner and Salop (1999), who provide extensive references to the literature in general. A third question that will arise – particularly in the context of adjudication – is, if information is to be gathered, exactly which information should be considered and in what order.

⁹ That is, we assume that if restricting competition is privately beneficial, the firm will enter into a contract with his rivals. Sanctions are effective in changing behaviour only if they lead to costs that exceed these benefits. One might react to this by saying that surely other factors than financial gain alone can have a bearing on the decision to stay within the bounds of the law. Particularly in fields such as traditional criminal law (i.e. excluding white collar crime) a perceived moral responsibility not to transgress will be of major importance in determining behaviour. This is certainly true. It should be realised, however, that the above discussion can easily be reformulated in terms of utilities (rather than euros). The profit maximization assumption implies that a person will prefer that course of action to which he attributes the highest utility. These preferences are likely to be influenced by the prospect of extra financial profit that the agreement could lead to and by the possible amount of a fine, but they may just as well be influenced by disutilities originating in moral perceptions. The outlook is in principle broader therefore. It should be noted, also, that even if it may be doubted whether individual persons will always act rationally, in antitrust offenders will generally be large corporate entities. Due to the fact that decision making processes in such firms will tend to incorporate shareholder interests, the arguments to expect their behaviour to conform to this assumption are much stronger.

¹⁰ The assumption here – and in the rest of this chapter – is that decision makers are risk neutral. A risk neutral decision maker weighs potential harms equally with potential benefits. In contrast, a risk averse decision maker places more weight on harms. Note, however, that even if a decision maker is risk averse, this does not render our decision-theoretic analysis irrelevant. It will simply mean that the downside potential of the decision will be given additional weight. In this regard, see Beckner and Salop (1999:52).

prescribed.¹¹ Such an investment in information would eliminate the possible costs of taking an erroneous decision. The firm would either know that alternative B is acceptable and earn €100 or, if it is not, it would implement agreement A. The expected benefit of the decision with the extra information is therefore €50 ($0.5 \times €100 - 0.5 \times €0$) whilst the expected benefit without the information is €5 (as calculated above). This means that the net incremental value of the information is €45. If the firm can obtain this information for less than this amount, therefore, it should do so (since the sum of error and investigation costs would remain lower than the total error costs in case no further investigation is engaged in).

Two key insights of the economic approach to decision making are reflected in these examples: (1) where information is limited, decisions should be based on a comparison of the expected value of the alternatives and (2) decision makers should try to minimize the combined costs of error and gathering information.¹² In the following sections we will build upon these insights to examine the decision making processes followed by policymakers and firms.

3.3. Calibration of the legal standard

In this section we adopt a policymaker's perspective and look at the rulemaking process. Our focus will be on the key characteristics of the effects-based standard that distinguish this method for analysing restraints from *per se* rules. In Section 3.3.1, we briefly revisit the reasons for working with a much more broadly pitched investigation and examine some of the implications of working with a standard rather than a rule. In Section 3.3.2, we study the considerations that may lead policymakers to divide the burden of proof between the enforcing party and the defendant.

3.3.1 Rulemaking: accuracy and timing

Components of the rule making calculus

Let us assume for the moment that the effects-based standard of Art 81 EC is composed of a number of rules regarding specific restrictive practices that are specified by policymakers *ex ante*. These rules should inform firms, litigants, and

¹¹ *Supra*, footnote 4.

¹² See Ehrlich and Posner (1974).

triers of fact on how to determine whether a restrictive agreement is caught. They must explain – in a greater level of detail than the broad language employed in the provision itself – the build-up and sequence of an acceptable argument that this provision has been infringed and they should designate the types of evidence that can be used to substantiate this argument. In his discussion on rulemaking in US antitrust Page (1989) suggests there are three variables that will determine the composition of these rules. These are: (1) policymakers' priors regarding the frequency with which a certain practice will lead to harm; (2) their estimates of the severity of the harm caused in such circumstances; and (3) their expectations as to the costs of producing the related evidence. Their respective roles can be appreciated by considering the following.

Theory developed in the Industrial Organization branch of economics will have a considerable impact on the formulation of such rules.¹³ It provides policymakers with models of reality that show how and under what circumstances certain behaviour will lead to consumer harm. Ideally, such a model will typically point to specific features of reality that allow us to distinguish a practice with restrictive consequences from efficient behaviour. The rules will direct the analysis of a restraint to this type of feature that is compatible with only one explanation. Related empirical studies will inform policymakers about the frequency with which specific practices can be expected to produce harmful and beneficial results.

On occasion, the rules of thumb derived from such empirical work may counsel against the investigation of the type of conclusive evidence advanced by a model. Consider the type of agreements subjected to *per se* rules, such as price fixing and market sharing. These practices are considered to be harmful in the vast majority of cases. Under a *per se* rule, therefore, it is sufficient that the mere existence of an agreement with stipulated features be shown, for the prohibition to apply, even if it is not the agreement itself, but its effect that is the reason for concern. Theorists might be able to show that under very specific circumstances such contracts lead to positive outcomes. But in any case this potential to contribute to consumer welfare will be minute. As a consequence, verification of the existence of these special circumstances would change the outcome of the decision in only a tiny fraction of cases. Treating a certain practice as a *per se* violation implies a policy decision, therefore, holding that

¹³ See the discussion on the impact of developments in Industrial Organization on antitrust policy in Chapter 5, Section 5.2.2.

the extra costs of a full inquiry into the effects in each case would be larger than the error costs that would be avoided by so doing.

The reason for subjecting other practices to the effects-based standard is that policy maker's estimates as to their effects are different. In line with modern economic insights, these type of agreements – vertical restraints, for example – are held to have a *non-negligible* potential both to decrease and increase consumer welfare. That is, they are expected by policymakers to produce positive result much more frequently than the type of practice subjected to *per se* rules. This uncertainty makes that requiring only that the existence of such an agreement be shown and then either allowing or prohibiting them, would poses a serious risk of errors. Such erroneous decisions are costly to society, (1) because they prohibit efficient behaviour or sanction harmful conduct and, crucially, (2) because they send out a message to firms in the market to repeat this error. Therefore, effects, or other circumstances than the agreement itself that have the potential to inform about effects, are taken into consideration in the assessment of these restraints.¹⁴

Differently so than with *per se* rules, however, these additional relevant circumstances are not specified in a great level of detail *ex ante*. For guidance on the application of the effects-based standard firms are very much dependent on case law and decision practice. Such decisions are generally geared towards the specifics of the case at hand and their implications for similar but different future cases will not always be unambiguous. In contrast, notices and guidelines may often be too abstract to resolve all questions that arise in an individual case. Gauging the outcome of

¹⁴ The same type of calculus will be used to determine whether new insights about the harmfulness of restrictive practices generated by Industrial Organization scholars should be incorporated in effects-based analysis. In general, such new insights will offer a different explanation for known behaviour, which can either be positive or restrictive in nature. As we saw, whether it is efficient to incorporate the associated sets of evidence into the legal standard depends, in the first place, on the impact on policymakers' prior conceptions about the frequency with which the effect described in the new model will occur in practice. This can be appreciated as follows. Before the development of new insights, policymakers will estimate that a practice with certain clear outward features will have a good or a bad effect in a proportion of g to b . A new model will shift this proportion. If the model advances an efficiency-based explanation, g will increase at the expense of b , and the opposite will happen if a hitherto unknown restrictive effect is shown. If policymakers consider the new model to be very powerful, and think that the explanation it offers will apply to many cases, then one can expect a substantial effect on the legal standard. The ultimate effect, however, is further determined by the two other factors mentioned. These are policy maker's beliefs about the effect on enforcement costs of changing the legal standard to account for the new model, and the error costs (in terms of wrong decisions and distorted incentives) that would be incurred if these changes are not or not fully made. Therefore, if priors change insubstantially, the added costs of gathering and reviewing the type of evidence that fits the model are large, and the harm to consumer welfare if the wrong outcome is reached in these cases is small, no effect on the legal standard may be expected.

effects-based analysis in individual cases will therefore often entail a considerable margin of interpretation and may lead to mistakes. This urges us to look deeper into policymakers' reasons for choosing a standard over a rule.

Timing: rules and standards

Let us start by clarifying the terminology. Rules and standards differ as to the extent to which the law is given content *ex ante* or *ex post* and, thus, as to the clarity of the signal that is sent to potential offenders.¹⁵ A rule will list distortive forms of conduct that firms may not engage in, whereas a standard will only proscribe distortive conduct as such, leaving the determination of which forms of conduct are harmful to adjudication, after the effect has been felt. Effects-based analysis clearly falls in the latter category, as it provides a case-by-case mechanism with which to determine whether agreements restrictive of the competitive process decrease consumer welfare. As argued by Kaplow (2000), the choice to employ a standard, rather than a rule, involves considerations regarding the cost of drafting rules of sufficient precision, the cost of errors due to the under or over-inclusiveness of rules, and the extra costs of legal proceedings due to the use of standards. To understand why effects-based analysis is standard-based,¹⁶ consider the following.

