Choosing Assessment Procedures in Antitrust Enforcement in Different Markets and Countries: A Review of and a Proposal for Extending the Decision-Theoretic Approach (with Applications to the Digital Platforms and Developing Countries) 1

Yannis Katsoulacos2 and David Ulph3

June 2022

WORK IN PROGRESS - PLEASE DO NOT QUOTE WITHOUT THE AUTHORS’ PERMISSION

Abstract

The appropriate choice of legal standards (LSs) in antitrust enforcement, that is, of the assessment procedures or decision rules, that provide the basis for how assessment of potentially anticompetitive conduct must be undertaken, in order to decide whether there is liability or not, has been hotly debated for many years. The debate has gained in intensity in recent years as a result of the concerns expressed by a significant number of academics and policy makers in many countries with the antitrust treatment of the major platforms.

This article starts with a review of the literature, especially on recent developments in the choice of LSs, to which the authors have contributed (2009, 2016, 2020, 2022). It then provides a detailed presentation of the methodology for defining legal standards along the continuum of legal standards (Arreda and Hovenkamp, 2017; Jones and Kovacic, 2017; Katsoulacos et al., 2020 and 2021; Katsoulacos and Ulph, 2022), depending on the screens or preconditions assessed at different stages of the continuum. This is followed by a detailed formal examination of how all the pertinent factors that could influence error minimisation interact to determine the optimal error-minimising LSs for different conducts and markets. Finally, we use this framework to examine how the choice of error minimising LSs depends on the context in which specific conducts are undertaken. Specifically, how this choice is affected when comparing developing jurisdiction/countries to developed countries / jurisdictions, a topic examined in Bageri and Katsoulacos (2022); and, how the choice is affected when the conduct is undertaken in digital multi-sided platforms.

Keywords: Legal standards, effects-based, Per Se, rule of reason.

JEL Classification: K21, L4.

1 Over the years we have benefitted enormously from discussions on the general issues dealt with by this paper with Svetlana Avdasheva, David Evans, Svetlana Golovanova, Frederic Jenny, Bill Kovacic, Pierre Regibeau, Patrick Rey, Thomas Ross and Jacob Seifert. Of course, all responsibility for errors, omissions and ambiguities lies with me. We would like also to thank for their research assistance, Vasiliki Bageri, Eleni Metisou and Galateia Makri who contributed in the context of the ELIDEK project “Optimal Design of Competition Policy Enforcement”.

2 Emeritus Professor, Athens University of Economics and Business. Affiliated Chair Professor Jiangxi University of Finance and Economics. Former Commissioner of Hellenic Competition Commission. Email: yanniskatsoulacos@gmail.com

3 Emeritus Professor, University of St. Andrews; Member of UK Competition Appeals Tribunal. Email:du1@st-andrews.ac.uk
1. Introduction: factors influencing the choice of legal standards

The appropriate choice of legal standards (LSs) in antitrust enforcement, that is, of the assessment procedures or decision rules that provide the basis for how assessment of potentially anticompetitive conduct must be undertaken in order to decide whether there is liability or not, has been hotly debated for many years. How widely divergent the opinions have been in this debate and how dominant specific points of view become, in terms of their influence on enforcement practice, has varied over time and across countries and continents. Broadly speaking, excluding hard-core horizontal agreements, for which there is broad unanimity that their treatment should rely on a strong presumption of illegality, for most other conducts that come under antitrust scrutiny, specifically, vertical restraints, concerted practices and monopolisation, or abuse of dominance, practices, the US (or North America) enforcement practice has differed quite significantly from that in the EU and the EC in particular as well as from other less mature jurisdictions, the latter being at present much closer to the EU than to US. The present article contributes to understanding the reasons for these differences.

