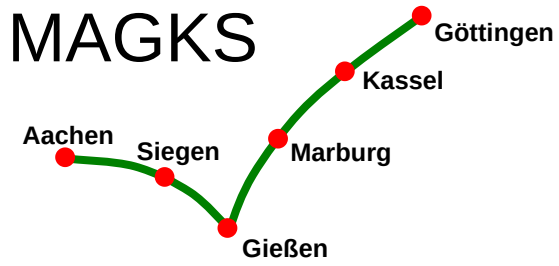
View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics



Joint Discussion Paper Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 16-2008

Wolfgang Kerber, Jürgen-Peter Kretschmer, and Georg von Wangenheim

Optimal Sequential Investigation Rules in Competition Law

This paper can be downloaded from
http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/index_html%28magks%29

Coordination: Bernd Hayo • Philipps-University Marburg
Faculty of Business Administration and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Preliminary draft: 10 September 2008

Optimal Sequential Investigation Rules in Competition Law

Wolfgang Kerber, Jürgen-Peter Kretschmer, and Georg von Wangenheim*

Abstract: Although both in US antitrust and European competition law there is a clear evolution to a much broader application of "rule of reason" (instead of per-se rules), there is also an increasing awareness of the problems of a case-by-case approach. The "error costs approach" (minimizing the sum of welfare costs of decision errors and administrative costs) allows not only to decide between these two extremes, but also to design optimally differentiated rules (with an optimal depth of investigation) as intermediate solutions between simple per-se rules and a full-scale rule of reason. In this paper we present a decision-theoretic model that can be used as an instrument for deriving optimal rules for a sequential investigation process in competition law. Such a sequential investigation can be interpreted as a step-by-step sorting process into ever smaller subclasses of cases that help to discriminate better between pro- and anticompetitive cases. We analyze both the problem of optimal stopping of the investigation and optimal sequencing of the assessment criteria in an investigation. To illustrate, we show how a more differentiated rule on resale price maintenance could be derived after the rejection of its per-se prohibition by the US Supreme Court in the "Leegin" case 2007.

JEL-Classification: K20, K21, K40, D81

Keywords: Law Enforcement, Decision-Making, Competition Law, Antitrust Law

* Wolfgang Kerber and Jürgen-Peter Kretschmer, Department of Business Administration and Economics, Philipps-Universität Marburg, Am Plan 2, D-35032 Marburg, Germany, E-mail: kerber@wiwi.uni-marburg.de, kretschmer@wiwi.uni-marburg.de; Georg von Wangenheim, Department of Economics, Universität Kassel, Nora-Platiel-Straße 2, D-34109 Kassel, Germany, E-mail: g.wangenheim@uni-kassel.de. We thank all participants of the 30th HOS Conference (Nürnberg, April 2008) and the 12th Annual Conference of the ISNIE (Toronto, June 2008) for a valuable and helpful discussion of a previous version of this paper.

1. Optimally differentiated rules in competition law and sequential investigation rules

Both in US antitrust law as well as in European competition law there has been a long-term development to replace per-se prohibitions by a broader application of rule of reason. In the US, the famous GTE Sylvania judgment of the Supreme Court (1977) marked the beginning of a step-by-step process of reversing former per-se prohibitions of vertical (and partly horizontal) restraints. The last important step was the Leegin decision of the Supreme Court (2007), which abolished a nearly 100 years old per-se prohibition of resale price maintenance (RPM).[1] A similar development towards more rule of reason can also be observed in European competition law, albeit much later, much slower and more cautious than in US antitrust law. To a large extent this development is driven by modern industrial economics, which can demonstrate that many business practices and agreements do not have always positive or always negative effects on competition or welfare; their pro- or anticompetitive effects may rather depend on a number of specific circumstances. The ensuing recommendation of a more differentiated and deeper case analysis suggests a broader use of the more flexible framework of rule of reason as a logical consequence of the new economic insights in industrial economics (Hovenkamp 2001: 268). Therefore the "more economic approach" in the EU with its focus on the relevance of economic effects has been the main driving-force towards a more rule of reason-like application of European competition law (Röller 2005; Carlton/Salinger 2007: 293). However, this development has also been seriously criticized. One important group of counter-arguments refer to the problems of increasing legal uncertainty and higher administrative and legal costs, especially if this development implies a tendency to a case-by-case analysis of the pro- and anticompetitive effects (Voigt/Schmidt 2005; Roth 2007).

An additional problem is that the decision to replace a per-se prohibition of a business behavior through a rule of reason does not clarify how the new more differentiated treatment should look like. For example, the consequence of the Leegin decision is that the competition authorities, the courts and the firms only know that RPM is no more per-se prohibited; however, a long process of court decisions can be expected to clarify the conditions and circumstances, under which RPM might be allowed or remain prohibited. In contrast to what most economists think on first sight, rule of reason does not imply a full-scale market analysis of all pro- and anticompetitive effects in each particular case. Both in US antitrust law and EU competition law practice, a number of intermediate solutions between simple per-se rules and a full-scale market analysis have been developed (Tom/Pak 2000). "Safe-harbor rules", "quick-look rules", "structured rule of reason" but also guidelines (with a list of assessment criteria) represent widespread practical solutions for the dilemma that full-scale analyses in each case are too costly (or even impossible) but that simple per-se rules might lead to too many decision errors by not being capable to discriminate sufficiently between pro- and anticompetitive behaviors. Theoretically, these intermediate solutions can be characterized as defining and structuring the application of a predefined and limited set of assessment criteria. Therefore simple per-se rules and full-scale case analyses of all

[1] *Leegin Creative Leather Products, Inc. v. PSKS, Inc.*, 551 U.S. ___ (2007), *Sup Ct* (USA), No. 06-480 (Leegin decision); *Continental T.V., Inc. v. GTE Sylvania, Inc.*, 433 U.S. 36 (1977); *National Collegiate Athletic Association v. Board of Regents of the University of Oaklohoma* 468 U.S. 85 (1984).

effects are only two extreme solutions of a continuum of a smaller or larger depth of case analysis with predefined differentiated rules for assessing the pro- and anticompetitive effects of certain kinds of business behavior. As a consequence, the transition towards more rule of reason does not imply the rejection of a rule-oriented application of competition law, rather it claims the need for a larger differentiation of these rules.

The relevant question is less about "per-se rules or rule of reason" but how many and what criteria should be used for the legal assessment of a business behavior. From that perspective it can be asked for the optimal degree of differentiation of competition law rules (Christiansen / Kerber 2006), i.e. for the optimal depth of case analysis for certain classes of business behaviors. If we define the degree of differentiation of rules as the number of assessment criteria and the extent of information that is used for an investigation about the legality of a behavior, then the notion of the differentiation of rules corresponds to the approach of Kaplow (1992, 1995) in regard to the optimal complexity (or precision) of law. We also follow Kaplow in his insistence of distinguishing strictly between the dimension of the complexity of law as the number of applied criteria and the dimension of an ex ante or ex post determination of these criteria (as the specific content of the law). The latter is at the heart of the well-known "rules vs. standards" debate.[2] In the following, we only will focus on the dimension of the optimal differentiation (complexity) of rules and will not discuss the question of ex ante vs. ex post determination of these rules and their evolution over time.

In the literature on competition law, the "error cost approach" is used for analyzing whether more general or specific rules (or: per-se rules or rule of reason) are appropriate (Easterbrook 1992; Beckner/Salop 1999; Tom/Pak 2000; Joskow 2002; Heyer 2005). Since competition authorities and courts can fail to distinguish properly between pro- and anticompetitive behaviors leading to the false prohibition of procompetitive behavior (overinclusiveness; type I error / false positives) and false acquittal of anticompetitive behavior (underinclusiveness; type II error / false negatives), welfare-decreasing decision errors will occur. A deeper investigation, which uses more criteria and information, might be able to reduce the welfare costs through decision errors by better discriminating between pro- and anticompetitive behavior. However, it can be expected that the direct and indirect legal and administrative costs for firms, competition authorities and courts will increase with the application of a more differentiated rule. Based upon the error costs approach, which intends to minimize the sum of error costs and regulation costs, Christiansen/Kerber (2006: 225-229) present a simple general model how the optimal degree of differentiation of rules can be derived: A rule is optimally differentiated, if the marginal reduction of welfare losses through more differentiation is equal to the marginal costs of more differentiation. Their analysis shows that the optimal differentiation depends on the distribution of the welfare effects of the class of business behavior, on the information value of the assessment criteria, and on the direct and indirect regulation costs (ibid.: 229-235). It should be noted that this problem of the optimal differentiation of rules (or the optimal depth of an investigation) is not specific to competition law but a general problem in many fields of the law.

[2] See Ehrlich/Posner 1974, Kaplow 1992, and Crane 2007; see also Kerber 2008.

The consideration of the investigation costs and the welfare losses through decision errors leads to an entirely new economic perspective on the discussion about the appropriate competition law rules. We can use the discussion on resale price maintenance as an example. In every textbook on industrial economics and competition policy, resale price maintenance is discussed by analyzing theoretically its procompetitive and anticompetitive effects (Motta 2004). Since it can be demonstrated that under certain conditions RPM can help to remedy free-riding problems or other inefficiencies and thus has welfare-enhancing effects, most economists and many legal scholars conclude that a per-se prohibition is definitely wrong, and recommend a rule of reason (Mathewson/Winter 1998: 82; Motta 2004: 377; Van den Bergh/Camesasca 2006: 234; Kneepkens 2007: 664). However, such a conclusion is premature from the error costs approach. From this perspective the decisive question is whether the transition from the simple per-se prohibition to a more differentiated rule about RPM is welfare-enhancing. This requires to answer the question whether such a new rule would reduce the error costs of prohibiting a welfare-enhancing RPM sufficiently to overcompensate the sum of additional costs of the new rule through the additional costs of differentiation (including the indirect costs through legal uncertainty) and the additional welfare losses through new errors by erroneously allowing RPM with anticompetitive effects (Christiansen/Kerber 2005: 228). This requires, e.g., knowledge about the probability that RPM is procompetitive and about the size of the welfare losses, if these decision errors occur. The traditional industrial-economics discussion about resale price maintenance has not taken this problem into account. As a consequence, the search for optimal competition law rules requires not only the insights of the traditional industrial-economics approach but also research on the optimal depth of investigation from a law and economics perspective.

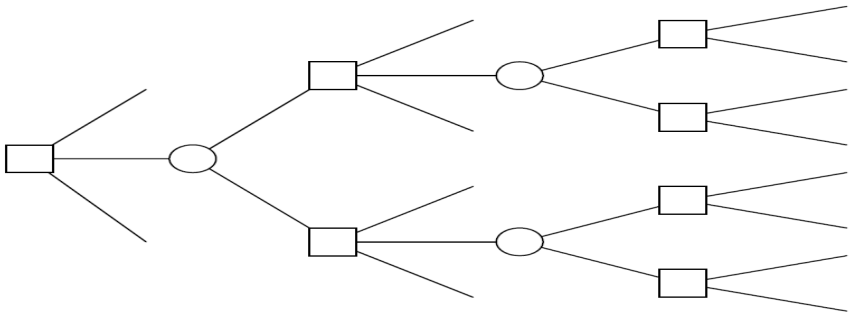
In this paper we want to analyze an additional aspect of optimally differentiated (or complex) rules. One can easily argue that it is not optimal to apply the same number of assessment criteria for all cases of a class of business behavior. For example, safe-harbor rules can exempt a business behavior from being prohibited as cartel on the only provision that the market shares of the firms are low, and only in the case of higher market shares a deeper investigation is carried out. Therefore market shares can be interpreted as a signal that allows us to sort cases of a class of behavior into different subclasses - into an unproblematic one, which only rarely entails welfare-decreasing cases, and into a subclass of more problematic cases, in which a deeper investigation is recommended for a better distinction of pro- and anticompetitive behavior.[3] From that perspective an investigation process can be interpreted as a step-by-step (or multi-stage) *sorting process*, in which a sequence of criteria is used for sorting the cases in (ever smaller) subclasses, which allow for better decisions. Safe-harbor rules can be welfare-increasing, because many unproblematic behaviors can be exempted with the use of only one low-cost criterion (market shares), and the bulk of the enforcement resources can be used for the smaller number of potentially more problematic cases with higher market shares. This additional option of a sequential investigation process transforms the question for the optimal number of

[3] Another wellknown example is the use of the Herfindahl-Hirschman-Index (HHI) in horizontal merger cases as a first criterion for sorting merger cases in different subclasses, in which different investigation procedures are being applied.

criteria into the more complex question for the optimal sequential structure of investigation. This encompasses two main questions: (1) the optimal sequence of assessment criteria and (2) the optimal stopping of the sequential investigation process. As the example of safe-harbor rules demonstrates, the optimal depth of investigation might depend on the information we get from early assessment criteria.