¹⁵ See Kaplow (1992, and 2000). The desirability of promulgating either rules or standards (the latter are also frequently referred to as principles) are the subject of a long-standing debate in legal theory that includes contributions by Dworkin (1967), Raz (1972), Kennedy (1976), and Alexander and Kress (1997). As is often the case, part of the confusion in this debate relates to differences in the definition of this distinction (in this regard, see Schauer, 1997). Despite the differences, however, most contributors do appear to agree on the more basic elements of this distinction. Generally spoken, they conceive of rules as specific prescriptions, whilst principles are seen as less specific or vaguer prescriptions (see Braithwaite, 2002). Goodin (1982) suggests that one can see this as a continuum. Prescriptions have a core meaning and a penumbra where their meaning is more uncertain. On rule-side of the spectrum the core is relatively large and the penumbra relatively restricted, whilst the opposite is the case on the standard-side of the spectrum. For our purposes, Kaplow's treatise of rules and standards – and his additional emphasis on the moment at which a prescription is best given substance – is particularly important because it adopts the perspective of policymakers. Most other contributions examine rules and standards through the lens of legal certainty. That is, they ask which type of prescription is preferable from a rule of law or moral point of view, focusing on the law's subjects rather than its drafters. In this regard, see the discussion in the text accompanying footnotes 18-20 and in Section 3.4.

¹⁶ It should be pointed out that in comparing EC and US antitrust law in Chapter 4, we will see that in practice effects-based analysis does not rely exclusively on standards. In the analysis of restrictive agreements with the potential to generate efficiencies both systems do rely substantially on standards, but make space for the application of rules where more certainty exists about the effects of a practice. In this regard, see the discussion (in Section 4.3.1.1) on truncated *rule of reason* analysis in US antitrust, the text accompanying footnotes 87-92 (in Section 4.3.1.2 on EC antitrust), and the discussion (in Section 4.4) of the implications of the comparison in terms of legal certainty (following footnote 140).

In traditional criminal law, for example, it will often be relatively easy to verify whether the law has been infringed. In principle, a broken window and a brick on the kitchen floor would suffice, as would a bullet wound on a corpse or an emptied out jewellery store. The difficulty in such cases will rather be to find the culprit. The situation is different in antitrust. Harm to consumers will often be much less evident. As we saw in Chapter 2, it tends to be spread out over numerous unidentified persons, who may be affected differently (some may buy at the higher price, others might spend their money on products that they value less). And to conceptualise harm in antitrust law we must always construe an image of the world without the restraint: how would consumers have fared in the absence of the restriction. In the case of *per se* infringements, this counter-factual will be rather straight forward: if the price fixing cartel had not been implemented prices would be lower, which is a very convincing reason to conclude that consumers would have been better off. When practices assessed under the effects-based standard are concerned, however, the construction of the relevant hypothetical becomes much more difficult to construct. The list of factors that may have to be taken into account is long. One may think of the market position of the parties, their suppliers and customers, about barriers to entry, capacity constraints, and product differentiation. All of these are matters of degree, and between them there may be complex inter-relations. A further complication is the fact that the world absent the restraint will also be the world absent the efficiency made possible by the restraint. In this context, European policymakers obviously estimate that *ex ante* formulation of sufficiently detailed (that is, neither too wide nor too narrow) rules covering all relevant situations is not feasible, or at least, prohibitively costly. These costs are avoided by use of the effects-based standard.

Errors in the application of standards

At the same time, it is important to note that the introduction of a standard implies a trade off between the costs of formulating sufficiently precise rules *ex ante* and the costs of erroneous decisions *ex post*. The uncertainty surrounding a legal standard may lead courts and other decisionmakers, incidentally, to sanction inefficient behaviour (a false positive, also referred to as a Type 1 error) or condemn efficient

behaviour (a false negative, also referred to as a Type 2 error).¹⁷ Apart from leading to an obvious and immediate waste of resources, such decisions may also produce wider adverse effects. This is because they send the wrong signal to other firms, who might copy this behaviour.¹⁸ And, more generally, it can be said that if a norm is given substance after the agreement has produced its effect on the market (*ex post*), firms will be at a relative disadvantage when it comes to assessing the legality of their intended agreement (*ex ante*). Instead of working with a rule that lists prohibited clauses (which are known at the time the agreement is concluded), they must (1) estimate – relying on their experience and knowledge of economics – what effect the agreement is likely to produce on consumers, and (2) assess how this effect would be evaluated in court. The fewer the constraints imposed by law (jurisprudence, secondary legislation, and decision practice) on the method by which proof of a certain claim must be provided, the wider the margin of appreciation that the trier of fact enjoys and the harder it will be for firms to predict the outcome at an early point. Staying in compliance with the law will, therefore, generally be less obvious when a standard is used.¹⁹ Ensuring compliance may also be expected to be more costly.²⁰ In addition, the broader scope of the *ex post* investigation of effects will come with costly extra investigatory efforts during litigation. We now turn to examine the division of the burden of proof in effects-based analysis, which is an instrument with which investigation and error costs can be minimised.

¹⁷ Note that in the context of antitrust, the application of the wider effects-based standard to potentially beneficial agreements can be expected to lead to more accurate results than if rules were applied (as was the case with the ‘old’ generation of block exemption regulations adopted by the Commission; in this regard see Chapter 2, Section 2.3). As argued by Kaplow (2000: 509), however, it is not generally true that standards will lead to more accurate precise results. In certain areas of the law, such as tax law, very detailed rules exist that can be expected to lead to more precise results than if an open-ended standard were used.

¹⁸ Generally, in a system based on deterrence these secondary effects will be of greater concern. This is so since firms are inspired to comply by means of a few interventions, rather than by screening all or the majority of potentially restrictive contracts, as done under the notification regime of Regulation 17/62 (*supra*, footnote 5). In this regard, see the discussion in Chapter 2, Section 2.3.

¹⁹ See Kaplow (2000: 510). In this regard, two remarks must be made. In the first place, compliance depends on more than the ability to assess the legality of an intended agreement. The expected profits associated with the available alternatives, the probability of being challenged, and the height of the fine that might be imposed will also play a role. A more detailed discussion follows in Section 3.4, and in Chapter 5, Section 5.3. Secondly, it has been suggested (Braithwaite, 2002) that in complex situations subject to continuous change principles may offer more legal certainty than rules. He gives an example of telecommunications regulation. What constitutes a telephone today may be something quite different tomorrow. In such instances a principle will provide more long-term certainty than a detailed rule that would have to be adapted continuously.

²⁰ Provided, of course, that firms have sufficient *ex ante* incentives to become informed about the law (in this regard, see Kaplow, 2000: 510) and on whether the evidence that will be taken into account *ex post* is of such a nature that it can be anticipated (in this regard, see Kaplow, 1994: 365). As regards the cost of self assessment, see the discussions in Chapter 6, Section 6.2, and Chapter 7, Section 7.2.

3.3.2 Dividing the burden of proof

The division of the burden of proof between the enforcing party and the defendant is an essential aspect of effects-based analysis. We saw that it is one of the main characteristics that distinguish this method of investigating restraints from *per se* (object-based) analysis. To understand its role within the enforcement mechanism, it is important to underline what we are dealing with here.

Like the question regarding the appropriate legal standard,²¹ the question of the assignment of the burden of proof relates to evidence. It concerns the question of which party carries the burden to produce evidence in support of his claim concerning the existence of a restriction.²² This assignment can be made in three ways. The burden can be placed entirely on either one of the parties.²³ Alternatively, the burden to produce the necessary evidence is divided between them. It is easily appreciated that these options are not interchangeable, in terms of their effects on enforcement. Assigning the burden of proof entirely to the Commission will limit its capacity to bring cases. A very low burden, on the other hand, might render its enforcement efforts unfocused. Hay and Spier (1997) offer an analysis of the conditions that determine the optimal assignment of the burden of proof. To do so, they use a model of litigation that is moulded on US civil litigation. As is explained below, antitrust enforcement by the European Commission is considerably different. Nonetheless, their analysis provides a useful starting point for our study of the division of the burden of proof in Art 81 EC proceedings.

Hay and Spier begin by considering the situation where the burden is assigned in its entirety to one party. They assume that only the party that carries the burden will introduce evidence and that he will do so if and only if evidence exists to support his claim.²⁴ This assumption urges them to consider the probability with which violations

²¹ See Section 3.2. There it was suggested that the legal standard informs firms and litigants about the kind of evidence that can be relied on to show that competition has illegally been restricted.

²² A third question concerns the standard of proof: what level of conviction must the trier of fact reach to be able to conclude that the evidence produced allows for a finding that a violation has occurred?

²³ In the context of European antitrust, this would mean that the Commission is charged with showing that the net effect of the agreement on consumer welfare is negative or that the defendant is charged with showing that the ultimate effect on consumer welfare is positive.