The debate on the appropriate choice of LSs has gained in intensity in recent years as a result of the concerns expressed by a significant number of academics and policy makers in many countries with the treatment of the major platforms. Even in the US, an increasing number of commentators have been arguing that the current antitrust doctrines, rules and antitrust enforcement “are too limited to protect competition adequately, making it needlessly difficult to stop anticompetitive conduct in digital markets” and growing market

---

4 We recognise that a distinction is drawn by legal scholars between “rules” (a term that, in the context of antitrust, they reserve for Per Se decision procedures) and “standards” (like the “rule of reason”) – see Blair and Sokol (2012), Jones and Kovacic (2017) and for a very recent excellent and extensive discussion (and references) Kovacic (2021). As e.g. Blair and Sokol (2012, p. 472) write “The rule of reason involves a more open-ended inquiry than that of a per se analysis, moving antitrust away from rules and toward a standard”. Also, can see Araiza (2011) for a discussion extending beyond antitrust. Below, for simplicity, we neglect this terminological distinction and refer to all the “decision procedures” (which might be the most appropriate term for economists) that we discuss and compare (including the Per Se rule) as legal “standards”.

5 In the error mimising framework we develop below, we can use the terms “legal standard” and “standard of proof” interchangeably: determining the error-minimising LS essentially means determining the error-minimising standard of proof. See also below. Our discussion has potential implications also for the allocation of the “burden of proof” though we do not delve into these here.

6 That is, the LS should be one of Per Se Illegality (in US) or by-object restriction (in EU). Though we recognise that these are not exactly equivalent LSs - see for an extensive discussion on this Katsoulacos and Makri, 2020 – for our purposes here they can be treated for much of the discussion as if they are, so below we will not distinguish between them. There are also some conduct (e.g. refusal to license know-how) for which there is broad agreement that they should be treated under Per Se Legality.

7 Below we will refer to them as developing jurisdictions or developing countries for convenience. Broadly speaking we have in mind jurisdictions that have developed the relevant laws and institutions and have been active in antitrust enforcement for at most the last 20 years. However, we also have in mind here countries that are “developing” in the sense of their general economic, technological, political and socio-cultural conditions and characteristics. Details can be found in Bageri and Katsoulacos (2022), including a description of a methodology developed for EBRD to construct a composite indicator of “development” in the above sense by Yannis Katsoulacos and Frederic Jenny, using a number of indices available from various international organisations, specifically: WB, WEF, OECD, UNCTAD, IMF, EBRD and WTO.
power (Baker et.al. 2020)\textsuperscript{8}. More generally, it has been argued, for US, that, “as a result of unsound economic theories and unsupported empirical claims about the competition effects of certain practices....antitrust rules constructed by the courts reflect a systematically skewed error cost-balance\textsuperscript{9}; they are too concerned to avoid chilling procompetitive conduct and the high cost of litigation, and too dismissive of the cost of failing to deter harmful conduct”. Also, they have “encouraged overly cautious enforcement policies and overly demanding proof requirements and have discouraged government enforcers and private plaintiffs from bringing meritorious exclusionary conduct cases”\textsuperscript{10}.

Below, our objective is to undertake a rigorous application of the error minimising approach to analyse these claims and, especially, whether more presumption-based illegality LSs would be more suitable for the antitrust (\textit{ex post}) assessment of potentially anti-competitive conduct by the major platforms rather than full blown rule of reason. We stress here that it is not our objective to analyse the large number of reports and policy papers that have argued in recent years that the key features of digital platform markets warrant also \textit{ante} regulatory action, in addition to \textit{ex-post} antitrust enforcement in view of their significant influence on on markets and society at large\textsuperscript{11}.

\textit{Factors influencing the choice of LSs: an outline}

\textsuperscript{8} “Joint Respose to the House Judiciary Committee on the State of Antitrust Law and implications for Protecting Competition in Digital Markets” by 12 of the most prominent economists and legal experts in US.

\textsuperscript{9} For a very systematic and extensive criticism of the view that the primary objective in antitrust enforcement is to limit false convictions rather than false acquitals, that has its origins in Easterbrook (1984), see Hovenkamp (2021) – that also contains many references to opposing views. Gavil and Salop (2020) and Baker (2015) are also very critical. Gavil and Salop (2020) point out that “Many of the assumptions that guided this generation-long retrenchment of antitrust rules were mistaken, and advances in the law and in economic analysis have rendered them anachronistic. This is especially the case with respect to exclusionary conduct” (p. 6).