So far only Beckner/Salop (1999) have introduced explicitly such a decision-theoretic perspective for analyzing antitrust investigation procedures and contributing to more rational decision procedures for competition authorities and courts (but also firms with their problems of compliance). After presenting the basics of decision theory, Beckner/Salop develop the structure of a multi-stage decision process, in which step by step decisions are taken, either for gathering additional information or for deciding the case at any stage of the process. This leads to a multi-branched decision tree (as in the following figure 1). Using also the aforementioned error cost approach Beckner/Salop discuss the questions of rule of reason, per-se rules, standard of proof, and truncated rule of reason standards from this perspective. One conclusion of their analysis is that the courts should not decide ex ante on the optimal amount of information but make the demand for additional information dependent on the outcome of the gathering of information on the first criteria; the latter should be the "least expensive, resolves the most uncertainty, and is most likely to affect its decision" (Beckner/Salop 1999: 54). They also conclude that a rule of reason should not include gathering the complete information simultaneously on both efficiency benefits and market power harms; rather the difficult and costly assessments of efficiencies should be located at the end of such an investigation process (with the consequence that they often might not be necessary any more) (ibid.: 62). Although their article has seminal quality by explaining the decision-theoretic approach and its relevance for antitrust law, Beckner/Salop do not present an analytical model of this approach or derive their specific conclusions from a rigorous theoretical analysis. In our paper we present first steps of such an analytical instrument for making complex legal investigations more rational. In particular, we model the structure of optimal sequential investigation procedures based upon such a decision theoretic approach.

Figure 1: Stylized multi-stage investigation process



= decision about permission / prohibition/ searching for additional information

= information (signal) from an investigation criterion

In our model we ask for optimal sequential investigation rules that maximize welfare.[4] We use the error costs approach that minimizes the sum of welfare losses through decision errors (by false positives and false negatives) and (all direct and indirect) regulation costs through the application of competition law. It is assumed that industrial economics and experience provides us with information about the probabilities that business behaviors or agreements or market structures have positive or negative welfare effects, i.e. that they are pro- or anticompetitive ("priors": Beckner & Salop 1999: 58; Heyer 2005), and with assessment criteria whose investigation helps to distinguish both.

In our basic model we assume not only a given set of those criteria but also a given sequence for investigating them step-by-step in this multi-stage process. Contrary to Beckner/Salop, who focus primarily on the specific US antitrust litigation process (with plaintiffs and defendants and burden of proof problems on each stage), we assume a simpler unilateral investigation process, as, e.g., by a competition authority (CA) in a public enforcement regime (as in the EU). The first stage consists of the decision of the CA, whether it allows or prohibits the behavior (the upper and lower branch) directly or opts for getting more information by investigating an additional criterion, as, e.g., the market share of the firm. This information can be interpreted as a signal that sorts the cases in two subclasses. For example, we know that for many behaviors the share of cases with anticompetitive effects is much lower, if there are low market shares, than in the subclass of the same behaviors of firms with large market shares. Therefore the signal allows us to sort the case into one of these two subclasses, with different distributions of anti- and procompetitive effects of this behavior. At the second stage, the same question arises again, whether now the behavior should be prohibited or allowed or whether the investigation should be continued. At each stage this decision is made by calculating whether the information value of an additional assessment criterion (measured by reducing the sum of error costs through a better discrimination of anti- and procompetitive behavior) is higher than the additional regulation costs by deepening the investigation.

What kind of information do we use in this procedure? In the industrial-economics discussion it is analyzed theoretically and empirically, whether certain behaviors lead to pro- or anticompetitive effects under certain conditions, e.g., if the market shares are high or low, or if high or low entry barriers exist. These additional conditions can be used as the assessment criteria in the investigation that help us to sort the cases into the different subclasses, which are represented by the different branches of the decision tree. We assume that we have information about the conditional distributions of anti- and procompetitive effects of the business behavior in all the subclasses of cases, i.e. under the conditions of a combination of signals.[5] In regard to

[4] Instead of maximizing welfare we could also use any other well-defined and consistent objective function for competition law, which also might consist of a bundle of goals. Therefore we need not discuss the current issue of consumer welfare standard vs. total welfare standard. However, it is important that we ask for the optimal investigation rules from a social welfare perspective. The problem would be different, if we would ask for optimal sequential investigation rules for a competition authority (or a firm) with an exogenously given budget.

[5] We do not assume to have information about the probability of getting a certain signal through assessing a criterion (as high or low market shares), if the behaviour has pro- or anticompetitive effects. Therefore

RPM, this would imply that we have information about the probability that a RPM has positive welfare effects, e.g., under the condition that the market shares of the manufacturer are smaller than 10%, and that RPM is not widely spread in this industry.[6] We also assume knowledge about the size of welfare losses in the case of decision errors and the costs of the investigation of additional assessment criteria. These information requirements are critical, and we will discuss this problem in section 4.

Our main question in this article about optimal sequential investigation is where we should stop at each branch within such a multi-stage and multi-branched decision tree. Mathematically this is solved by backward induction, because the information value of the first criteria also depends on the potential of reducing error costs by criteria that might be applied later. We emphasize that the result can be interpreted as a *rule* (for a class of behavior or agreements) that determines a sequence of assessment criteria and how far these criteria should be investigated in each of the branches of this multi-stage decision tree. If a competition authority has to investigate a case, then it only has to apply the optimal sequential investigation rule for this class of cases. In that respect, our concept of an optimal sequential investigation rule can be seen as a sophistication of the concept of optimally differentiated (or complex) rules. Per-se rules, full-scale rule of reason, and other intermediate solutions as safe-harbor rules and structured rule of reason can be interpreted as special cases of this kind of (optimal sequential investigation) rules, and their expediency can be assessed by using such an analytical approach. Our intention is to present a formal model that allows the analysis of the problem of optimal sequential investigation processes, and that can be viewed as an additional step for the development of a general analytical instrument for deriving optimal sequential investigation rules.

Our paper is structured as follows: In the next section 2, we present our basic model which focuses on the problem of optimal stopping within such a multi-stage decision process (optimal depth of investigation). We use the example of resale price maintenance to illustrate how such an approach can be applied for deriving optimal investigation rules. Whereas in section 2 we assume the existence of a given sequence of assessment criteria, in section 3 we address the question of the optimal sequencing of a set of criteria. We show that this is an extremely complex problem that has not been solved generally in decision theory. However, we show that under certain assumptions fairly good heuristic solutions are possible. Section 4 entails a brief critical discussion of our approach and additional extensions and applications.

2. Optimal Sequential Investigation: The Basic Model

2.1 Definitions and assumptions

The starting point of our model is a Competition Authority (CA), which has to decide on

---

we have a different information input as in the usual case of Bayesian updating (see the next section 2.1.).

[6] We will discuss this example in some more detail in section 2.3.

allowing or prohibiting a wide array of business behaviors brought before it. The cases under investigation are to be assessed according to their welfare effects: in a perfect world without lack of information, welfare enhancing behaviors should be allowed; welfare reducing behaviors should be prohibited. The problem of the CA is, however, to decide in an imperfect world, where gathering information is costly and there exists a trade-off between perfection of the Authority's decision and saving costs of collecting information.

Let $w \in \{B, H\}$ be the true state of the world, where $w = B$ means that the business behavior under investigation is socially beneficial and $w = H$ means that it is socially harmful. We use the words 'procompetitive' and 'anticompetitive' as synonyms for 'socially beneficial' and 'socially harmful', respectively.

The aim of the CA's investigation process is to identify the true state of the world. The CA starts its investigation process by characterizing the behavior as belonging to a particular class, for which it knows — based on theoretical analyses and on experience — the size of the possible positive and negative welfare effects, $L_B > 0$ and $-L_H < 0$, respectively,[7] as well as some prior probability p that the welfare effect is positive. $1-p$ is the probability that the welfare effect is negative.

Starting from this *a-priori* information, the CA decides whether to gather more information or not. We assume that the additional information may be structured into *k* distinct criteria, which are labeled by an index $i \in K = \{1, 2, \ldots, k\}$. For each of these criteria, the CA may gather exactly one signal z_i whose outcome $m_i \in M \equiv \{-1, +1\}$ is dichotomous, and is $m_i = +1$ ('positive') with probability q_i^w, and $m_i = -1$ ('negative') with the remaining probability $1 - q_i^w$, for the true state of the world $w \in \{B, H\}$. The interpretation of the dichotomous signals may differ between the various criteria. For example, 'positive' and 'negative' may translate into 'high' or 'low' for the degree to which RPM is spread in a market or into 'manufacturers' or 'retailers' for the initiators of an RPM agreement. For simplicity, the signals are assumed to be stochastically independent and distinct in the sense that every criterion is a unique and exclusive indicator of a specific informational result (Heckerman et al 1992: 96). Generally, the signals are neither perfectly uninformative nor perfectly informative, i.e. the conditional probability (the likelihood) q_i^w that the outcome of signal z_i is positive is not the same for both states of the world and the outcome of signal z_i does not perfectly relate to one of the two states of the world: $0 < q_i^H \neq q_i^B < 1$.[8] Stochastic independence of signals allows us to further assume without loss of generality that

[7] We choose *L* as a descriptor of the welfare effect because loss functions will be at the center of our analysis.

[8] However, the general analysis also allows for perfectly informative signals at some points of the investigation process, but this is not referred to as the standard case and we concentrate on the case of criteria with imperfectly informative signals.

$0 < q_i^H < q_i^B < 1$.[9]

The CA may collect only one signal at a time. In other words, each time after receiving a signal, the CA has to make a decision on how to proceed further. Since the CA may collect up to *k* signals, there are up to $k+1$ decisions to be made: at each stage $j \in \{0,1,\ldots,k-1\}$, a decision d_j hast to be made, if that stage is reached. The decision has to select from the set of decision possibilities of the CA, $D_j = \{d^+, d^-, d^0\}$, which comprises three elements: two terminating decisions (permitting the behavior under investigation, d^+, and prohibiting it, d^-) and one continuative decision (collecting a j+1st signal, d^0). For the last stage *j* = *k* the set of decision possibilities comprises only the two terminating decisions, $D_k = \{d^+, d^-\}$.