²⁴ This assumption follows from their assumption regarding the standard of proof (*supra*, footnote 22). In their model courts are instructed to find that a violation has occurred if, on the basis of the evidence produced, they consider it more probable that a violation has occurred than not. As a consequence, if evidence exists to support the claim of the party that does not bear the burden, he will choose not to introduce it. This can be appreciated as follows. The party that carries the burden (B) will not make

occur.²⁵ Suppose that this probability is very small.²⁶ In that case, a rule that assigns the burden to the plaintiff will result in fewer instances where costs are incurred to produce evidence, than a rule that places the burden on the defendant.²⁷ They then refine this basic model by considering what happens when, at the start of the proceedings, the court is presented with an initial signal regarding the occurrence of a violation in the case at hand.²⁸ If this signal provides only a weak indication, the burden is best assigned to the plaintiff.²⁹ If, on the other hand, the signal is strong, the burden is best placed on the defendant, particularly if he can produce the evidence at lower cost.

At this point, we are also able to see the possibilities offered by dividing the burden between the parties.³⁰ Assume that (1) the evidence by which it can be shown that a violation has occurred is composed of several elements, (2) that the first of these elements – the initial signal – is not very telling, but that (3) at least two other

expenses to show the court the available evidence, since it does not support his claim. In that case the standard of proof instructs the court to decide in favour of the other party (N). N does not need to make expenses and show what evidence exists to win the case. It is argued below that this crucial assumption in the model developed by Hay and Spier cannot be transposed to the setting of European antitrust.

²⁵ For example, the likelihood that a failed surgical operation is the result of negligence on behalf of the surgeon or the probability that inter-firm contracting results in consumer harm.

²⁶ For the moment, it is assumed that parties incur the same costs in producing the same evidence. The implications of possible differences between the parties in terms of the costs of gathering evidence are discussed below.

²⁷ The authors use the following expression to illustrate their analysis. The burden should be assigned to the plaintiff if:

$$p(I) \times C_p < p(\neq I) \times C_d$$

where $p(I)$ stands for the probability that an infringement occurred and C_p denotes the plaintiff's costs, whilst $p(\neq I)$ stands for the probability that an infringement did not occur and C_d denotes the defendant's costs.

²⁸ Here, one can think of a clause in an agreement that by its nature limits the commercial freedom of one of its signatories, or of a letter in which the defending manufacturer refuses to supply a distributor. Using Bayes' rule, Hay and Spier express this probability that the law was infringed or was not infringed, given the signal (E), as follows:

$$p(E|I) \times p(I) / p(E) \text{ and } p(E|\neq I) \times p(\neq I) / p(E)$$

where $p(E|I)$ indicates the probability of observing the signal in the event that a violation has occurred, $p(I)$ stands for the probability that an infringement occurred, and $p(E)$ is the probability that we observe the signal, whilst $p(E|\neq I)$ indicates the probability of observing the signal in the event that an infringement has not occurred, and $p(\neq I)$ denotes the probability that an infringement has not occurred.

²⁹ By plugging the conditional probabilities given in footnote 27 into the expression of footnote 26 and simplifying, Hay and Spier show that the burden should be assigned to the plaintiff if:

$$p(E|I) \times p(I) \times C_p < p(E|\neq I) \times p(\neq I) \times C_d$$

This expression makes it easy to appreciate that if $p(E|I)$ and $p(I)$ are low and costs not too far apart, the burden is best assigned to the plaintiff.

³⁰ Hay and Spier (1997) do argue that shifting the burden of proof may be optimal in circumstances where the initial signal provides a high level of certainty that a violation occurred, but they do not provide theoretical support for this argument. What follows in the main text is but a rough description of the conditions under which a division could be beneficial in the context examined by Hay and Spier. Since our focus is on the rather different situation of European antitrust, there is no need here to fully develop this argument.

types of evidence exist that are both substantially more indicative of harm, with (4) type A slightly less indicative and less costly to produce and type B slightly more indicative and more costly to produce, such that (5) evidence of type B will exonerate the defendant in a fraction of cases where exclusive reliance on evidence of type A would lead to a condemnation. In such cases dividing the burden of proof may allow for savings vis-à-vis the situation where the burden is entirely placed on the plaintiff.³¹ This is because the defendant can be expected to invest in showing evidence of type B only in the relatively few cases where it will save him. If this fraction of cases is sufficiently small and the costs of showing both types of evidence at the same time not too large in comparison to the error costs that are thus avoided, shifting the burden can consume less resources vis-à-vis the situation where the plaintiff always has to show evidence of type B or where the case is decided on the basis of the plaintiff's showing of evidence of type A (which implies accepting a margin of error).

There are considerable differences between US civil litigation and the administrative law setting of antitrust enforcement by the European Commission. Particularly the vital assumption made by Hay and Spier – that only the party that carries the burden will introduce evidence – does not reflect the realities of European antitrust. When the Commission investigates a case it will make a full assessment that covers all aspects of Art 81 EC, including the existence of efficiencies.³² Defendants, in turn, will generally also provide arguments and evidence on all accounts.³³

There are good reasons to assume, however, that the shift of the burden of proof allows the Commission to incur less investigation costs than it would if it were fully responsible for both elements of the analysis (that is, Art 81(1) and (3) EC). The

³¹ In the antitrust context A and B must be seen as proxies for consumer harm. A might, for instance, be a showing of market power by means of market shares (where a high level of concentration will lead to a finding of an infringement), whilst B could represent the ultimate effect of the agreement on output (in terms of volume, quality, etc) and prices.

³² See e.g. Ortiz Blanco (2006: 384).

³³ This state of affairs is clearly reflected in the Commission's infringement decisions. In *Master Card* (decision of 19 December 2007, published on the Commission's website) the defendant presented numerous detailed arguments regarding market definition (which falls under the Commission's burden) that were supported by several expert studies (see e.g. para. 254, 260, 291, and 293); in *Morgan Stanley / Visa* (Commission decision of 3 October 2007, published on the Commission's website) the defendant adduced evidence regarding a drop in merchant service charges to show that the Commission's allegations under Art 81(1) EC did not stand up to scrutiny (see para. 199); the decision in *EATA* ([1999] OJ L193/23) shows that the parties adduced several arguments – supported by evidence – concerning market definition (see e.g. para. 80, 81, 89, and 93); and in the case of *Van den Bergh Foods* ([1998] OJ L246/1), the defendant commissioned several studies relating to the definition of the relevant market (see para. 115).

Commission's Guidelines on the application of Art 81(3) EC indicate that it relies to a considerable extent on arguments and evidence provided by the defendant to complete its initial assessment of the efficiencies generated by an agreement. Explicitly referring to Art 2 of Regulation 1/2003 concerning the division of the burden of proof in Art 81 EC,³⁴ the guidelines point out that this provision can be invoked as a defence and that in so doing the undertaking concerned must substantiate its claims so that they can be verified and balanced against the negative effects identified in the analysis under Art 81(1) EC.³⁵ In other words, the Commission does not intend to deepen its initial assessment on this point by gathering detailed evidence should the defendant fail to do so. Recent decision practice reflects this stance. The Commission's analysis under Art 81(3) EC centres on the arguments and evidence presented by the defendant, whilst evidence introduced by the Commission itself generally plays a minor role.³⁶ Arguably, if the Commission carried the entire burden under Art 81 EC,

³⁴ Supra, footnote 3. As regards the dividing line between Art 81(1) and (3) EC, see the discussion in Chapter 2, Sections 2.3 and 2.4.3.

³⁵ [2004] OJ C101/97, at para. 41 ('...the burden of proof under Article 81(3) rests on the undertaking(s) invoking the benefit of the exception rule.'). 55 ('...efficiency claims must be substantiated so that they can be verified.'). 56 ('In the case of claimed cost efficiencies the undertakings invoking the benefit of Article 81(3) must as accurately as reasonably possible calculate or estimate the value of the efficiencies and describe in detail how the amount has been computed. They must also describe the method(s) by which the efficiencies have been or will be achieved. The data submitted must be verifiable so that there can be a sufficient degree of certainty that the efficiencies have materialised or are likely to materialise.'). and 57 ('In the case of claimed efficiencies in the form of new or improved products and other non-cost based efficiencies, the undertakings claiming the benefit of Article 81(3) must describe and explain in detail what is the nature of the efficiencies and how and why they constitute an objective economic benefit.'). These Guidelines were discussed above in Chapter 2, Section 2.4.