\textsuperscript{10} Baker et.al (2020; p. 4-5). This situation has “been defended with reference to mistaken and unjustified assumptions – including erroneous claims that markets self-correct quickly, monopolies best promote innovation, firms with monopoly power can obtain only a single monopoly profit, vertical restraints...... almost invariably benefit competition even in oligopoly markets, courts and enforcers are manipulated by complaining competitors, and courts cannot tell whether exclusionary conduct harms competition or benefits it” (p. 5). The authors go on to devote a distinct section on legal rules. See also discussion below of the work of other commentators.

\textsuperscript{11} For a good review see the OECD Report of 2 December 2021. Among the main reports we mention Cremer et.al. (2019), Furman et.al. (2019) and Stigler Committee (2019). Another excellent discussion with interesting concrete proposals is that of Rogerson and Shelanski (2020). They note that: “...light handed pro-competitive” (LHPC) regulation could increase levels of competition in markets served by digital platforms while helping to clarify the platforms’ obligations with respect to interrelated policy objectives, notably privacy and data security. Key categories of LHPC regulation could include interconnection/interoperability requirements (such as access to application programming interfaces (APIs)), limits on discrimination, both user-side and third-party-side data portability rules, and perhaps additional restrictions on certain business practices subject to rule of reason analysis under general antitrust statutes. These types of regulations would limit the ability of dominant digital platforms to leverage their market power into related markets or insulate their installed base from competition. In so doing, they would preserve incentives for innovation by firms in related markets, increase the competitive impact of existing competitors, and reduce barriers to entry for nascent firms. See Also Caffara et.al. (2021) and Caffara and Scot Morton (2021).
A large number of broad considerations influence the choice of LSs and have been the subject of an extensive literature. The most important are: the desire to minimise decision errors\textsuperscript{12}; the desire to minimise implementation / enforcement costs\textsuperscript{13}; the deterrence effects and the legal uncertainty effects of different LSs\textsuperscript{14}; reputational concerns of the Competition Authorities (CAs); the substantive (or liability) standards applied\textsuperscript{15}. The first four considerations are encapsulated in the so-called normative or welfare maximising approach to the choice of LSs (Katsoulacos and Ulph, 2009, 2015, 2016, 2020). Reputational concerns can be important, given that when decisions are reached by different LSs, they encapsulate economic analysis to a different extent and degree of sophistication, and thus may be treated differently by Appeal Courts, leading to annulment rates of decisions that differ depending on the LS used – decision annulment influencing negatively the reputation of CAs (Avdasheva et.al., 2019; Katsoulacos, 2019b). Finally, the adoption of non-welfarist substantive standards\textsuperscript{16} leads to optimal LSs that are closer to Per Se (Katsoulacos, 2019a).

In this paper, we focus on the consideration that has had the greatest influence on thinking in this area and that has been discussed most extensively and for a longer period than all others: the desire to minimise the welfare costs of decision errors\textsuperscript{17} (see, Easterbrook, 1984; Beckner and Salop, 1999; Hylton and Salinger, 2001; Evans and Padilla, 2005; Katsoulacos and Ulph, 2009 and for a very recent authoritative non-technical review applied to exclusionary conduct, by Gavil and Salop, 2020\textsuperscript{18}). As the latter note:

“...It has been recognized for decades that decision theory is useful for understanding and formulating legal standards. Making legal decisions based on probability, inferences, and presumptions is consistent with a decision-theoretic approach to legal rules. Decision theory provides a methodology for information-gathering and decision-making when outcomes are uncertain, information is inherently imperfect, and information is costly to

\textsuperscript{12} And, hence, on the factors that influence decision errors, on which our analysis here is dedicated. See below for additional discussion and references.

\textsuperscript{13} Grant and Sanghvi (2021) focus on these. They consider “the per se rule a profoundly economic approach to the problem that the demand for judicial resources exceeds its supply” (p. 99), recognising however that “Administrative convenience alone is not enough to justify the per se rule”, as the Leegin (2007) decision, to which they point out, shows (footnote 5), as well as the multitude of the other cases in which per se has been abandoned in US (and many other countries) – see also Kovacic (2021). Clearly, the welfare cost of decision errors and implementation costs are two sides of the same coin: choosing the most suitable LS must take both into account. On the other hand, it is important to stress that in many cases (hard-core horizontal cartels been the most obvious) Per Se rules minimise decision error costs – the existing analyses on decision errors and the one presented here explain exactly when this is the case (abstracting from implementation costs).