After collecting a signal at any stage $j \in K$, the CA knows a history h_j which consists of a sequence of *j* pairs $\left(i_{\tilde{j}}, m_{\tilde{j}}\right) \in K \times \{-1,1\}$, in which the first entry is the index of the signal collected at stage $\tilde{j}$ and the second entry is the outcome of this signal. For any given history h_j the CA may determine updated probabilities of the true state of the world, which we denote by $\hat{p}\left(h_j\right)$ and, respectively $1-\hat{p}\left(h_j\right)$, by iterated Bayesian updating. With $q_j^w = q_i^w$ as the probabilities that signal *i* collected at stage *j* yields a positive outcome and h_j^m denoting the history h_j when the last signal yielded outcome $m \in \{-1,+1\}$, standard Bayesian updating yields

(1) $$\hat{p}(h_j^{+1}) = \hat{p}(h_{j-1})q_j^B / \Delta_j > \hat{p}(h_{j-1}) > \hat{p}(h_j^{-1}) = \hat{p}(h_{j-1})(1-q_j^B)/(1-\Delta_j)$$

with $\Delta_j = \hat{p}(h_{j-1})q_j^B + \left(1-\hat{p}(h_{j-1})\right)q_j^H$. To close the iterative updating we define $\hat{p}\left(h_0\right) \equiv p^B$.[10]

This updating of course requires knowledge of the conditional probabilities q_j^w. In the investigation process of real CAs, though, hardly anybody will be able to state these conditional probabilities. What administrators in the agencies however do know — or at least claim to know — are the posteriors $\hat{p}\left(h_j\right)$. Obviously, for any given sequence of signals, knowledge of the prior probabilities p and the posterior probabilities $\hat{p}\left(h_j\right)$ for all possible histories is equivalent to knowing the priors p and the likelihoods q_i^w of getting a positive signal.[11] We therefore

[9] Stochastic independence is a sufficient, but not a necessary condition. The inequality would only be violated if not only the size but also the sign of $q_i^B - q_i^H$ depended on the outcome of another signal.

[10] Thereby, *j* = 0 denotes the initial decision stage and the notation fulfils the general logic that the investigation of *j* criteria is necessary to calculate the posteriors of stage *j*.

[11] Reversing Bayesian updating yields: $q_j^B = \frac{\hat{p}\left(h_j^{+1}\right)}{\hat{p}\left(h_{j-1}\right)}\Phi$ and $q_j^H = \frac{1-\hat{p}\left(h_j^{+1}\right)}{1-\hat{p}\left(h_{j-1}\right)}\Phi$ with $\Phi = \frac{\hat{p}\left(h_{j-1}\right) - \hat{p}\left(h_j^{-1}\right)}{\hat{p}\left(h_j^{+1}\right) - \hat{p}\left(h_j^{-1}\right)}$. Assuming reversibility of Bayesian updating in this sense places strong requirements on the consistency of the CA's ideas about the history dependent probabilities $\hat{p}\left(h_j^m\right)$: Knowledge of $\hat{p}\left(\tilde{h}_j^{+1}\right)$ and $\hat{p}\left(\tilde{h}_j^{-1}\right)$ for

follow actual thinking inside CAs and present our model of sequential investigation as the *mirror image* of Bayesian updating, which is given by the *sorting* of cases (into ever more refined subclasses) according to the observed signals.[12]

We assume that the CA aims at minimizing the expected costs from its decisions. Costs accrue from collecting more information and from making wrong decisions. We call the former 'investigation costs' and the latter 'error costs'. Investigation costs of collecting signal z_i are $c_i > 0$. Unless otherwise stated, we will assume $c_i = c \; \forall \; i \in K$ for the remainder of the analysis. Error costs are given by a loss function $L(w,d)$, which represents the welfare loss when the true welfare effect is *w* and the CA chooses decision $d \in \{d^+, d^-\}$. Normalizing the loss from making the correct decision to zero ($L(B,d^+) = L(H,d^-) = 0$), we get $L(H,d^+) = L_H$ and $L(B,d^-) = L_B$.

2.2 Investigation and decision process

To determine the optimal decision process,[13] we first simplify the analysis by assuming that the sequence of criteria to be investigated is given, and that the CA's problem reduces to the question after which histories it should stop investigation and hand down a terminating decision. For notational ease, we assume that the signals are to be collected in their natural order, i.e. the *i*th signal may only be collected at stage $j = i$. Then the CA has to make a trichotomous decision at each stage except for the last stage, where only the two terminating choices are possible. While it is easy to decide on the terminating choices, the decision whether to terminate or to continue investigation of the case is more complex, since the costs resulting from continuation depend on the signals and choices of all later stages to be eventually reached (Lewis/Sheehan 1990: 374). For example, the total expected costs associated with the decision to collect signal 1 in a case with $k = 2$ (see Figure 2 for the decision tree) depend on the outcome of the first signal, on how the CA decides after the two possible outcomes and possibly on the outcome of the second signal and the consequential decisions of the CA. We therefore start with the comparison of the costs of the terminating decisions and then turn to the decision to continue investigation by looking at it in a backward manner.

---

any given history $\tilde{h}_{j-1}$ implies well-defined values of $\hat{p}(h_j^{+1})$ and $\hat{p}(h_j^{-1})$ for any other history h_{j-1} of the same length.

Given our assertion that administrators in CAs usually fail to know the values of the conditional probabilities q^B and q^H, but rather rely on their (claimed) knowledge of the posterior probabilities for all histories, one may doubt whether they do correctly calculate the unconditional probabilities of receiving a positive or a negative outcome of a signal. If they do not, a legal optimal sequential investigation rule should acknowledge this and should be structured accordingly. However, this goes far into behavioural economics and is beyond the scope of our paper.

[12] see e.g. Lindgren et al. (1978: 220-221), Marschak/Miyasawa (1968: 146ff.) or Hirshleifer/Riley (1992: 187 ff.)

[13] The analysis is based on Lewis/Sheehan (1990: 368-375), Vos (2007: 612-620) and Gorry/Barnett (1968: 492 ff.).

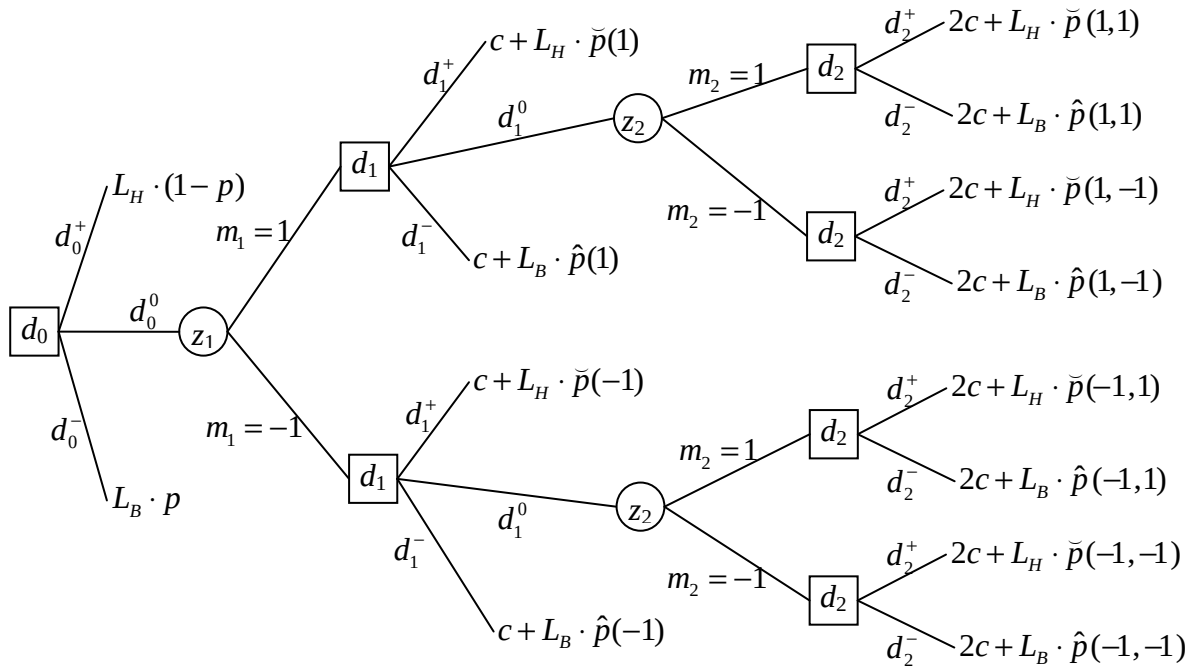


Figure 2: Decision tree with $K = 2$; we abbreviate $\breve{p}(h_j) \equiv 1 - \hat{p}(h_j)$.

Let $EC(h_j, d)$ be the expected costs from handing down decision $d \in D_j$ after history h_j and define $EC^*(h_j) = \min_{D_j}\left[EC(h_j, d^+), EC(h_j, d^-), EC(h_j, d^0)\right]$ as the minimal expected decision costs after history h_j and $d^*(h_j)$ as the optimal decision after history h_j, i.e. the decision which achieves $EC^*(h_j)$.

It is easy to see that the total expected costs from permitting or prohibiting the behavior under scrutiny after history h_j are $EC(h_j, d^+) = j \cdot c + L_H \cdot \left(1 - \hat{p}(h_j)\right)$ and $EC(h_j, d^-) = j \cdot c + L_B \cdot \hat{p}(h_j)$, respectively. Obviously, at stage j the CA will select d^+ among the two terminating decisions if and only if

(2) $$\hat{p}(h_j) \geq L_H / (L_H + L_B).$$[14]

Finally, $EC(h_j, d^0)$ is given by the expected value of the minimal expected decision costs of the next stage, where the expectation is taken over the probabilities of the next signal. From the CA's viewpoint, these probabilities depend on the history of signals, but not on the true state of the world. They are thus given by:

[14] For simplicity, we assume that the CA permits if it is indifferent between permitting and other decisions.

(3)
$$q_j\left(h_{j-1}\right)=\hat{p}\left(h_{j-1}\right)q_j^B+\left(1-\hat{p}\left(h_{j-1}\right)\right)q_j^H=\frac{\hat{p}\left(h_{j-1}\right)-\hat{p}\left(h_j^{-1}\right)}{\hat{p}\left(h_j^{+1}\right)-\hat{p}\left(h_j^{-1}\right)}$$
$$\text{and}\quad 1-q_j\left(h_{j-1}\right)=\frac{\hat{p}\left(h_j^{+1}\right)-\hat{p}\left(h_{j-1}\right)}{\hat{p}\left(h_j^{+1}\right)-\hat{p}\left(h_j^{-1}\right)}$$

With these probabilities, we get the following expression for the expected costs from handing down decision $d_j=d^0$:

(4) $$EC\left(h_j,d^0\right)=q_{j+1}\left(h_j\right)EC^*\left(h_{j+1}^{+1}\right)+\left(1-q_{j+1}\left(h_j\right)\right)EC^*\left(h_{j+1}^{-1}\right)$$

We can then determine the optimal decision and the consequential minimal expected decision costs in a backwards iterated manner. At stage *k*, the optimal decision is given by:

(5) $$d_k^*=\arg\min_{d\in D_k}\left[EC\left(h_k,d\right)\right]\quad\text{where}\quad D_k=\left\{d^+,d^-\right\},$$

which is equivalent to

(6) $$d_k^*(h_k)=\begin{cases}d^+ & \text{if } \hat{p}\left(h_k\right)\geq L_H/\left(L_H+L_B\right)\\ d^- & \text{otherwise}\end{cases}.$$

As a consequence, the minimal expected decision costs at stage *k* are given by

(7) $$EC^*\left(h_k\right)=\min\left[k\cdot c+L_H\left(1-\hat{p}\left(h_k\right)\right),k\cdot c+L_B\hat{p}\left(h_k\right)\right].^{15}$$

Starting from these results for $j=k$, we can iterate backwards for $j<k$ by following

(8) $$d_j^*=\arg\min_{d\in D_j}\left[EC\left(h_j,d\right)\right]\quad\text{where}\quad D_j=\left\{d^+,d^-,d^0\right\}$$

as decision rule, which is well-defined and yields

(9)
$$EC^*\left(h_j\right)$$
$$=\min\left[j\,c+L_H\left(1-\hat{p}\left(h_j\right)\right),\ j\,c+L_B\hat{p}\left(h_j\right),q_{j+1}\left(h_j\right)EC^*\left(h_{j+1}^{+1}\right)+\left(1-q_{j+1}\left(h_j\right)\right)EC^*\left(h_{j+1}^{-1}\right)\right]$$

as minimal expected decision costs at stage $j<k$.