³⁶ The four infringement decisions – where most evidence can be expected to surface – listed in Appendix A are most interesting to look at in this regard. In the case of *Master Card* (supra, footnote 33) the discussion concerning Art 81(3) EC (para. 666-753) fully concentrates on refuting arguments and questioning the evidence presented by the defendant (to the effect that a fixed default fee for inter-bank services served to expand the over all output of its credit card scheme). By comparison, the Commission introduces very little evidence (at para. 696 it discusses evidence that it has presented in the assessment under Art 81(1) EC). The same can be said for the Commission's decision in the case of *Morgan Stanley / Visa* (supra, footnote 33, at para. 311-319 and also at 254-296). In para. 313, the Commission states that 'Visa should identify and determine the specific nature of the efficiencies generated by the Rule as applied to Morgan Stanley (that is to say, it should demonstrate a causal link between the application of the Rule and the efficiencies), and demonstrate how and when each claimed efficiency would have been achieved. In order to establish that the efficiencies generated outweigh their anticompetitive effects, it should establish the likelihood and magnitude of such efficiencies'. Similarly, in the cases of *EATA* (supra, footnote 33, at para. 187-237) and *Van den Bergh Foods* (supra footnote 33, at para. 222-238), the Commission's assessment under Art 81(3) EC consists of a discussion (and refutation) of a whole list of efficiency arguments presented by these defendants, rather than that it discusses the results of its own investigation. The above is not to say that there may not be individual decisions where the Commission makes considerable efforts in the assessment under the third paragraph. The exemption decision in the case of *Bass* ([1999] L186/1), where the Commission set out to quantify efficiencies claimed by the defendant (a detailed discussion of this case follows in Chapter 4, Section 4.3.1), provides an example. Particularly in the current context of ex post enforcement under Regulation 1/2003 (supra, footnote 3), however, these are likely to be exceptions.

an initial assessment that only filters out cases that produce obvious efficiencies would not suffice to discharge its burden under the third paragraph. This would make Art 81(3) lose much of its effectiveness and would turn Art 81 EC as a whole into a rather blunt instrument.

To structure the analysis made in the coming chapters, it is worth examining these implications of dividing the burden of proof in effects-based analysis from a theoretical perspective. For purely descriptive purposes, a number of formulas are introduced.³⁷ Let us assume that the value of enforcement to society (V) is determined by the amount of harm to consumers that is averted (B) and the legal costs that are incurred in so doing (C):

$$V = B - C \quad (1)$$

Let us begin by considering the benefits of enforcement.³⁸ In principle, these benefits depend on the number of cases pursued by the Commission in which it takes a good decision (that is, terminates an agreement that is harmful to consumers). Dividing the burden of proof affects these benefits in two ways. Firstly, by lowering the Commission's per case investigation costs, it increases the total amount of cases that the Commission is able to bring (assuming that it disposes over a fixed budget).³⁹ Secondly, by allowing the Commission to proceed on the basis of *prima facie* evidence, a risk is introduced that selection errors are made. To appreciate why,

³⁷ The formulas are introduced so as to facilitate the appreciation of the discussion that follows. We will see that a broad variety of factors weigh in on the costs and benefits of dividing the burden of proof. After discussion of each of these factors, the argument made is summarized in the form of a descriptive formula. By integrating these formulas at the end, the reader is offered a simple and short summary of the entire argument made in this section. Please note that these formulas are not used as the basis of a formal model.

³⁸ The benefits of enforcement are built up of two components. They include the harm that is avoided (1) by terminating the restrictive agreement in the individual case at issue and (2) because in the future firms in a similar position will abstain from signing an agreement with features singled out in the decision. In this regard, see Section 3.2 above.

³⁹ It can be assumed that with each additional case that the Commission is able to pursue and in which it takes a good decision, the benefits of enforcement grow in absolute terms. If the Commission sets the right enforcement priorities and deals with the cases that yield most enforcement benefits first, these positive returns on additional enforcement efforts can be expected to decline gradually. The harm that is avoided by terminating an additional restrictive agreement will be smaller than the harm that was prevented in previously challenged cases and the deterrent effect of an additional decision where many exist already will be relatively smaller as well. In the context of European effects-based analysis, where, as we will see in Chapters 4 and 5, decisions are few and there is thus a scarcity of guidance, it seems fair to assume that additional enforcement efforts – provided that they result in insightful decisions – will produce substantial positive returns. This is an argument in favour of dividing the burden of proof, therefore.

consider the situation where upon a showing of a restrictive clause in an agreement, the entire burden of proof under Art 81 EC would shift to the defendant. This would allow the Commission to challenge firms on the basis of a rather superficial investigation of their agreement. Since we know that restrictions are not necessarily harmful, we must expect that such a division would result in relatively unfocused enforcement. Amongst the restrictions selected to be challenged, there will some that produce no harm to consumers.⁴⁰ The proportion of cases in which the Commission gets it right – that is, in which it selects a case that is worth pursuing – will depend on the quality of the initial signal. The expected benefits of enforcement can therefore be expressed as follows:

$$B = p(I|E) \times (R/C_C) \times W \quad (2)$$

(where $p(I|E)$ stands for the probability that the investigation produces the right result (I) when evidence of a certain type (E) is used in the initial assessment under Art 81(1) EC, R stands for the part of the Commission's budget devoted to antitrust enforcement under the effects-based standard of Art 81 EC, C_C denotes the Commission's costs of enforcement per case, and W stands for the welfare increase due to successful enforcement.⁴¹

Now let us consider the costs of dividing the burden of proof. As suggested, the Commission's per case investigation costs will go down.⁴² At the same time, however,

⁴⁰ This argument is further developed below. It is worth noting that there is a certain similarity between this argument and the long-standing debate in US antitrust on the appropriateness of the treble damages provision of Clayton Act Section 4 (15 U.S.C.A. §14). The question there is whether multiplying damages by three might give private claimants incentives to file suit in cases with little merits, that is, in cases with a low probability of success. In this regard, see e.g. Landes (1983), Baumol and Ordover (1985), Easterbrook (1985), Cavenagh (1987), and Waller (2003). Here, it is submitted that making it too easy for enforcers to present a *prima facie* case could lead to low returns on enforcement efforts.

⁴¹ In practice, the benefit derived from each case will depend on the seriousness of the infringement and may, thus, vary per case. Here, it is assumed that all cases resulting in the termination of a restriction produce benefits of the same magnitude. More specifically, the term W should be thought of as representing the cumulative benefit derived from all cases challenged – in the situation without a division of the burden of proof, where no selection errors are made – divided by the number of cases that are brought. It is appropriate to do so, since we are studying the division of the burden of proof from the perspective of policymakers. They will not be interested in individual cases, but in enforcement in general and, thus, in average enforcement costs.

⁴² We may expect that in the longer run this effect on investigation costs will diminish. This can be explained as follows. Cases assessed under the effects-based standard will generally produce both negative and positive welfare effects. If the Commission sets the right enforcement priorities and deals with the cases that yield most enforcement benefits first, this means that the cases where the difference between these two effect is largest will be prioritized. With each additional case this difference will grow smaller and more detailed and compelling evidence will needed to show that the negative effect

we may expect the defendant's costs to rise by a similar margin. The net cost effect per case will then be neutral, but – given that the Commission is able to challenge more restrictions – total defence costs and, thus, the overall costs of antitrust enforcement will rise. This may be different when the defendants have a cost advantage over the Commission in gathering the evidence concerned, as may be the case with evidence regarding improvements in production or distribution brought about by the agreement (the first condition of Art 81(3) EC).⁴³

This, however, is not the whole story regarding the cost effects of dividing the burden of proof. An extra error cost component will have to be taken into account. It was suggested above that some harmless agreements may be subjected to scrutiny and may even be made the object of decisions. This tendency towards Type 2 errors is explained by the structure of the investigation under this provision. Art 81(1) EC essentially functions as a first filter that distinguishes potentially harmful agreements (for example: agreements that restrict the competitive process and enhance market share) from agreements that are clearly not intended to be caught (for instance: a non-compete clause in a partnership or agency agreement). Some of the restraints caught by Art 81(1) EC will only produce harmful effects for consumers. We saw, however, that others will produce opposing effects – certain vertical restraints may, for example, simultaneously increase the price and improve the quality of a product. Such opposing effects are examined under the scope of Art 81(3) EC. This implies that the criterion used in the analysis under the first paragraph (a restriction of the competitive process, possibly combined with an increase in the parties' market share) inevitably catches a larger group of agreements than will ultimately be held to infringe Art 81 EC. Since we are dealing with the introduction of an over-inclusive first test, there can be no risk that dividing the burden of proof will lead to mistaken acquittals (errors of the first type).⁴⁴

prevails. As a consequence, per case investigation costs will gradually increase. This effect may eventually undo the effect of dividing the burden of proof.

⁴³ It may also be the case, however, that the Commission has a cost advantage over the defendant as regards certain elements of the latter's burden of proof. As suggested, when it comes to efficiencies it is intuitively appealing to say that the firm that will realise the claimed cost reductions or quality improvements should be better positioned to produce supporting evidence. This is not necessarily the case though when it comes to the other elements of Art 81(3) EC, since they require that the effect of these efficiencies on competition in the market and consumers is shown. Arguably, with its extensive investigative powers, the Commission is better able to obtain data about market-wide effects.