\textsuperscript{14} Easterbrook (1984); Lemley and Leslie (2008). See also, however, Katsoulacos and Ulph (2015 and 2016) who distinguish between different types of legal uncertainty and show that under many circumstances the rule of reason remains superior.

\textsuperscript{15} See below for more details.

\textsuperscript{16} Such as “protecting the competitive process” or “non-disadvantaging rivals”. See for more details below and Katsoulacos (2019a).

\textsuperscript{17} More precisely, in the words of Beckner and Salop (1999), “minimising the expected consumer welfare costs of erroneous decisions” (p. 50). For early applications of this error-cost approach to legal rules see Ehrlich and Posner (1974) and Posner (1973).

\textsuperscript{18} Especially section II. They provide a number of references to earlier work in decision theory and optimal statistical decisions in footnote 62, p. 16.
obtain. This methodology is a rational process in which a decision-maker begins with initial beliefs (i.e., presumptions) based on prior knowledge and then gathers additional information (i.e., evidence) to supplement the presumption in order to make a better, more accurate decision” (p. 16).

In particular, Katsoulacos and Ulph (2009), extended by their 2016 paper and followed by Seifert (2020), Katsoulacos and Ulph (2020) and, especially, Katsoulacos and Ulph (2022), provide models\(^{20}\) that examine all the factors that a CA or a court must take into account and derive simple representations, in terms of conditions expressed by simple formulae, of exactly the way that these factors interact and influence the error-minimising choice of LSs, and whether /when / how much additional information should be gathered when examining specific categories of conduct\(^{21}\).

2. The continuum of legal standards

The starting point in Katsoulacos and Ulph (2022) is the recognition that the task usually facing CAs and Courts is to choose to what extent their assessment should rely on additional distinct economic analyses and information gathering investigations that improve our ability to correctly discriminate between genuinely harmful and benign conducts of the same type. They think of the additional assessment tests as lying along a sliding scale or continuum, at the extremes of which are, on the one hand, assessments based purely on presumptions (namely, the Strict Per Se rule, that relies on just the characterization of the conduct) and, at the other, assessments based on the findings of all potential case-specific economic analyses and tests of the market(s) that could influence the conduct’s impact (full Effects-based, or rule of reason\(^{22}\)). The idea that “the modes of antitrust analysis represent a continuum, or “sliding scale” with different fact finding requirements for different situations” was initially developed in the Antitrust Law treatise of Areeda and Hovenkamp\(^{23}\). This idea’s articulation, that best represents the approach in Katsoulacos and Ulph (2022) and in this paper is that of Jones and Kovacic (2017). As they note “the general progression in U.S. doctrine has been toward recognition of an analytical continuum whose boundaries are set, respectively, by categorical rules of condemnation (per se illegality) or acquittal (per se legality) and an elaborate, fact-intensive assessment of reasonableness (Rule of Reason). These poles are

\(^{19}\) Our 2022 working paper contains a more detailed description of what is presented in the current paper. We build on the seminal contribution of Breckner and Salop (1999), and the papers of Hylton and Salinger (2001) and of Evans and Padilla (2005).

\(^{20}\) The models in Katsoulacos and Ulph (2022) are preliminary versions of the model developed below.

\(^{21}\) We do not think that the error-minimising framework should be applied on a case-by-case basis – see also discussion in Section 6 below. The categories we refer to are for example, tying, bundling, exclusive dealing, royalty rebates, refusal to deal, margin squeeze, etc.

\(^{22}\) We will use the terms “effects-based” (popular in Europe, also as “economics-based”) and “rule-of-reason” (used in US) interchangeably though, as has been pointed out, Vickers (2007), under the latter there is greater discretion afforded to an agency / court than under the former. Intermediate LSs are described in detail below.