We will sometimes refer to the net information value of the criterion studied at stage *j* given by the costs to be saved by studying this criterion:

(10) $$NIV_j\equiv\min\left[EC\left(h_{j-1},d^+\right),EC\left(h_{j-1},d^-\right)\right]-EC\left(h_{j-1},d^0\right).$$

The decision rules (5) and (8) jointly define investigation *rules* for every possible class of business practices, which may differ according to the characteristic of each class (i.e. the priors), the set of criteria and the posteriors. The rules assign to each observed signal the (optimal) action to be taken by the CA, namely either terminating the investigation by deciding between permission and prohibition or continuing by the investigation of another criterion. The realized paths through the decision trees specify the degree of differentiation of each class, i.e. the number and sequence of investigated criteria.

[15] One may replace *k* by any *j*<*k* in equations (6) and (7) whenever one is – for whatever reason – sure that at stage *j* one of the terminating decisions is taken.

It is not surprising that the investigation rules also depend on the size of the welfare effects and the investigation costs. Here a number of questions can be analyzed, and also some of our restrictive assumptions can be revoked and extended. The model implies that, firstly, an increase of the quantitative welfare effects of pro- or anticompetitive behavior deepens the investigation of any class or leaves it unaffected. Secondly, a general increase of investigation costs leads to an earlier end of the investigation process of any class or leaves it unaffected. Thirdly, the relative size of the pro- and anticompetitive welfare effects influences the probability threshold when deciding in favor of permission or prohibition. For example, if the error costs through an erroneously allowed (anticompetitive) RPM (false negative) are much higher than the error costs through a falsely prohibited (welfare-enhancing) RPM (false positive), then already a small share of anticompetitive cases of RPM can lead to the recommendation of a prohibition, if additional investigation does not pay off. Therefore, a large quantitative welfare effect can offset a low probability of occurrence.[16] However, the general determination of stopping criteria remains the same. Another modification of the model involves the introduction of the possibility of different investigation costs for different assessment criteria, which also relates to the problem of an optimal sequencing of the assessment criteria (see section 3).

2.3 Application: Resale Price Maintenance as an Example

For illustrating how such an analysis can be applied to a specific competition law problem, we use the already mentioned example of Resale Price Maintenance (RPM) with the Leegin decision as a starting-point.[17] The Supreme Court decided that the simple per-se prohibition is no more appropriate and RPM should be decided according to the rule of reason, which requires the search for a more differentiated approach.[18] The discussion how such a more differentiated rule for RPM in US antitrust law should look like is still in the beginning. One may expect that some discussion about the so far undisputed per-se prohibition of RPM will now emerge in EU competition law, too.[19]

Due to the large number of positive, negative and ambiguous effects of RPM, a modern industrial-economics perspective on RPM casts sever doubts on whether a per-se prohibition of RPM is still the appropriate rule.[20] However, a complete economic assessment of RPM is too complex to be carried out in every case. In the following numerical example we do not aim at

[16] Only in the case of symmetric losses, the CA should decide in favour of the more probable welfare effect. If, however, L_H is assumed to be twice as large as L_B, the probability threshold when deciding in favour of permission equals $\hat{p}^B = \frac{1}{3}$ - see equation (2). The assumption of symmetric losses indicates equal undesirability of false positives and false negatives (Lewis/Sheehan 1990: 370). See for a similar conclusion also Beckner/Salop (1999: 61-62) and Heyer (2005).

[17] *Leegin Creative Leather Products, Inc. v. PSKS, Inc.*, 551 U.S. ___ (2007), *Sup Ct* (USA), No. 06-480

[18] *Leegin*, p. 19, 28

[19] See Carlton/Heyer (2007: 135), Hay (2007: 35), Kneepkens (2007: 664), and Cavanagh (2008).

[20] For overviews about the economic assessments of RPM see, e.g., Mathewson/Winter (1998), Scherer/Ross (1990: 548-558), Rey/Caballero-Sanz (1996: 27-30), Carlton/Perloff (2005: 425-437), Motta (2004: 302-378), Kneepkens (2007), Kerber/Schwalbe (2008: 331-335).

providing a sophisticated ("state of the art") analysis of how RPM should be treated in the future or whether the decision of the US Supreme Court was right in the first place. This would be beyond the scope and intention of this article. Nor do we want to claim that the applied assessment criteria and the assumed conditional probabilities in our example are the most realistic ones. The following example only serves as a brief illustration of how our model can be applied for searching for a more differentiated rule for RPM after the Leegin decision.

We visualize the example in Figure 3, in which the numbers in the decision nodes (rectangles) are the probabilities that RPM is procompetitive, given the information collected up to this decision node ($\hat{p}\left(h_j\right)$). The **bold** numbers at the end nodes are the expected total costs (investigation costs plus error costs) accruing from deciding for the respective end node. The numbers in parenthesis under the continuation branches (d^0) are the expected cost from selecting continuing investigation, given optimal further decisions. Finally, the *italic* numbers at the branches from signal nodes to decision nodes are the unconditional probabilities that the outcome of the signal is positive (upper branch: $q_j\left(h_{j-1}\right)$) or negative (lower branch: $1-q_j\left(h_{j-1}\right)$). We have underlined the minimal expected costs at every decision node and marked the path of the optimal decision rule by bold lines.

We chose specific numbers according to the following intuition. We assume that RPM usually is anticompetitive but in some cases has welfare-enhancing effects. We express this in our numerical example by assuming that in $p=\hat{p}\left(h_0\right)=15$ per cent of the cases RPM has positive welfare effects, whereas in 85 per cent they are negative. We also assume that the welfare losses through false positives and false negatives are equal and normalized to 1, investigation costs are c = 0.02, and there are three assessment criteria — market shares, role of initiator of RPM in the market, and completing a full-fledged welfare analysis.[21] We assume that one observes "low market shares" (say, below 10 per cent) with a probability of $q_1^B=0.82$ if the RPM agreement under investigation is beneficial and with a probability of $q_1^H=0.22$ if it is not. For $q_2^B=0.66$ of the beneficial RPM agreements, so we assume, one observes that the manufacturer initiated the agreement and for $q_2^H=0.37$ of the harmful RPM agreements. Finally, we assume for the rest of a full-fledged welfare analysis that it yields the outcome "this RPM agreement is beneficial" with a probability of $q_3^B=0.71$ if it actually is and with a probability of $q_3^H=0.4$ if it actually is not. Hence, we say that "low market share", "manufacturer initiated", and, respectively "full-fledged welfare analysis yields positive welfare effect of specific RPM agreement" are the signal outcomes $m_i=+1$. These conditional probabilities of the signal outcomes, which the CA does not consider directly, yield the history-dependent probabilities that the specific RPM agreement

[21] Another self-suggesting criterion would be the extent to which RPM prevails in the market as a whole. Wide spread use of RPM would indicate a welfare loss, as it tends to impede competition among retailers. While considering this criterion as well would make our model more realistic, we refrain from including it in the model to keep the presentation as simple as possible.

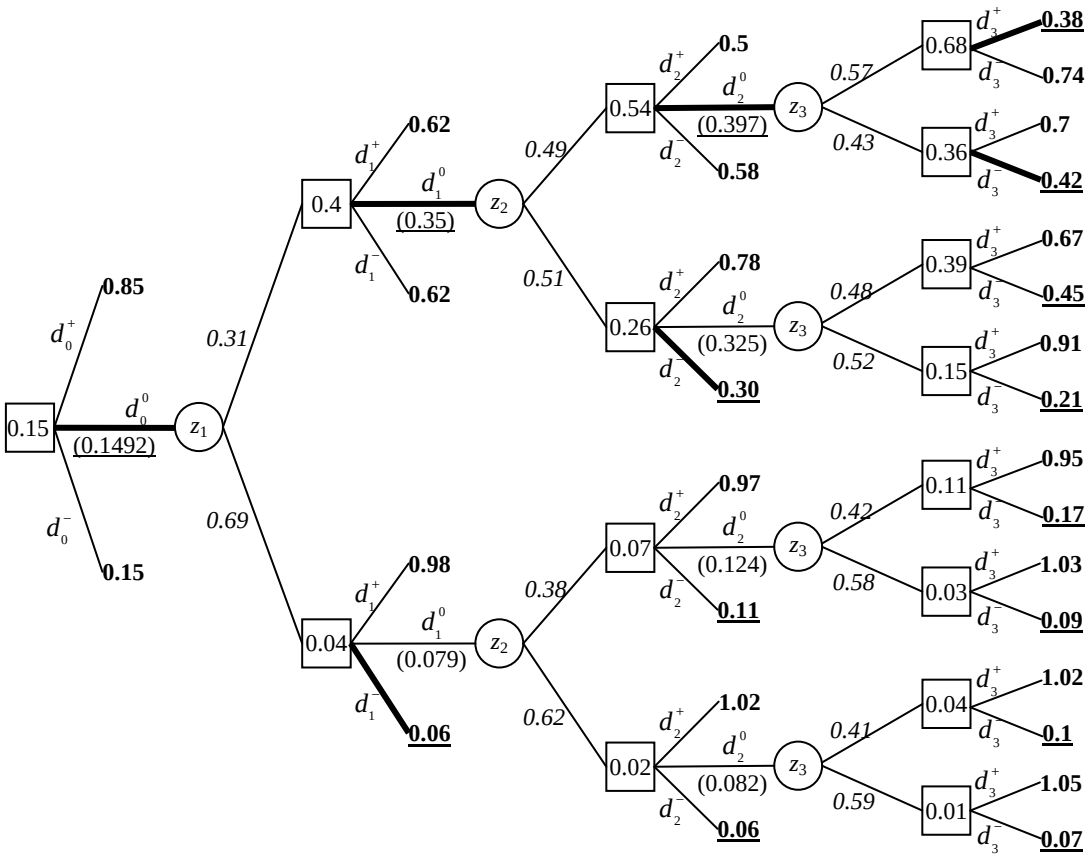


Figure 3: Numerical example for deriving a differentiated rule for RPM with three assessment criteria.

is welfare enhancing as stated in the decision nodes of Figure 3. It is these probabilities on which authorities rely when making decisions.

We derive the optimal sequential decision rule from this multi-branched decision tree by backward induction relying on the total expected costs indicated at the end nodes. At the last stage, prohibiting the RPM induces lower expected total costs in all subclasses except for the low-market-share RPM initiated by manufacturers with a positive assessment from the in-depth investigation. Hence, in all second-level subclasses except for the low-market-share RPM initiated by manufacturers the detailed welfare analysis fails to affect the final decision and thus induces investigation costs without any chance to reduce error costs. As a consequence, an in-depth analysis of the welfare effects will only be performed in the subclass of low-market-share RPM initiated by manufacturers.

In all other subclasses of first level, prohibiting RPM is the decision minimizing expected total error costs. So there is no reason to investigate the role of the initiator of RPM in the market within the subclass of high-market-share RPM. Prohibition of RPM is the best action in any way. For low-market-share RPM, investigation of the initiator is worthwhile, however, as indicated by a comparison of the expected total costs from investigating this criterion (cf. the number in

parenthesis below the d_1^0-branch) and the expected total costs indicated at the end nodes d_1^+ and d_1^-. The bold lines in Figure 3 summarize this argument by indicating what decisions should be made at which decision node.