⁴⁴ This is not to say that Type 1 errors cannot occur in European antitrust. It is true that under Regulation 1/2003 (supra, footnote 3) the Commission no longer clears agreements (note that commitment decisions constitute implicit prohibitions of the agreement as originally conceived by the

In cases where these Type 2 selection errors are made, the benefits of enforcement do not accrue.⁴⁵ In fact, the expenditures made by the Commission and the defence in these cases are resources wasted. If such selection errors are not filtered out during the proceedings and result in an infringement decision or – which is more likely in the context of effects-based analysis under Art 81 EC – in the modification of the original agreement, additional losses will be incurred, since this would mean that an agreement that is harmless to consumers (or even beneficial) and profitable to the firm in question is terminated. These costs will be magnified by the effect on other firms active on the Common Market, who may end up taking unnecessary precautions or signing harmful contracts.⁴⁶

Moreover, we should realise that easing the test that the Commission has to satisfy will affect the benefits of enforcement in yet another way. The fewer efforts we require the Commission to make to ensure that a restriction of the competitive process results in sufficient harm to justify the legal costs needed to bring it to an end, the greater the risk becomes that amongst a pool of harmful agreements to challenge, the Commission will select a number that would not have been given priority if it were held to a more stringent test.⁴⁷ Both types of welfare losses – terminating harmless cases and selecting low-value cases – are not reflected in the discount on enforcement benefits discussed above and must therefore be added to the overall costs of dividing the burden of proof:

$$C = (R/C_C) \times [((1-e) \times C_C) + ((1+e) \times C_D) + (p(T|E) \times W)] \quad (2)$$

where e represents a factor ranging between 1 and 0 by which the Commission's enforcement costs are discounted and by which the defendant's enforcement costs (C_D) are marked up,⁴⁸ and $p(T|E)$ denotes the probability that relying on a specific

defendants), but it remains possible that a good decision is mistakenly annulled in appeal. This, however, would have nothing to do with the division of the burden of proof. What is argued is that shifting the burden of proof will only engender Type 2 errors, for the reasons mentioned above.

⁴⁵ This is reflected in expression (2) by multiplying B by $p(I|E)$, which implies that in $p(\neq I|E) \times 100\%$ of the cases enforcement will not produce positive results.

⁴⁶ The effects of errors on the behaviour of firms who are in the process of designing and implementing agreements with possible antitrust implications are examined in the following section.

⁴⁷ This could be because the proceedings in question produce few benefits in comparison with other restrictions that may be investigated and challenged (with the risk that these benefits are outweighed by the investigation costs).

⁴⁸ The implications of possible cost advantages for the defendant are considered below. See the text accompanying footnote 52.

type of evidence in the investigation under Art 81(1) EC will lead to the kind of welfare losses described above.⁴⁹

These considerations may help us decide whether in a given situation dividing the burden of proof may increase the efficiency of enforcement. By inserting (2) and (3) into (1) we obtain the following expression:

$$V = (p(I|E) \times (R/C_C) \times W) - (R/C_C) \times [((1-e) \times C_C) + ((1+e) \times C_D) + (p(T|E) \times W)] \quad (4)$$

Suppose that policymakers are faced with the choice either to allow the enforcing party to make a *prima facie* case on the basis of evidence of type E_1 or to burden him with the task of coming forward with all the evidence, in which case the type of error costs discussed before will not emerge. Dividing the burden of proof will be beneficial if:

$$(p(I|E_1)/(1 - e_1) \times W) - (p(T|E_1)/(1 - e_1) \times W) - (((1 + e_1)/(1 - e_1)) \times C_D) > W - C_D^{50} \quad (5)$$

This expression reflects a very basic intuition, notably, that the sum of (1) the benefits of enforcement with a divided burden (which are reflected in the term $p(I|E_1)/(1 - e_1) \times W$), (2) the extra error costs (represented by $- p(T|E_1)/(1 - e_1) \times W$), and (3) the extra defence costs (represented by $- ((1 + e_1)/(1 - e_1)) \times C_D$) must be larger than when the burden remains undivided.⁵¹ It is easily appreciated, also, that this will sooner be the case if defence costs rise by less than per case enforcement costs decline, that is, if defendants are at an advantage when it comes to providing proof regarding specific aspects of the legal standard.⁵² Finally, the same reasoning can be applied to examine whether it is appropriate to alter the Commission's burden of proof in effects-based

⁴⁹The two types of error costs discussed above are lumped together in the term $p(T|E) \times W$. As was done when describing the benefits of enforcement (see footnote 41), the error costs caused by targeting harmless agreements and by bringing low-value cases are assumed to be equal in all cases. This term $p(T|E)$ may thus be thought of as indicating the proportion of cases in which costs of the magnitude of W are incurred (this proportion is obtained by taking all the error costs spread out in smaller and larger amounts over cases involving harmless agreements and cases involving low-value infringements and dividing them by W). This probability will increase as the test to which the enforcer is held becomes less stringent.

⁵⁰ This is obtained by (1) equating the situation with a divided burden to the situation where the Commission is responsible for the whole burden and (2) simplifying the result. Note that in the latter situation no division-related error costs are incurred (see footnote 39) and the factor e is set at 0.

⁵¹ Since it is assumed that in either scenario the Commission disposes over the same budget (and depletes it), enforcement costs (C_C) have no influence on the trade off and are left out of this expression.

⁵² Since the term $((1 + e_1)/(1 - e_1))$ will then become smaller.

analysis under Art 81 EC.⁵³ Whether it would be efficient to shift the burden to the defendant at an earlier or later stage of the investigation depends on whether this would result in a more favourable mix of genuine infringements terminated, defence costs, and error costs.

To deepen our understanding of the effects of errors, the final section of this chapter examines how firms that are in the process of designing and implementing an agreement with possible antitrust implications react to unfocused enforcement, that is, to uncertainty about the precise location of the legal standard.

3.4 Firm behaviour under uncertainty about the legal standard

Before entering into a detailed discussion of firm behaviour under conditions of uncertainty, we do well to briefly consider the role of legal certainty in modern systems of law and, in particular, in European law. The protection of legal certainty is regarded as a crucial requirement for the rule of law.⁵⁴ In modern society, the powers of the state to formulate and apply law and legal commands are subjected to clear restrictions.⁵⁵ According to legal theory, laws should be validly made and publicly promulgated, of general application, stable, clear in meaning and consistent, and ordinarily prospective. In addition, their application should be impartial, provide parties who are sanctioned an opportunity to be heard, and deliver predictable, consistent decisions in individual cases.⁵⁶ The law's subjects, in turn, are expected to behave prudently and to inform themselves about the law and – to a certain extent – the evolving circumstances that determine its interpretation.⁵⁷ Together, these prescriptions fulfil a fundamental ordering function. They enable the law's subjects to set out a course of action (and to make the related investments) which they may legitimately expect to remain free of state intervention.

⁵³ In this regard, think of a setting as described in the text accompanying footnotes 31 and 14.

⁵⁴ For a more extensive discussion of these matters, see Maxeiner (2007).

⁵⁵ The rule of law is used here in the narrow – legal – sense of the word. When used in a broader sense this term is often infused with ideals of a liberal and democratic state, such as democracy, constitutionalism, human rights, and a free-market economy. In addition, it may extend over the fairness of the formulation or application of rules. As such, it is a much more contested concept that means different things to different people. In this regard, see e.g. Peerenboom (2004) and Maxeiner (2006).

⁵⁶ These requirements can, for example, be found in Fuller (1971), Summers (1999), and Peerenboom (2004).

⁵⁷ See Usher (1998: 59) and Popelier (1997: 167) as regards the requirements imposed on those subject to European law.

These concerns are reflected in European law. The protection of legal certainty is recognised as a general principle by both the ECJ and the European Court of Human Rights.⁵⁸ In the words of the latter institution, this principle requires that ‘all law must be sufficiently precise to allow the person – if need be, with appropriate advice – to foresee, to a degree that is reasonable under the circumstances, the consequences which a given action may entail.’⁵⁹ Yet the principle of legal certainty in European law is not absolute. The ECJ recognises that a certain measure of flexibility is essential if the Community’s policy objectives are to be met in situations of complex, changing, or evolving circumstances.⁶⁰ In the context of antitrust law, this is, for example, reflected in the *Remia* judgment, where the Court considered that the Commission enjoys a margin of appreciation in the evaluation of complex economic situations.⁶¹ From the point of view of undertakings, the obvious result is reduced determinacy of the law. The present section contributes to our analysis of how well the effects-based standard performs its ordering function by exploring economic insights concerning firm behaviour under conditions of reduced legal certainty.

With uncertainty, potential offenders are unable to determine with precision which commercial practices are likely to violate antitrust laws. This is to say that firms reckon with the chance of being found in violation when considering whether to sign a contract that will not harm consumers. Alternatively, they may take into account the possibility that a pact that does contravene the antitrust laws will be cleared. There are two separate strands of literature that are instructive with respect to how firms may operate under such uncertainty.⁶² Calfee and Craswell (1984) and Goetz (1984) reach

⁵⁸ See e.g. Popelier (1997: 162), Usher (1999: 65), Tridimas (2006), and Groussot (2006). It is a fundamental concept of the national legal orders, also. Maxeiner (2007: 549) provides an overview of member state law with references, and detailed information on the importance of the principle of legal certainty in German law. Popelier (1997) discusses both German and Belgian law in addition to European law. In this regard it should be noted that the general principles of EU law, which are prescriptions that are not explicitly expressed in the treaties, are to a considerable extent influenced by concepts developed in the national legal orders. The other source from which general principles are derived are the essential characteristics of the EU’s legal order itself. See e.g. Usher (1998) and Tridimas (2006). For ECJ case law on the issue of legal certainty, see e.g. *Duff v. Minister of Agriculture* (Case C-63/93, [1996] ECR I-569) and – in relation to the direct effect of Art 81(1) EC – *Brasserie de Haecht v. Wilkin* (Case 48/72, [1973] ECR 77). As to the case law of the ECHR, see e.g. *Sunday Times v. United Kingdom*, judgment of 26 April 1979, Series A, no. 30, 2 EHRR 245, and, more recently, *Korchuganova v. Russia*, judgment of 8 June 2006 in case 75039/01. The sources cited above provide extensive references to the case law of both Courts.

⁵⁹ *Sunday Times v. United Kingdom* (supra, footnote 53, at para. 49).

⁶⁰ See Popelier (1997: 167).

⁶¹ Case 42/84 *Remia et al. v. Commission*, [1987] ECR 2545, at para. 34.

⁶² For an overview of this literature see Schwartz (2000). Note that the focus is on the implications of uncertainty for the behaviour of potential offenders, and not so much on the negative effects of a

the conclusion that – depending on the circumstances – operators may respond either by over-complying or violating. Png (1986) and Polinsky and Shavell (1989), by contrast, find that under-deterrence is the only negative effect that can result from uncertainty. Section 3.4.1 discusses how differences in the underlying situations that are modelled lead to these diverging conclusions. Section 3.4.2 examines which of these models can be expected to describe the situation in European antitrust most accurately.

3.4.1 Analysing the effects of uncertainty

Binary settings

Png (1986) and Polinsky and Shavell (1989) analyse situations in which the choice is between engaging in a potentially prohibited practice and desisting from it. It can easily be explained that, in such a setting, uncertainty increases incentives to violate. At the beginning of this chapter we discussed an example of a firm having to choose between a more and a less restrictive form of distribution agreement. The first alternative affords the firm some measure of market power and produces extra profits. But, as a consequence, the firm also estimates that there is a probability that the agreement will be caught and terminated, in which event a fine is imposed and the profits are foregone.⁶³ To make his choice, we said, the firm will compare B's expected costs and benefits. If the result is positive, it will choose B as it is more profitable than the neutral option A.

Now consider what happens if uncertainty regarding the location of the legal standard is introduced in the form of (1) the possibility that even if B turns out to have negative effects, the court might acquit him, or (2) the (perhaps remote) possibility that even if he chooses A, he might still be found in violation. Both possibilities tend to make option B relatively more attractive. The possibility of false acquittal (a Type 1 error) increases the net expected value of option B, since the probability of the agreement being challenged and terminated becomes smaller. The prospect of false

mistaken decision *per se*, i.e. the immediate welfare loss of prohibiting a beneficial agreement or sanctioning harmful conduct. We look therefore not at the welfare effects in individual cases, but at the effects of such decisions on all those third parties that base their actions on their best understanding of the relevant case law.

⁶³ Note that in this example (and differently so than in the example given in Section 3.2) the firm is uncertain about whether the agreement will be discovered by the antitrust authorities or by private enforcers. It entertains no doubt about the illegal nature of the form of agreement it contemplates.

conviction (a Type 2 error) makes choosing A potentially costly, and therefore lowers the net expected value of this option. Under these circumstances, therefore, where firms choose between violating and not violating, uncertainty about the outcome of an investigation by the Commission will always strengthen incentives to break the law. This is the result, also, that is obtained by Schinkel and Tuinstra (2006) in their treatise on the effects of imperfect competition law enforcement on incentives to engage in price fixing behaviour.⁶⁴

Continuous range of options

By contrast, Calfee and Craswell (1984) and Goetz (1984) urge us to consider a situation in which firms have more than two options to choose from. When studying price fixing and *per se* conduct in general, it is natural to present the trade off that firms face as a binary affair. Either they collude or they do not and there is no intermediate form. As is argued in more detail below, in the effects-based field firms may often face a choice between more than two options. This has significant implications for the possible effects on firms' decisions of uncertainty about how the antitrust authority will evaluate these forms of agreement.

Imagine that our firm has a choice between three or more different ways to design his distribution contracts. Rather than supplying all willing resellers in an area, the firm may, for instance, consider whether to work with a number of selected distributors, or he may appoint an exclusive distributor. If he opts for selective distribution, he may be able to choose between various methods for selecting his retailers. In case he decides in favour of exclusivity, he will have a choice whether to

⁶⁴ These authors describe a situation with three types of industries (combinations of firms). Two of these industries have the option either to behave competitively or to collude; the third industry always behaves competitively. One of the two industries that may collude can choose between a more and a less collusive alternative, but it is clear that both these options are squarely illegal. On this basis the authors examine how the choice between these alternatives is influenced by imperfect competition law enforcement, that is, by the possibility that assessment errors are made by antitrust authorities. As suggested, they obtain the same general result as Png (1986) and Polinsky and Shavell (1989). In particular, they show that the incidence of anticompetitive behavior may increase in the enforcement error, essentially for two reasons. The first is that the expected sanction for law breaching decreases, due to possibility of firms escaping without a penalty, even when monitored. The second reason is that firms that would otherwise behave perfectly competitive are induced to collude as a precautionary measure when they face the risk of being unjustly sanctioned when obeying the law. This implies that imperfect enforcement may be counter-productive in that it stimulates the very behavior it was designed to prevent. Note that below it is argued that firms' inability to determine with certainty which forms of behaviour that will be assessed under the effects-based standard (as opposed to *per se* analysis) will be held in violation and which forms will be excused, may also produce another perverse effect, which is that firms may end up being unnecessarily cautious in signing agreements and, thus, forego legitimate profits.

allow active and passive sales into this territory by resellers appointed in another area, etc. Some of these alternatives will be legal, others will be caught.

Like in the binary situation, the firm will simply choose that contract to which it assigns the highest net expected benefit (that is, the largest combination of expected profits and fines). Put differently, whether at any starting point the potential offender will find it more profitable to decrease or increase the restrictiveness of his agreement, depends on the pace at which expected profits and sanctions change in moving from the one option to the other. If, at a certain point, the expected sanction costs increase faster, as behaviour becomes more restrictive, than the anticipated profits do, a potential offender will choose the less restrictive course of action.⁶⁵

Legal certainty, in this context, implies that the firm is able to distinguish, without doubts, the forms of the agreement that will be considered illegal and those that are in the clear. Assuming that the expected cost of fines outweighs the benefits of violating, the firm will then choose the most restrictive option that is still legal. Uncertainty, on the other hand, can be thought of as follows. The firm considers that as they become more restrictive, the four options are increasingly likely to be challenged and terminated. But it does not exclude the possibility that if it implements the least restrictive form it will be held in violation, nor does it completely discount the possibility that the most restrictive form will be cleared.⁶⁶ It will prove helpful in understanding the arguments that follow, to express this in the form of a graphical image.

⁶⁵ See Calfee and Craswell (1984). This means that absolute levels of precaution and liability costs at a certain point are not determinative. Even if in absolute terms expected fines are lower than the costs of precaution at a certain point, the potential offender would still choose to take even more precaution if this would have more effect in terms of minimizing the sum of precaution and liability costs.

⁶⁶ Both strands of the literature that are discussed in this section deal with *ex ante* uncertainty on the part of firms in determining how the antitrust authority will assess different forms of conduct that they may engage in. In each context, however, the nature of uncertainty is different. In the type of binary setting such as described by Schinkel and Tuinstra (2006), the antitrust authority is less well informed about the true nature of firm conduct (in their case, about cost structures) than firms are. *Ex ante*, the question will be whether the authority will take the right decision in case the firm engages in either of the two available options. The situation in the effects-based field, where firms face multiple options, is different. Here, even firms can be expected to dispose of imperfect information, since the welfare effects of these types of agreements are more difficult to assess. *Ex ante*, firms will therefore not be fully certain about the precise impact of their agreement on the market and, crucially, on how this impact will be assessed *ex post* by the antitrust authority.

Figure 3.1

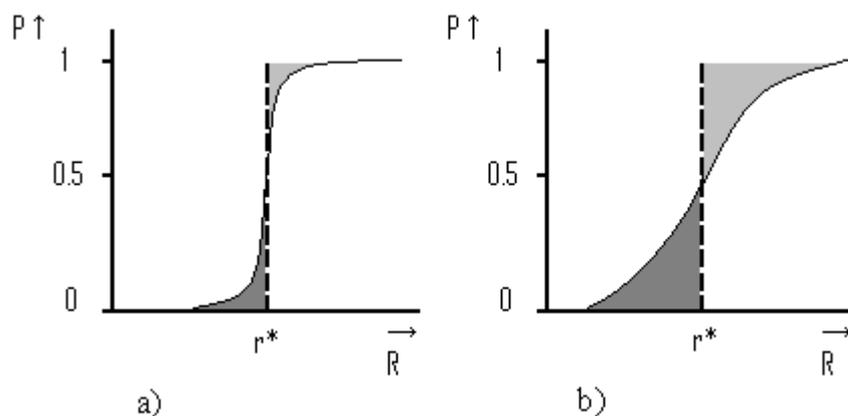


Figure 3.1 displays two possible distributions of probabilities, both of which map the combination of a firm's best guesses of the probability of being found to have violated the law that it associates with each of the possible forms of actions that it can choose from. The horizontal axis measures the firm's options in terms of behaviour. These represent the various forms of agreements that it can conclude. As we move from left to right on this axis, the behaviour becomes more restrictive and, eventually, harmful to consumers. The vertical axis measures the probability of violating the law. Thus, in each of the three graphs the curved line shows the probability violating the law that the potential offender in question associates with each possible form of restriction.