The specific numbers we chose hence yield the following decision rule: "If market share is above 10 percent ($m_1 = -1$), then interdict RPM, otherwise investigate who initiated the agreement. If it was a retailer, also interdict RPM, otherwise complete a full-fledged welfare analysis and decide according to its result." This happens to be similar to what could be a general rule after the Leegin decision: There is a broad consensus that the anticompetitive effects of RPM depend critically on the market shares. Small market shares may allow for an exception, but high market shares always trigger prohibition of RPM agreements. Given the outcome of the signal "market share", the CA sorts the case at hand into one of the two resulting subclasses. It will turn out that in the subclass "high market share" (lower branch of the first signal node) the probability of still having procompetitive RPM — to remember: this is the data the CA knows or thinks to know — is small enough to terminate the information gathering process at this stage with the decision to prohibit RPM. Additional information cannot alter this probability enough to justify the additional investigation costs.

For the other subclass, "low market share" anticompetitive RPM is more likely than procompetitive RPM as well, but the difference is smaller. Sorting the cases into further subclasses according to who initiated the agreement may lower the expected total costs despite the investigation costs. In fact, with the numbers we chose, it does. Initiation by manufacturers ($m_2 = +1$, as in the Leegin case) suggests that RPM increases efficiency, while initiation by retailers ($m_2 = -1$) suggests that its sole goal is price collusion among retailers.[22]

However, even when the manufacturer initiated the RPM agreement (and his market share is low), the probability of having one of the few cases of socially beneficial RPM, although being larger than 0.5, is still not large enough to completely rely on the two criteria hitherto studied: The rest of a full-fledged welfare analysis is reliable enough (the sum of q_3^B and $\left(1-q_3^H\right)$ is large enough) to justify its investigation costs. This last investigation step may reduce this probability so clearly below one half, that it is worthwhile conducting it. This is not the case, when initiation by a retailer lowers the probability of the RPM agreement being welfare enhancing down to 26 percent. In fact, even a positive result of the last investigation step would not bring back this probability above one half.

Admittedly, such an analysis seems to be complex and error prone, given the large amount of information on probabilities required and the limited ability of human beings making decisions in CAs. Had such an analysis to be performed for each single case under investigation of a CA, the costs would be likely to outweigh the benefits from improving the decisions. However, the

[22] Also in the Leegin decision, these criteria are given as relevant investigation factors (see *Leegin*, p. 17-18).

starting point of our analysis was sorting into a limited number of case classes and what we are looking for is a general rule on the order and depth of information collection for each class. Such a rule need not be perfect, but should — and in many cases outside competition law does — reflect the general structure of our argument: neither per-se rules nor complete in depth analyses nor intermediate rules with a generally fixed level of investigation depth are optimal, but sequential investigation rules with stop rules which depend on the information gathered at previous stages.

For the case and the numerical example at hand, the following differentiated rule would result for assessing the permissibility of an RPM agreement:
(1) In a first step, investigate the market share. If the market share is larger than 10%, the RPM agreement is prohibited.
(2) If the market share is not larger than 10%, then investigate whether this RPM agreement was initiated by the manufacturers or by the retailers: If it was initiated by the retailers, then this RPM agreement is prohibited.
(3) If it was initiated by the manufacturing firm, then the RPM agreement is permissible if an in-depth investigation of efficiency effects and anticompetitive effects yields a net positive result, otherwise it is prohibited.
We actually find such differentiated rules in many fields of law, ranging from penal law to nuisance law.[23]

One should note that the exact structure of the legal rule may depend both on the informativeness of the various assessment criteria and on the costs of collecting the corresponding signals. In fact, in our example presented so far, the differentiated rule would reduce the expected total costs only to 0.1492 as compared to 0.15 in the case of the simple per-se prohibition. As a consequence, for example an increase in the investigation costs by only 6 per cent to 0.0212 for each criterion investigated would reverse the result: the difference of the expected total costs would still be roughly 0.0008, but now to the advantage of the simple per-se prohibition. Similarly, increasing the investigation costs only of the last criterion, i.e. of the in-depth analysis of welfare effects, to 0.03 would also reverse the result. However, allowing for different costs of the various criteria not only affects the optimal termination stage of the differentiated analysis, but may also affect the optimal ordering of the criteria. We deal with this and further variations in the next section.

3. Optimal Sequencing

3.1 Introduction

So far we have assumed the existence of a given sequence of investigation criteria. We limited our analysis to the problem of the optimal investigation depth in each of the branches of the decision tree. However, a general solution of the optimization problem would require also the

[23] For the latter see Gomez-Pomar/von Wangenheim (2008)

solution of the more complex problem of optimal sequencing. Generally, optimal rules minimize the total expected cost at every point of the investigation procedure (Vos 2007: 612). This optimization calculus for sequential investigation procedures depends on the set of possible criteria *K* and on the sequencing of these criteria. Realistic investigation procedures must include the option of investigating different criteria in a way, which depends on previously observed signals. Even in the simplest case of a *stage-consistent sequence* (i.e. every criterion can be investigated only once, only one criterion can be investigated at every investigation point and the criteria are the same for every point of each investigation stage, no matter which signals have been observed before getting to a certain investigation point) all possible sequences of criteria have to be examined and solved by backwards induction to determine the optimal sequence (Gorry/Barnett 1968: 496-497). Furthermore, the "number of sequences grows exponentially with the number of tests" i.e. 'criteria' in our analysis (Heckerman et al. 1991: 2).[24]

Therefore, by determining the stopping criteria for every class several possibilities of realized investigation paths must be considered. Each possibility determines a concrete investigation procedure, which encompasses all terminal decisions in dependence of the observed signals and investigated criteria. By changing the sequence of the criteria, the investigation procedure is also changed.

3.2 Flexible sequencing problem

The most general case is the case of a *flexible sequence*: the information value of the criteria is different according to previously observed signals and investigated criteria. For example, the information value of investigating market entry barriers may differ according to previously observed market shares and the CA may prefer the investigation of different criteria at the next stage. In general, the posterior probabilities of each criterion are not the same at every stage. The CA, therefore, must have knowledge about any possible posterior distribution with respect to every possibility of previously investigated criteria and observed signals. Further, and more importantly, the number of possible sequences grows more than exponentially as in the traveling-salesman problem. In our example of only three criteria, there are six ways of ordering them, if the ordering is the same for all branches of the decision tree. If the ordering may vary across the branches, the number of possible orderings increases to twelve. For four criteria the respective numbers are $4! = 24$ and $4 \cdot 3^2 \cdot 2^4 \cdot 1 = 576$, respectively. With six criteria, a computer calculating one thousand sequences per second would need more than 500 years to calculate all $6 \cdot 5^2 \cdot 4^4 \cdot 3^8 \cdot 2^{16} = 1.6 \cdot 10^{13}$ possible orderings when the ordering may vary across the branches.

While the basic analysis and determination of investigation rules via stopping criteria is not affected, the huge number of possible sequences for any larger number of criteria makes simple

[24] See also Kalagnanam/Henrion (1990: 272), Heckerman et al. (1992: 97) and Tseng/Gmytrasiewicz (1999: 3)

heuristic solutions to manage the complex problem of optimal sequencing inevitable. In the following parts we examine conditions, which allow for a predetermined sequence (*bracketing condition*) and the determination of nearly optimal sequences (*myopic approach*). Still, such a heuristic aims at minimizing the total expected costs (i.e. the sum of expected investigation costs and losses).[25]

3.3 'Bracketing condition'

An "easy case" is given if some or all criteria are equally costly and satisfy the following 'bracketing condition':

We say that a criterion i *'brackets'* another criterion $i' \neq i$ for a history h_j of collected signals and their outcomes, if

$$\hat{p}\left(h_j \cap (i,1)\right) > \hat{p}\left(h_j \cap (i',1)\right) > \hat{p}\left(h_j \cap (i',-1)\right) > \hat{p}\left(h_j \cap (i,-1)\right)$$

or

$$\hat{p}\left(h_j \cap (i,-1)\right) > \hat{p}\left(h_j \cap (i',-1)\right) > \hat{p}\left(h_j \cap (i',1)\right) > \hat{p}\left(h_j \cap (i,1)\right)$$

is satisfied, where $h_j \cap (i,m)$ denotes the extension of history h_j by the collection of signal i and its outcome m and $h_j \cap (i',m)$ is defined accordingly.

Proposition: If criterion i brackets criterion i', then the latter should *not* be investigated before the former.

The intuition behind this proposition is simple. Consider the first version of the inequality sequence and assume that $\hat{p}\left(h_j \cap (i',1)\right) \geq L_H / \left(L_H + L_B\right) > \hat{p}\left(h_j \cap (i',-1)\right)$. Then both criteria may determine the choice of a terminating decision. Suppose both criteria do and thus exclude further investigations. Then the expected losses from permitting the behavior under investigation are smaller, if the decision is based on criterion i rather than on criterion i' since $L_H\left(1-\hat{p}\left(h_j \cap (i,1)\right)\right) < L_H\left(1-\hat{p}\left(h_j \cap (i',1)\right)\right)$. Similarly, expected losses from prohibiting the behavior under investigation are smaller, if the decision is based on criterion i rather than on criterion i' since $L_B\hat{p}\left(h_j \cap (i',-1)\right) > L_B\hat{p}\left(h_j \cap (i,-1)\right)$. Hence the decision should be based on criterion i, which implies that i has to be investigated before i' as both end the investigation process. Now suppose that only criterion i determines the choice of a terminating decision. Then collecting signal i' only adds costs c without changing the decision. Finally note that the case that only i' determines the choice of a terminating decision is impossible. A formal proof of the Proposition based on Hirshleifer/Riley (1992: 188) is given in appendix A.

[25] See Vos (2007: 611) and Marschak/Miyasawa (1968: 144)

Corollary: If criterion i brackets all other criteria, it should be investigated first. If among the remaining criteria, one criterion i' brackets all other remaining criteria, than criterion i' should be investigated next. This rule should be repeated as often as possible.
If each criterion i brackets all criteria $i' > i$, then the each criterion i should be investigated at stage $j = i$, (i.e. the optimal sequence is the natural order of the criteria).

The proof of this Corollary is obvious.

The corollary may provide a complete ordering of the criteria to be investigated. Up to which criterion the investigation should actually be carried out then depends on the costs of investigation as discussed in Section 2 of the paper.

According to the sorting approach the CA has knowledge about the sorting characteristics of the different criteria at each possible stage, reflected by the corresponding posteriors. The smaller posterior risk of criterion i is equivalent to a higher discriminatory power of the signal observation, which results in lower expected loss compared to any other criterion.

Obviously, the bracketing condition yields a complete ordering of all possible investigation criteria only in rare cases. We therefore discuss alternative approaches to find a reasonable ordering of criteria possibly to be investigated.

3.4 'Myopic approach'

The "easy case" of the 'bracketing condition' makes high demands on the characteristics of the different criteria. If these cannot be met, the problem of intractability of computing all possible sequences can nevertheless be alleviated but not solved. Gorry/Barnett (1968) proposed a 'myopic approach'. As in the model presented here, they assume that only one criterion can be investigated at every stage, i.e. one test (criterion) at a time. The idea of their heuristic is to cut off all further investigation stages in the selection of the next criterion to be studied and to calculate and compare the optimized total costs of handing down a terminating decision based on any of the hitherto unused criteria at stage j.[26] Formally, the myopic approach suggests to select the criterion with the highest 'myopic information value' for investigation in any stage, where the 'myopic information value' is given by the reduction in costs resulting when only this one more criterion is investigated, i.e. by

[26] See Gorry/Barnett (1968: 498-499), Kalagnanam/Henrion (1990: 272) and Heckerman et al. (1991: 4)

$$
\begin{aligned}
&\min\Big[\, EC\left(h_{j-1}, d^{+}\right), EC\left(h_{j-1}, d^{-}\right)\Big] \\
&-\Big[\, q_i\left(h_{j-1}\right)\, EC_j^{*}\left(h_{j-1}\cap(i,1), d\in\left\{d^{+}, d^{-}\right\}\right) \\
&\qquad +\left(1-q_i\left(h_{j-1}\right)\right) EC_j^{*}\left(h_{j-1}\cap(i,-1), d\in\left\{d^{+}, d^{-}\right\}\right)\Big],
\end{aligned}
\tag{11}
$$

where $q_i\left(h_{j-1}\right)=\frac{\hat{p}\left(h_{j-1}\right)-\hat{p}\left(h_{j-1}\cap(i,-1)\right)}{\hat{p}\left(h_{j-1}\cap(i,+1)\right)-\hat{p}\left(h_{j-1}\cap(i,-1)\right)}$ is the unconditional probability that signal *i* yields outcome $m=+1$ at stage *j* given history h_{j-1} (cf. equation (3), and $EC_j^{*}(\cdot)$ is the optimized total costs of handing down a terminating decision based on criterion *i* at stage *j* given history h_{j-1}. One should note that this myopic information value differs from the net information value defined in equation (10) by the cost reduction which may be possible by investigating criteria beyond the signal *i*. Obviously, maximization of the myopic information value is equivalent to solving the following minimization problem:

$$
\begin{aligned}
&\min_{i\in K_j}\Big[\, q_i\left(h_{j-1}\right)\, EC_j^{*}\left(h_{j-1}\cap(i,1), d\in\left\{d^{+}, d^{-}\right\}\right) \\
&\qquad +\left(1-q_i\left(h_{j-1}\right)\right) EC_j^{*}\left(h_{j-1}\cap(i,-1), d\in\left\{d^{+}, d^{-}\right\}\right)\Big],
\end{aligned}
\tag{12}
$$

where K_j is the set of criteria not yet investigated at stage *j*.