Consider first the vertical dotted line starting at point r^* on the horizontal axis. This is the point at which the potential offender estimates that a potential suit challenging his behaviour would be equally likely to go either way. Claims regarding behaviour to the left are more likely to be rejected, and cases involving conduct that are situated to the right are more likely to be granted. This suggests that r^* represents the legal standard as perceived by potential offenders. The second feature of the graphs depicted in Figure 3.1 that is of interest here is the way uncertainty is spread around the perceived standard. The spread can be taken as a measure of the uncertainty that potential offenders face. Graphs a) and b) can be used to illustrate this. Potential offenders that work with a distribution such as depicted in graph a) face considerable uncertainty only as they approach the standard quite closely (from both directions). These firms can be said to enjoy a high degree of legal certainty. Potential

offenders that reckon that probabilities are distributed as in graph b) remain uncertain about the appraisal of behaviour that is much less (and much more) restrictive.

Simple numerical examples suffice to show that – in this context – uncertainty may lead to both under-deterrence and over-deterrence. Assume that our firm faces a choice between four different forms of distribution agreement (E, F, G, and H), that these forms are increasingly restrictive of the competitive process, that they present mounting profit opportunities and, thus, increasing sanction costs.⁶⁷ Assume, also, that the standard as perceived by the firm corresponds to the actual legal standard as would eventually be applied by the Commission should formal proceedings be initiated,⁶⁸ such that the dividing line (r^* in graphs a) and b) of Figure 3.1) runs between F (legal) and G (illegal). Now, consider the situations where our firm makes the estimations represented in Tables 3.1 and 3.2 (p represents the probability of being successfully challenged by the Commission; B stands for the profits generated by the agreement; C indicates the fine that will be imposed in case of termination; and EV represents the expected value of each option). Both situations reflect a high degree of legal certainty: options E and F are considered very unlikely to be successfully challenged, whilst the opposite is true for options G and H. Costs and benefits are differently distributed, however. In Table 3.1 expected profits rise faster than the expected sanction and the reverse is true for Table 3.2. In both situations, comparing the net expected benefits will lead the firm to choose the most restrictive legal option F.

Table 3.1

Options	P	B	C	EV
E	0	50	50	50
F	0.05	90	70	82
G	0.95	140	90	<0
H	1	200	110	<0

⁶⁷ Fines are calculated on the basis of a proportion of the turnover in the affected market, adjusted for the seriousness of the infringement (see the Commission's guidelines on the method of setting fines pursuant to Article 23(2)(a) of Regulation 1/2003 ([2006] OJ C210/2)). As options become more restrictive (serious) and profitable (higher profits mean a larger turnover) sanction costs should therefore also rise. As is reflected in the tables presented below, however, firms' perceptions of the pace at which sanction costs rise may differ as they depend on (1) how much the extra profits increase overall turnover and (2) the expectations about the Commission's appreciation of the seriousness of the restriction (repeat offenders may, for example, be punished more severely for the same act). In this regard, see also the discussion on fining issues in Chapter 6, Section 6.2.

⁶⁸ This means that the negative effects of the restriction start outweighing the benefits exactly at the point where the potential offender considers it most likely that his conduct would be condemned if challenged. This is the case in graphs a) and b) of figure 3.1.

Table 3.2

Options	P	B	C	EV
E	0	100	20	100
F	0.05	120	40	112
G	0.95	140	80	<0
H	1	160	160	<0

It is easily appreciated, by considering Tables 3.3 and 3.4, that when legal certainty deteriorates inefficient results may be obtained. The firm still considers that options E and F are more likely than not to be cleared and that alternatives G and H are more likely to be held in violation. But the differences between the various options in terms of the probability of being successfully challenged by the Commission have grown much smaller. As a consequence, in the situation reflected in Table 3.3, the expected benefits of making the agreement more restrictive increasingly outweigh the expected sanctions, leading the firm to choose the option that poses the greatest threat to consumers (under-deterrence). The opposite result is achieved in the situation captured by Table 3.4. There, it is the rate of change in expected sanctions which outweighs the benefits and induces the firm to opt for a less restrictive alternative than is necessary to avoid consumer harm, which means that it foregoes legitimate private benefits.

Table 3.3

Options	P	B	C	EV
E	0.43	50	50	7
F	0.48	90	70	13.2
G	0.52	140	90	20.4
H	0.57	200	110	23.3

Table 3.4

Options	P	B	C	EV
E	0.43	100	20	48.4
F	0.48	120	40	43.2
G	0.52	140	80	25.6
H	0.57	160	160	<0

More generally, it can be said that it is the spread of a potential offender's probability distribution that determines the seriousness of such inefficiencies, since it influences the pace at which sanction costs are perceived to rise. We can appreciate this by comparing the situations depicted in graphs a) and b) of Figure 3.1. Because in

the latter situation the probability of being found to violate the law starts rising at an earlier point when moving from left to right on the horizontal axis, expected sanction costs initially increase faster than is the case when uncertainty is concentrated in a narrow range. Conversely, in the area immediately surrounding the standard, these costs change at a much slower pace (as in the numerical example). And in the region farther to the right, the effect of uncertainty is to dilute the threat of sanctions. As we said, the ultimate effect on behaviour also depends on the rate of change in foregone profits. Nevertheless, we are able to see that the risk of potential offenders taking too much or too little precaution becomes larger as uncertainty is more spread out.

Since the models developed in the two strands of the literature discussed above point in different directions – Png (1986) and Polinsky and Shavell (1989) find that uncertainty will result in under-deterrence, whilst Calfee and Craswell (1984) and Goetz (1984) find that it may lead to both under-deterrence and over-deterrence – it is important to examine which circumstances prevail in European antitrust. To this end, the next section takes a closer look at circumstances under which firms make decisions that may lead to adverse effects on European consumers.

3.4.2 Contracting in the effects-based field

As we saw, what is crucial here is the number of alternative actions that a firm can take. But note that it is not sufficient that the firm can choose between more than two types of behaviour to determine that the model advanced in Calfee and Craswell (1984) and Goetz (1984) applies. To use an example involving a *per se* violation, cartel members may fix the price of their products at many different levels. But they will surely know that whether they opt for a 5%, a 10% or a 15% mark-up makes no difference whatsoever for the illegality of their conduct. In such cases the real choice is binary – whether to form a price fixing cartel or not – which means that the framework drawn up by Schinkel and Tuinstra (2006), Png (1986) and Polinsky and Shavell (1989) is more suitable. It is only when a firm is able to choose between more than two options that differ in terms of the probability that they will be held illegal that over-deterrence becomes a legitimate concern.

A survey of the Commission's decisions listed in Appendix A suggests that most involve conduct that will have been chosen from a range of options. It was suggested above already that this may be the case when setting up a network. This can be

illustrated by reference to the facts related in the Commission's decision in the case of *Bass*⁶⁹ (and the nearly identical cases of *Whitbread*,⁷⁰ and *Scottish and Newcastle*⁷¹). UK breweries prevented their tenanted pub-operators from obtaining supplies from competitors, as well as from buying their own brands from wholesalers. The Commission exempted these agreements, but it is clear from reading these decisions that this was only because they applied to a quickly decreasing number of retailers. At the time this form of distribution agreement was implemented – on a much larger scale – antitrust intervention must have been a possibility taken seriously. Lowering the level of restrictiveness by allowing intra-brand competition through the wholesale channel or by allowing tenants to buy from whichever supplier they preferred would each have reduced this probability by a certain margin.

The Commission's more recent commitment decision in the case of *Repsol* points in the same direction.⁷² This firm operated a complex system of agreements to regulate the distribution of its fuel products by petrol stations in Spain. Depending on the situation of the individual reseller, these contracts included different stipulations regarding (1) the way the distributor was tied to Repsol (by means of proprietary and or contractual rights), (2) the duration of the distributorship, and (3) the reseller's freedom to determine the prices charged to consumers. The Commission's initial assessment led it to oppose several more restrictive forms of these agreements out of fear that they foreclosed the market. Clearly, raising entry barriers is a gradual affair. As in the case of *Bass*,⁷³ different variations of the restrictive components in the agreements could have been used to produce results that were either more or less likely to provoke intervention (for example, by allowing more inter-brand competition or by reducing the duration of the agreements).