Myopia of this approach also extends to the stop rule: while the original model requires considering all costs of further investigation stages when deciding on whether to study the next criterion, the myopic approach requires the CA only to decide whether the maximal myopic information value is still positive. If it is not, no more criteria should be studied.[27] Further costs reductions resulting from improving the decision at later investigation stages are neglected.

Both simplifications induce larger costs as compared to the complete procedure to find the optimal sequence of criteria to be investigated or not. On the one hand, the myopic approach may induce a suboptimal sequence as we show by an example in Appendix B even for stochastically independent criteria. On the other hand, the myopic approach therefore tends to result in a too early termination of the investigation process even when the sequence is correct.[28] The reason is simple: by neglecting the possibility of more than one additional investigation stage, the myopic approach also neglects the possibility of saving costs by investigating more than one additional criterion. It is noteworthy that the myopic approach may never stop the sequence of investing criteria too late: should the complete study of possible sequences of investigating the existing criteria indicate that terminating the investigation is the best alternative, then there obviously is no more criterion which is worthwhile to be investigated in the next stage; whence the myopic criterion will also prescribe termination of the investigation.

[27] See Mookerjee/Mannino (1997: 683-684), Mookerjee/Dos Santos (1993: 116 and 118), Beckner/Salop (1999) and Kalagnanam/Henrion (1990: 273-274)

[28] The investigation of a single criterion might have no positive myopic information value, while the net information value for a sequence of criteria is positive (Kalagnanam/Henrion 1990: 272 and Mussi 2002: 103).

Despite the two pitfalls discussed in the previous paragraphs, as Gorry/Barnett (1968: 505) argue, the myopic approach may be seen as a good heuristic in practice, where exact knowledge of the various probabilities as well as the exact size of possible benefits and harms of the behavior under scrutiny is lacking. In addition, if the bracketing condition is satisfied, the myopic approach always results in optimal sequencing and the optimal termination rule.

We finally note that the myopic approach also provides reasonable results if the investigation costs c_i vary for the different criteria. All one has to change in this case is that the decisive criterion is not simply the maximal myopic information value of the remaining criteria, but the myopic information value relative to the investigation costs.[29] As before, the stopping criterion for a certain path is generally given by the condition of a negative myopic information value of all remaining possible criteria.

4. Applications and Discussion

In this article we presented a decision-theoretic model that can be used as an instrument for deriving optimal rules for a sequential investigation process of an administrative agency (as a competition authority) that has the task to determine whether a certain behavior should be allowed or forbidden because of its welfare effects. Although a deeper investigation can help to discriminate better between anti- and procompetitive behavior, both the limited possibilities of reducing the error costs (through the uncertainty of making correct decisions; Heyer 2005: 385) and the rising investigation costs lead to the conclusion that in most cases only a more or less limited depth of investigation into the welfare effects of business behavior is optimal, leading to the notion of an optimal complexity (or differentiation) of legal rules (measured by the optimal number of assessment criteria that are applied in an investigation). In this article, we analyze a sequential investigation process, which can be interpreted as a step-by-step sorting process into ever smaller subclasses of cases that help to discriminate better between pro- and anticompetitive cases. The result are rules for the sequential investigation of certain classes of business behavior, which consist of a set of assessment criteria, their optimal sequencing and the extent of the investigation, which can be different within a multi-branched decision tree. For deriving optimal sequential investigation rules, therefore three problems must be solved: (1) What assessment criteria should be used? (2) In what sequence should these assessment criteria be applied? (3) What is the optimal depth of investigation in the different branches? In our basic model in section 2 we focused on the third question of optimal stopping rules within a decision tree with a given sequence of assessment criteria. The second question was dealt with in section 3 that showed that optimal sequencing can be a very difficult problem. However, under certain assumptions (bracketing condition) this problem can be simplified and also heuristic solutions (myopic approach) are possible that might lead to satisfactory solutions. We do not focus on the first question.

[29] See Mookerjee/Dos Santos (1993: 116 and 118) and Kalagnanam/Henrion (1990: 273-274); a similar result is given in Beckner/Salop (1999: 48 and 51), i.e. to gather less costly and/or potentially dispositive information first.

Two additional extensions of the basic model might be very interesting from the perspective of the practical application of competition law. (1) Often the investigation of a criterion does not simply lead to the observation of a dichotomous signal ($m=+1$ or $m=-1$, e.g. high or low market share), but to the observation of a specific value within a given range of values. This additional information can be used better, if we define (instead of only two) three or more signal values (e.g., high, intermediate, and low market shares) using specific ranges. The use of the HHI in merger analysis is an example; here the whole range of possible signal values is divided into several subintervals,[30] within which values are treated equally. For an optimal assessment, also the exact specification of these subintervals will play a crucial role. (2) Another important generalization is to take into account a probability distribution for a range of welfare effects, which would replace our simple assumption of only dichotomous welfare effects, $w \in \{B, H\}$.[31] Although the model would be more complicated, the general calculus of determining stopping criteria is not changed for both extensions. In case (2), the problem of optimal sequencing can be simplified by first-degree stochastic dominance as well (as through the bracketing condition in the basic model). Therefore our model might offer extensions that reduce the unrealistic assumptions of dichotomous signals and welfare effects.

In section 2.3 we applied our approach to the problem of the appropriate competition law rule for RPM as an example. Most important is that our analysis demonstrates clearly the very different economic perspective that leads to different relevant questions. The decisive question now is whether it is worthwhile to replace the old per-se prohibition rule through another rule. This requires the comparison of alternative rules in regard to its welfare effects in terms of the sum of error costs and investigation costs. The analysis demonstrates clearly that it is not sufficient to show the existence of welfare-enhancing RPMs for rejecting the per-se prohibition of RPM. Rather it is necessary to have information about the probability of welfare-enhancing or anticompetitive RPMs (priors) and the size of their positive and negative welfare effects, about the investigation costs for discriminating between anti- and procompetitive RPMs, and about the additional error costs that emerge through erroneous permission of anticompetitive RPMs. The theoretical analyses of RPM by industrial economics are not able to give reliable answers to these empirical questions. It is an interesting fact in the Leegin decision, that it is Justice Breyer (in his dissenting opinion), who asked - without recurring to an explicitly economic argumentation - exactly these questions. His different stance is not based upon a rejection of the arguments of industrial economists about the potential positive welfare effects of RPMs, but upon his scepticism whether it is worthwhile to replace the per-se prohibition rule through another more differentiated (and so far unknown) rule.[32]

[30] Of course, the classification into just two subintervals corresponds to the assumption of dichotomous signals.

[31] See, e.g., Christiansen/Kerber (2006: 229 ff.)

[32] Justice Breyer's questions about RPM are: "[H]ow often are harms or benefits likely to occur?" "How easily can courts identify instances in which the benefits are likely to outweigh potential harms?" What additional errors are made through courts and juries by taking into account complex economic criteria? (*Leegin*, dissenting opinion, pp. 9-10)

These information problems raise critical questions whether such a decision-theoretic approach for deriving optimal sequential investigation rules is feasible at all. However, three considerations should be taken into account: (1) Our knowledge about the effects of business behavior in competition would be grossly underestimated, if we think we would not have considerable experience about the probability and size of beneficial or harmful behavior or do not have reliable assessment criteria which help us to sort cases in more problematic or unproblematic ones. The problem is more that so far we have not analyzed competition problems from this economic perspective. (2) Consequently, there is much future research to be done. This refers on one hand to a much more precise and sophisticated theoretical analysis how such a comparison of alternative sequential investigation rules should be done and what kinds of empirical knowledge is necessary. Here our model can be seen only as a first step. Equally important is empirical research that helps to make better estimations about the distributions of anti- and procompetitive behaviors, about the effectiveness of assessment criteria for discriminating between beneficial and harmful cases etc. (3) The third argument refers to the empirical fact that the competition law practitioners have always applied implicitly such a kind of analysis, without making it explicit and without a clear theoretical foundation. All the intermediate solutions in competition law between per-se rules on one hand and a full-scale market analysis (as the extreme form of rule of reason) on the other hand, as, e.g., safe harbor rules (with market share thresholds), quick look rules, or the definition of ranges of HHIs that sort mergers into subclasses of different probabilities for their anticompetitiveness, use implicit notions of probabilities for harmful or beneficial effects and the potential size of error costs. Our approach offers the chance to make these notions explicit and discuss them critically from a clear theoretical perspective.

In our model, we asked for optimal sequential investigation rules for a competition authority, which assesses a certain business behavior. It would be no problem for our approach to use it also for structuring a sequential procedure in the case of private litigation, in which a plaintiff sues a defendant for infringing competition law and claiming damages. This is the example Beckner/Salop (1999) use in their application of such a decision-theoretic multi-stage investigation approach. The problems of the optimal sequence and the optimal depth of investigation are similarly relevant in a private trial, in which first the plaintiff has to prove the fulfillment of certain criteria, and afterwards the defendant has the possibility to rebut a presumption by providing information about other criteria etc. Empirical evidence about private litigation in US antitrust law shows clearly that most cases are already decided in early stages of such a multi-stage process. Only in rare cases all criteria with an overall balancing of all positive and negative effects are investigated,[33] i.e. the question of optimal sequencing of criteria and optimal stopping is also very important in this context.

Our approach fits very well into a new discussion both in US antitrust law and EU competition law about the necessity of clear structured procedures how to assess whether a business behavior should be prohibited or allowed. Within EU competition policy, the awareness increases that

[33] See Carrier (1999)

despite the results of modern industrial economics, which suggest more differentiation, it is not possible to solve generally competition cases by carrying out economic analysis of all positive and negative effects on a case-by-case basis. Therefore also the proponents of the "more economic approach" claim the necessity for developing structured tests with robust rules. A similar development can be observed in the US discussion. Also here economic and legal scholars increasingly postulate the importance of developing and applying "structured rules of reason" (e.g., Tom/Pak 2000 and Kolasky 2008b). In this respect, Michael Salinger (Carlton/Salinger 2007: 92-93) states the following: "[…] that the role of economics is to help inform somewhat more formulaic rules. And these are rules that would be based explicitly on a recognition of the risk of error." In a similar fashion, already Frank H. Easterbrook (1984: 39) had defended a 'filter-approach' for screening categories of cases by also referring to high costs of judicial decision-making of the "traditional Rule of Reason" and saying that "courts should establish rules, recognizing that one cost of decision by rule is occasional over- and under-breadth." Importantly, such a structured rule of reason approach also addresses the sequence of antitrust issues (Beckner/Salop 1999: 67). The structuring is partly also justified by the uncertainties and problems of the US court system (esp. the juries), which might be alleviated, if clearer rules for the assessment of business behavior would exist (Kolasky 2008a). In any case, the problem of uncertainty and decision errors in competition law is an increasingly important topic on both sides of the Atlantic, and there is a growing insight that the analysis of the intermediate solutions between per-se rules and a full-scale rule of reason are the most promising perspective for finding suitable solutions for different categories of business behavior.[34] However, despite a number of articles that address these issues, so far a systematic theoretical analysis of these questions is still missing. Our approach might be helpful for such a research.

References

Beckner, C. F. and Salop, S. C. (1999): Decision Theory and Antitrust Rules, *Antitrust Law Journal* 67 (1), 41-76.

Blackwell, D. (1953): Equivalent Comparison of Experiments, *The Annals of Mathematical Statistics* 24 (2), 265-272.

Blackwell, D. and Girshick, M. A. (1966): Theory of Games and Statistical Decisions, New York: Wiley.

Carlton, D. W. and Perloff, J. M. (2005): Modern Industrial Organization, Fourth Edition, Boston: Addison Wesley.

Carlton, D. W. and Heyer, K. (2007): The Year in Review: Economics at the Antitrust Division, 2006-2007, *Review of Industrial Organization* 31 (2), 121-137.

Carlton, D. and Salinger, M. (2007): Economic Analysis of Competition Practices in the EU and the U.S.: A View from Chief Economists, *Competition Policy International* 3 (2), 81-98.

Carrier, M. A. (1999): The Real Rule of Reason: Bridging the Disconnect, *Brigham Young University Law Review* 22 (4), 1265-1365.

Cavanagh, E. D. (2008): Vertical Price Restraints After Leegin, *Working Paper ExpressO*,

[34] See, e.g., Beckner/Salop (1999), Cavanagh (2008), Heyer (2005), Klein (1996), Kolasky (2008b), Leary (2004), Lemley/Leslie (2007), Piraino (1991) and Tom/Pak (2000)

http://works.bepress.com/cgi/viewcontent.cgi?article=1001&context=edward_cavanagh

Christiansen, A. and Kerber, W. (2006): Competition Policy with Optimally Differentiated Rules Instead of "Per se Rules vs. Rule of Reason", *Journal of Competition Law and Economics* 2 (2), 215-244.

Crane, D. A. (2007): Rules Versus Standards in Antitrust Adjudication, *Washington and Lee Law Review* 64, 49-110.

DeGroot, M. H. (1961): Uncertainty, Information and Sequential Experiments, *The Annals of Mathematical Statistics* 33 (2), 404-419.

DeGroot, M. H. (1970): Optimal statistical decisions, New York: McGraw-Hill.

Easterbrook, F.H. (1984): The Limits of Antitrust, *Texas Law Review* 63 (1), 1-40.

--- (1992): Ignorance and Antitrust, in: Jorde, T.M. and D.J. Teece (eds.), *Antitrust, Innovation, and Competitiveness*, New York: Oxford University Press, 119-136.

Ehrlich, I. and Posner, R. A. (1974): An Economic Analysis of Legal Rulemaking, *Journal of Legal Studies* 3, 257-286.

Gomèz-Pomar, F. and Wangenheim, G.v. (2008): The Complex Economics of Nuisance Law: Bargaining under Property and Liability Rules with Public Law Standards, *Working Paper*.

Gorry, G. A. and Barnett, G. O. (1968): Experience with a Model of Sequential Diagnosis, *Computers and Biomedical Research* 1, 490-507.

Hay, G. A. (2007): The Quiet Revolution in U.S. Antitrust Law, *The University of Queensland Law Journal* 26 (1), 27-37.

Heckerman, D.; Horvitz, E. and Middleton, B. (1991): An Approximate Nonmyopic Computation for Value of Information, *Working Paper*.

Heckerman, D. E.; Horvitz, E. J. and Nathwani, B. N. (1992): Toward Normative Expert Systems: Part I, The Pathfinder Project, *Methods of Information in Medicine* 31 (2), 90-105.

Heyer, K. (2005): A World of Uncertainty: Economics and the Globalization of Antitrust, *Antitrust Law Journal* 72 (2), 375-422.

Hirshleifer, J. and Riley, J. G. (1992): The analytics of uncertainty and information, Cambridge: Cambridge University Press.

Hovenkamp, H. (2001): Post-Chicago Antitrust: A Review and Critique, *Columbia Business Law Review* 2001 (2), 257-338.

Joskow, P. L. (2002): Transaction Cost Economics, Antitrust Rules, and Remedies, *Journal of Law, Economics, and Organization* 18 (1), 95-116.

Kalagnanam, J, and Henrion, M. (1990): A Comparison of Decision Analysis and Expert Rules for Sequential Diagnosis, *Uncertainty in Artificial Intelligence* 4, 271-281.

Kaplow, L. (1992): Rules Versus Standards: An Economic Analysis, *Duke Law Journal* 42, 557-629.

Kaplow, L. (1995): A Model of the Optimal Complexity of Legal Rules, *Journal of Law, Economics, and Organization* 11(1), 150-163.

Katz, M. L. and Shelanski, H. A. (2007): Merger Analysis and the Treatment of Uncertainty: Should we expect better?, *Antitrust Law Journal* 74 (3), 537-574.

Kerber, W. (2008): "Rules vs. Standards" or Standards as Delegation of Authority for Making (Optimally Differentiated) Rules, in: Eger, Thomas et al. (eds.), Internalization of the Law and its Economic Analysis, Wiesbaden: Gabler 2008, pp. 489-498.

Kerber, W. and Schwalbe, U. (2008): Economic Principles of Competition Law, in: Hirsch, G., Montag, F., and F.-J. Säcker (eds.), Competition Law: European Community Practice and Procedure - Article-by-article Commentary, London: Sweet & Maxwell 2008, pp. 202-393.

Klein, J. (1996): A Stepwise Approach to Antitrust Review of Horizontal Agreements,

Department of Justice, Before the American Bar Association's Antitrust Section Semi-Annual Fall Policy Program, 1-19.

Kneepkens, M. (2007): Resale Price Maintenance: Economics Call for a More Balanced Approach, *European Competition Law Review* 12, 656-664.

Kolasky, W. (2008a): Different roads, *Global Competition Review*, 17-21.

--- (2008b): Reinvigorating Antitrust Enforcement in the United States: A Proposal, *Antitrust* 22 (2), 85-90.

Leary, T. B. (2004): A Structured Outline for the Analysis of Horizontal Agreements, *The Conference Board 2004 Antitrust Conference, Working Paper*, 1-18.

Lemley, M. A. and Leslie, C. R. (2007): Categorical Analysis in Antitrust Jurisprudence, *Stanford Law School Working Paper*, No. 348.

Lewis, C. and Sheehan, K. (1990): Using Bayesian Decision Theory to Design a Computerized Mastery Test, *Applied Psychological Measurement* 14 (4), 367-386.

Lindgren, B. W.; McElrath, G. W. and Berry, D. A. (1978): Introduction to Probability and Statistics, 4. ed., New York: Macmillan.

Marschak, J. and Miyasawa, K. (1968): Economic Comparability of Information Systems, *International Economic Review* 9 (2), 137-174.

Mathewson, F. and Winter, R. (1998): The Law and Economics of Resale Price Maintenance, *Review of Industrial Organization* 13 (1-2), 57-84.

Mookerjee, V. S. and Dos Santos, B. L. (1993): Inductive Expert System Design: Maximizing System Value, *Information Systems Research* 4 (2), 111-140.

Mookerjee, V. S. and Mannino, M. V. (1997): Sequential Decision Models for Expert System Optimization, *IEEE Transactions on Knowledge and Data Engineering* 9 (5), 675-687.

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

Mussi, S. (2002): Sequential decision-theoretic models and expert systems, *Expert Systems* 19 (2), 99-108.

Piraino (1991): Reconciling the Per Se Rule and Rule of Reason Approaches to Antitrust Analysis, *Southern California Law Review* 64, 685-739.

Rey, P. and Caballero-Sanz, F. (1996): The Policy Implications of the Economic Analysis of Vertical Restraints, Economic Papers #119, European Commission, DG Economic and Financial Affairs, Belgium: Bruxelles.

Röller, L.-H. (2005): Economic Analysis and Competition Policy Enforcement in Europe, in: van Bergeijk, P. A. G. and Kloosterhuis, E. (eds.): *Modelling European Mergers. Theory, Competition Policy and Case Studies*, Cheltenham: Elgar, 11-24.

Roth, W.-H. (2007): The 'More Economic Approach' and the Pluralist Tradition of European Competition Law (Comment), in: Schmidtchen, D.; Albert, M. and Voigt, S. (eds.): *The More Economic Approach to European Competition Law*, Tübingen: Mohr Siebeck, 37-57.

Scherer, F. M. and Ross, D. (1990): Industrial Market Structure and Economic Performance, Third edition, Boston: Houghton Mifflin.

Tom, W. K. and Pak, C. (2000): Toward a Flexible Rule of Reason, *Antitrust Law Journal* 68 (2), 391-428.

Tseng, C.-C. and Gmytrasiewicz, P. J. (1999): Time Sensitive Sequential Myopic Information Gathering, *IEEE Proceedings of the 32nd Hawaii International Conference on System Sciences*.

Van den Bergh, R. J. and Camesasca, P. D. (2006): European Competition Law and Economics: A Comparative Perspective, London: Sweet & Maxwell.

Voigt, S. and Schmidt, A. (2005): Making European Merger Control More Predictable,

Dordrecht: Springer.
Vos, H. J. (2007): A Bayesian sequential procedure for determining the optimal number of interrogatory examples for concept-learning, *Computers in Human Behavior* 23 (1), 609-627.

Appendix A

In this appendix, we prove the proposition in Section 3.3. The 'should' in the proposition comes from the idea that the CA can reduce expected total costs by acting in this way. When discussing the proposition, we assumed stochastically independent signals. The bracketing relations therefore do not depend on the history of signals. We will introduce a simplified notation to keep the proof as clear as possible.

The comparison of a bracketing criterion *i* with any possible criterion *i*' out of *K* must hold for every possible posterior probabilities. Thereby, a higher reduction of expected costs is equivalent to a lower posterior expected cost. This indicates more informativeness. Furthermore, because (i) the bracketing condition is fulfilled for any stage in the decision process, (ii) investigation costs and expected loss additively yield expected cost and (iii) investigation costs are assumed to be equal for every criterion, the investigation costs are not included in the following analysis. We, therefore, can regard expected loss (*EL*) instead of expected cost (*EC*). The proof of more informativeness of an information system compared to other information systems is known as *Blackwell`s Theorem*.[35] It is used here to prove lower expected loss of one (information) criterion compared to any other criterion: if the posteriors of criterion *i*' are derived by a stochastic transformation of the posteriors of another criterion *i*, then criterion *i* is more informative than criterion *i*' and, hence, the expected loss of *i* is lower than the expected loss of *i*'. Therefore, the posteriors of *w* of criterion *i*' can be expressed as a convex combination of the posteriors of *w* of criterion *i*, regarding both signals. Starting point of the analysis is any stage *j* in the investigation process with the given history of signals h_j.

The proof is organized as follows: the first part derives conditions, which hold for the unconditional signal-probabilities if the bracketing condition is fulfilled; the second part then proves the resulting lower expected loss of criterion *i*.