The exemption decision in *UEFA*⁷⁴ (and the very similar commitment decisions in the cases of *Premier League*,⁷⁵ and *DFB*⁷⁶) can be used to illustrate that the same may apply to horizontal schemes. UEFA managed a joint sales operation of broadcast rights by football clubs participating in the UEFA Champions League. After

⁶⁹ Supra, footnote 36.

⁷⁰ [1999] OJ 88/26.

⁷¹ [1999] L186/28.

⁷² Decision of 14 April 2006, published on the Commission's website.

⁷³ Supra, footnote 36.

⁷⁴ [2003] OJ L291/25.

⁷⁵ Decision of 22 March of 2006, published on the Commission's website.

⁷⁶ Decision of 19 January of 2005, published on the Commission's website.

negotiations between the Commission and the organisations involved, the exclusivity claimed by joint venture was trimmed down so as to ensure that individual clubs could market their rights independently in case the joint venture received no reasonable offer. But it is, again, easy to imagine alternative ways to market these rights that – *ex ante* – would have been more or less likely to be held illegal.⁷⁷ More generally, the Commission’s conditional exemptions and commitment decisions frequently suggest that negotiations with the parties – which are often spread out over different stages of the proceedings – involve bargaining over a range of less restrictive alternatives, rather than two unique possibilities.⁷⁸

Cases that involve strictly exclusionary behaviour, such as *Van den Bergh Foods*⁷⁹ and *Morgan Stanley / Visa*,⁸⁰ are the best candidates for the binary model, where uncertainty can only lead to under-deterrence.⁸¹ The first case involved Ireland’s principal ice cream manufacturer Van den Bergh, which had denied its competitor Mars access to freezer cabinets installed in outlets with limited floor space. The latter case revolved around Visa’s refusal to allow Morgan Stanley, a competitor in a different geographical market, to become a member of its credit card network in the UK. At first sight it would appear that both firms had only two options: to accept or to refuse. We must realise, however, that such requests – and the question whether to engage in anticompetitive conduct in general – will seldom present itself without prior notice as these two examples could be taken to suggest. Mostly, the conduct at issue in effects-based analysis will be the result of a sequence of decisions taken over a longer period of time. All examples given above – including the latter two – involve large framework agreements which the firms and organisations in question had put in place to regulate the many different aspects of their complex relations with competitors or retailers.

⁷⁷ A similar argument can be developed for the following horizontal cases included in Appendix A: *P&O / Stena Line* ([1999] OJ L163/61), *TPS* ([1999] OJ L90/6), *P&I Clubs* ([1999] OJ L125/12), *Cegetel* ([1999] OJ L218/14), *EATA* (supra, footnote 33), the exploitative aspects (rebates) of the *Cannes* agreement (decision of 4 October 2006, published on the Commission’s website), and *Master Card* (supra, footnote 33).

⁷⁸ In this regard the decisions listed in Appendix A where during the proceedings the defendants make amendments to their agreement and the Commission later makes clearance conditional upon further modifications are interesting to look at: *IMA (EPI)* ([1999] OJ L106/14), *GEAE / P&W* ([2000] OJ L58/16), *Reims II* ([1999] OJ L275/17), *BSkyB* ([1999] OJ L312/1), *Visa MIF* ([2002] OJ L318/17), *Lufthansa / SAS / United* ([2002] C264/5), *Reims II Renewal* ([2004] OJ L56/76), *Argev / Ara* ([2002] C250/3), and *Alitalia / Air France* (decision of 7 April 2004, published on the Commission’s website).

⁷⁹ Supra, footnote 33.

⁸⁰ Id.

⁸¹ See the discussion above on the works of Schinkel and Tuinstra (2006), Png (1986) and Polinsky and Shavell (1989).

Now, even if at the final stage the choice is binary, at the moment the agreement is drawn up more options may be available. The case of *Van den Bergh Foods* is a prime example.⁸² Van den Bergh provided its network of distributors with free-of-charge freezer cabinets on the condition that they were used to stock the supplier's brand only. Given that most shops selling ice cream had little floor space, in practice almost all vendors had but one cabinet and therefore many of the agreements were in fact exclusive purchasing agreements. At the time these distribution contracts were developed, the company could also have imposed no restriction at all. Alternatively, it could have stipulated a minimum percentage of Van den Bergh ice cream stocked in the cabinet. And at the other end of its range of options, Van den Bergh could have flat-out prohibited selling any other brands in the entire shop.

It is thus of crucial importance that we focus on that stage of the decision making process at which the firm realises that the action it contemplates might fall foul of the antitrust laws. Arguably, in the effects-based field this moment will mostly come early, at the stage where the agreement is given form.⁸³ Remember that firms whose agreements are subjected to individual scrutiny will often be large undertakings (unable to benefit from a block exemption regulation) or organisations whose market wide coordination is necessary for the existence or proper functioning of the market itself.⁸⁴ Such firms and organisations can be expected to be acutely aware of the European antitrust laws and to obtain the necessary advice on the risks they run in implementing these large scale contractual operations. As suggested, at this early stage more than two options will generally be open. Cases such as *Morgan Stanley /*

⁸² *Id.*

⁸³ Recent US case law offers an interesting example of a case where this point may have come much later: *Pepsi v. Coca Cola* (315 F.3d 101 (2nd Cir. 2002)). This case involved the distribution of fountain syrup, which is thickened cola that is mixed with carbonated water at the point of sale. Early on in the 20th century, these two competitors chose for different distribution channels. Pepsi relied on independent bottlers with whom it shared in the ownership of fountain syrup rights by means of licensing agreements. Coca Cola retained its rights and chose to work with exclusive distributors instead. For a variety of reasons, this choice eventually allowed Coca Cola to manage this channel more effectively, resulting in a larger share of this business. Towards the end of the 1990s, Pepsi decided to compete more aggressively in the fountain syrup channel. As a consequence, it shifted its attention away from bottlers towards distributors, many of whom had contracts with Coca Cola. At this point, the exclusivity clauses that had been in Coca Cola's contracts for almost a century suddenly became a serious problem. It should be realised, however, that the circumstances of this case are rather exceptional. Ordinarily, a much shorter time span will be involved and market conditions and antitrust enforcement will therefore not change as much in the period between the implementation of the agreements and the anticompetitive conduct.

⁸⁴ Such as *IMA* (supra, footnote 78), *Reims II* (supra, footnote 78), *CEDED* ([2000] OJ L187/47), *Eurovision* ([2000] OJ L151/18), *IFPI Simulcasting* ([2003] OJ L107/58), and *Argev / ARA* ([2002] OJ C250/3).

Visa,⁸⁵ where the choice is binary from the start, appear to be exceptional in the effects-based field.⁸⁶ This means that we must assume that uncertainty about the effects-standard may result in both under-deterrence and over-compliance.⁸⁷

3.5 Conclusions and implications

It was argued at the beginning of this chapter that a rigorous evaluation of the Commission's effect-based decision practice requires a clear view on the role of the legal standard in the enforcement system as a whole, as well as on the optimal calibration of the legal standard. Such an objective benchmark is absent in earlier assessments of the European *rule of reason*.

We have seen that the legal standard functions at two levels. It regulates the investigation of restraints in the course of legal proceedings by signalling to parties in litigation what arguments and associated categories of evidence must be presented in order for their claims to succeed, and it serves to inform firms as to whether their intended agreement is of the type that may suffer the consequences of antitrust enforcement. The calibration of the legal standard can be said to be optimal if it incorporates policymakers' appraisal of modern economic insights in such a way as to minimise the combination of investigation costs incurred in proceedings and error costs produced by firms deviating from the legal standard. Accordingly, the following chapters will set out to characterise European antitrust in terms of (1) policymakers' priors regarding the frequency with which a certain practice will lead to harm, (2) their expectations as to the costs of producing the related evidence, and (3) their assessment of the consequences of non-compliance.

We have also seen that compliance problems may ensue if the required level of precision in showing harm is low, leaving potential offenders uncertain about the legality of their intended agreements. With uncertainty, firms planning to sign a distribution or joint venture agreement may either take unnecessary precautions, or

⁸⁵ *Supra*, footnote 33.

⁸⁶ The behaviour at issue in the case of *ECO System/Peugeot* ([1992] OJ L66/1) might be of a similar nature. Note that this decision is not included in Appendix A because it was adopted well before the reference period.

⁸⁷ Note that this conclusion does not call into question the adequacy of the contributions by the authors mentioned in footnote 81 for describing the situation in the *per se* field. Note, also, that the conclusion drawn here corresponds to the approach taken by Cass and Hylton (2001) in their analysis of US antitrust. Without elaborating on their reasons for doing so, they apply a model involving uncertainty that allows for under and over-deterrence, along the lines of Craswell and Calfee (1986).

cross the line instead. This was argued to be different from the situation in the field of *per se* analysis, where firms will typically face a choice (whether to participate in a cartel or not) that does not allow for the third option of taking too much precaution. This, too, is important for the analysis engaged in below, since over-compliance must be addressed by different means than under-deterrence.