Obviously, the expected loss may only be reduced by investigation if the terminating decision depends on the signal for both criteria.[36] Assuming without loss of generality that the signal $m = +1$ ($m = -1$) is more conclusive to the occurrence of $w = B$ ($w = H$), this implies that

$$\text{(A1)} \quad \hat{p}_i(+1) > \hat{p}_{i'}(+1) \geq \frac{L_H}{L_H + L_B} > \hat{p}_{i'}(-1) > \hat{p}_i(-1)$$

[35] See Blackwell (1953), Blackwell/Girshick (1966: 324 ff.) and Hirshleifer/Riley (1992: 188 ff.)
[36] See DeGroot (1961: 405-406), e.g., for the generally non-negative information value.

by equation (1) and the bracketing condition, where $\hat{p}_i(+1)$ and $\hat{p}_i(-1)$ abbreviate $\hat{p}\left(h_j \cap (i,+1)\right)$ and $\hat{p}\left(h_j \cap (i,-1)\right)$, respectively, and the corresponding abbreviations are used for i'. We will use these abbreviations in the rest of this appendix. From (A1) we know that $EL_i^*(+1) = L_H \hat{p}_i(+1)$ and $EL_i^*(-1) = L_B \hat{p}_i(-1)$, and correspondingly for i'. The analysis follows Hirshleifer/Riley (1992: 188 ff.).

First, the expression of the posteriors of criterion *i*' as a convex combination of the posteriors of criterion *i* ('bracketing condition'), is given by:

(A2) $\hat{p}_i(+1)a_+ + \hat{p}_i(-1)(1-a_+) = \hat{p}_{i'}(+1)$ and $\hat{p}_i(+1)a_- + \hat{p}_i(-1)(1-a_-) = \hat{p}_{i'}(-1)$.

The welfare effect probabilities at any stage *j* can always be expressed as the unconditional-signal-probability-weighted average of the corresponding posteriors of any criterion at the next stage, i.e.

$$\hat{p}_i(+1)q_i + \hat{p}_i(-1)(1-q_i) = \hat{p}\left(h_j\right) = \hat{p}_{i'}(+1)q_{i'} + \hat{p}_{i'}(-1)(1-q_{i'}).$$

Inserting (A2) yields

$$\begin{aligned}&\hat{p}_i(+1)q_i + \hat{p}_i(-1)(1-q_i)\\ &\quad = \left(\hat{p}_i(+1)a_+ + \hat{p}_i(-1)(1-a_+)\right)q_{i'} + \left(\hat{p}_i(+1)a_- + \hat{p}_i(-1)(1-a_-)\right)(1-q_{i'})\end{aligned}$$

which simplifies to:

(A3) $q_i = a_+ q_{i'} + a_-(1-q_{i'})$.

Now, posterior expected loss of *i*', given by $EL_{i'}(h_j, d^0) = q_{i'} \cdot EL_{i'}^*(+1) + (1-q_{i'}) \cdot EL_{i'}^*(-1)$, expands to

(A4) $$\begin{aligned}EL_{i'}(h_j, d^0) &= q_{i'} \cdot L_H(1-\hat{p}_{i'})(+1) + (1-q_{i'}) \cdot L_B \hat{p}_{i'}(-1)\\ &\overset{(A2)}{=} q_{i'} L_H\left(\left(1-\hat{p}_i(+1)\right)a_+ + \left(1-\hat{p}_i(-1)\right)(1-a_+)\right) + (1-q_{i'})L_B\left(\hat{p}_i(+1)a_- + \hat{p}_i(-1)(1-a_-)\right)\\ &= q_{i'}\left(a_+\left(1-\hat{p}_i(+1)\right)L_H + (1-a_+)\left(1-\hat{p}_i(-1)\right)L_H\right) + (1-q_{i'})\left(a_-\hat{p}_i(+1)L_B + (1-a_-)\hat{p}_i(-1)L_B\right)\\ &\overset{(A1)}{>} q_{i'}\left(a_+\left(1-\hat{p}_i(+1)\right)L_H + (1-a_+)\hat{p}_i(-1)L_B\right) + (1-q_{i'})\left(a_-\left(1-\hat{p}_i(+1)\right)L_H + (1-a_-)\hat{p}_i(-1)L_B\right)\\ &= \left(q_{i'}a_+ + (1-q_{i'})a_-\right)\left(1-\hat{p}_i(+1)\right)L_H + \left(q_{i'} + 1 - q_{i'} - q_{i'}a_+ - (1-q_{i'})a_-\right)\hat{p}_i(-1)L_B\\ &\overset{(A3)}{=} q_i\left(1-\hat{p}_i(+1)\right)L_H + (1-q_i)\hat{p}_i(-1)L_B\\ &= EL_i(h_j, d^0)\end{aligned}$$

Q.E.D.

Appendix B

In this appendix we show that the myopic approach may fail to find the optimal sequencing of the assessment criteria when the bracketing condition is (at some point of the investigation) violated.

To simplify the calculation, the number of criteria is restricted to $K = 3$. It is further assumed that criterion z_1 brackets the other two criteria. Therefore, criterion 1 is more informative and should be investigated first. The other two criteria do not satisfy the bracketing condition. Let the likelihoods that criterion *i* yields a positive signal be given by the following table:

Criterion	z_1	z_2	z_3
true state of the world = B	0,82	0,71	0,68
true state of the world = H	0,18	0,4	0,35

Then the optimal decision tree may be derived by trying all four remaining possible orderings (criterion 1 is the first, due to bracketing, criteria z_2 and z_3 may be first or second both on the upper and the lower branch). Doing so, we derived the optimal tree as given in figure 5.

For the upper branch (first signal is $m_1 = +1$), the expected costs from investigating exactly one more criterion is given by $0.536 \cdot 0.326 + 0.464 \cdot 0.427 = 0.372864$ if criterion z_3 is the one investigated and $0.574 \cdot 0.345 + 0.426 \cdot 0.423 = 0,378228$ if criterion z_2 is the one investigated. Hence the myopic criterion fails to coincide with the optimal ordering but prescribes the reverse order. The reason is simple: the myopic approach underestimates the information value of criterion z_2, since for one outcome the criterion is not terminal but calls for criterion z_3, which reduces the minimal costs associated to the decision node marked by a bold frame in Figures 4b and 5 from 0.423 (Figure 4b) to 0.398 (Figure 5), i.e. by a difference of 0.025. The same is true for criterion z_3, but here the effect is weaker: the expected costs only decline from 0.427 to 0.417, i.e. by a difference of only 0.01. The difference in total expected costs for the entire decision tree amounts to 0.0002, i.e. by 0.01 per cent.

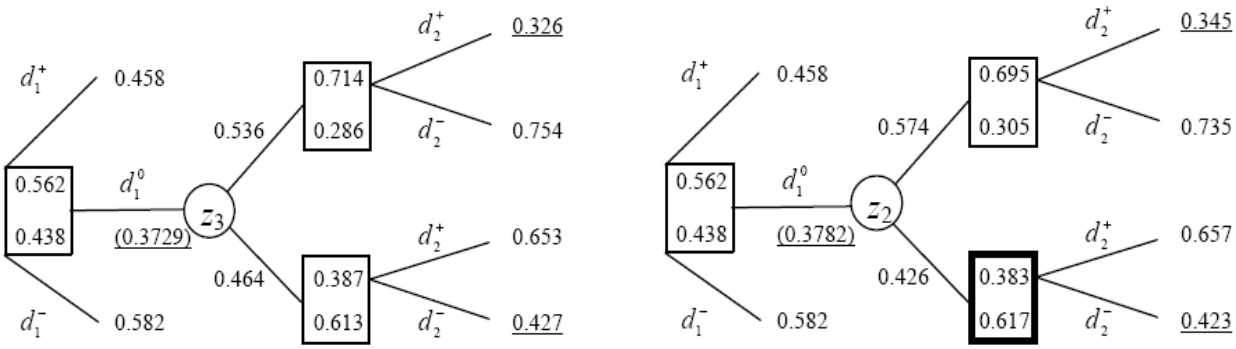


Figure 4: Expected total cost of criterion 3 (left) and criterion 2 (right) at second stage according to the myopic approach (second number in decision node rectangles is $1 - \hat{p}(h_j)$)

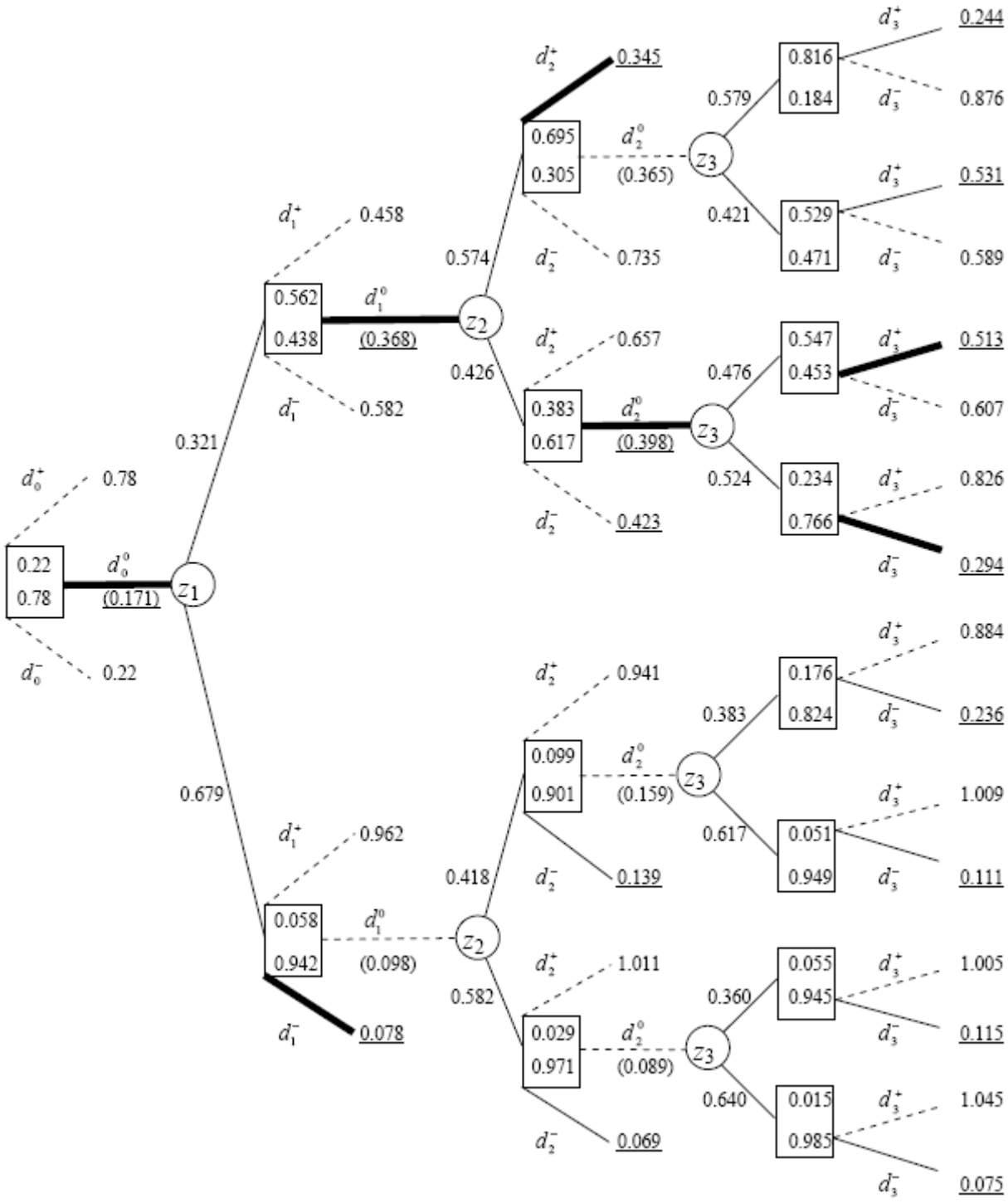


Figure 5: Optimal decision tree (second number in decision node rectangles is $1-\hat{p}\left(h_j\right)$)