\(^{23}\) 4th Edition, 2017. See also the detailed discussion in Hovencamp (2018); as noted there this was an idea discussed in all three previous editions of the Areeda and Hovencamp treatise, p.123).
connected by a range of intermediate tests that seek to combine some of the clarity and economy of bright-line rules with the greater analytical accuracy that a fuller examination of evidence can produce."\(^{24}\)

In Katsoulacos and Ulph (2022), *the continuum is described by a sequence of steps or stages, in each of which additional screens are examined, using further blocks or components of economic analysis, generating additional information, building on the information already gathered in previous steps*. The question examined is when it is best to add additional steps of economic analysis and hence move to a LS closer to full effects-based. The objective of each step of the information gathering and analysis process is to examine whether certain *preconditions or screens* are satisfied that are considered necessary for demonstrating liability (welfare harm) – such as significant extant market power, potential for exclusion, potential for consumer harm and potential for efficiencies. Then, decision error costs across steps or stages can be compared in order to determine the optimal number of stages, that defines the error-minimising LS. Or, to use the term some authors prefer to use, defines the error-minimising *standard of proof*.

To appreciate the usefulness of this approach, one could for example think its application for comparing whether, when assessing tying arrangements, a Modified Per Se Illegality LS, under which we rely, in order to reach a decision, on certain contextualisation tests and the existence of significant market power, is preferable (in terms of decision errors) than strict Per Se Illegality under which there is no pre-requirement of extant market power. Or, whether a Disadvantaging Rivals (truncated effects-based) LS is preferable to Modified Per Se Illegality – where, under the former, for illegality, significant market power is not enough, it is also required to demonstrate that rivals are likely to be excluded (in a broad sense) from the market by the conduct. Or, whether a full effects-based is preferable to the Disadvantaging Rivals LS. As noted by Evans and Padilla (2005), first, Strict Per Se and then later Modified Per Se Illegality have been the standards favoured for tying by both US and EU jurisdictions until about the end of the 1990s\(^{25}\) and since then it has been decided to move to LSs closer to effects-based.

To give another example, the approach can be used to clarify and make precise why it may make sense to recommend that antitrust laws should be updated in order “to recognise that under some circumstances conduct that creates a risk of substantial harm should be unlawful

---

\(^{24}\) Also, Kovacic (2021), Gavil and Salop (2020; p. 3, also referring to Gavil, 2012), Gavil (2008), p. 139 and Italianer (2013, p. 2), referring to Justice Stevens who was one of the first to point out that one should think of legal standards (for dealing with restraints under US Section 1) as forming a *continuum* with Per Se and Rule of Reason being at the opposite ends of this *continuum* (on Judge Stevens see also Azaira, 2019, who notes that “Justice Stevens has suggested that a judge better performs her role by paying careful attention to facts and context, as opposed to unthinkingly applying rigid legal rules”). As Italianer notes, the US Supreme Court has explicitly recognized that “the categories of analysis cannot pigeonholed into terms like “per se” or... “rule of reason”. No categorical line can be drawn between them. Instead, what is required is a situational analysis moving along what the Court referred to as a “sliding scale””.

\(^{25}\) Under this LS, tying is presumed to violate the law (i.e. it is considered presumptively illegal) when undertaken by dominant firms. See also Ahlborn et.al (2004) and Evans et.al. (2006).
even if the harm cannot be shown to be more likely than not." Or, it may be used to examine how the choice of error minimising LSs depends on the context in which specific conducts are undertaken. Below, we discuss, in sections 5 and 6, how this choice is affected when comparing developing jurisdiction/countries (as we defined these countries above) to developed countries / jurisdictions, a topic examined in Bager and Katsoulacos (2022); and, how the choice is affected when the conduct is undertaken in digital multi-sided platforms. We start, in section 3 below with a detailed presentation of the methodology for defining legal standards along the continuum depending on the screens or preconditions assessed. This is followed in section 4 by a detailed formal examination of how all the pertinent factors interact to determine the optimal error-minimising LSs for different conducts and markets.

3. Potential investigation stages: defining LSs along the continuum depending on the screens that have been assessed

As indicated above, a CA can decide whether a conduct violates competition law by undertaking one or more investigations, in each of which it successively examines a screen or precondition for identifying harm. Assuming here that the objective is to identify whether the conduct is harmful to consumer welfare the CA could consider that this objective has been satisfied, depending on the type of conduct investigated, in a number of ways, differing in terms of whether some or all and which screens are examined. Specifically, to reach a liability decision, the CA’s investigations can cover one or more of the following stages each of which is associated with the examination of a specific screen or precondition:

Stage 0: Initial characterisation of the conduct. This includes a detailed examination of all the relevant features of the conduct with a focus on those features that according to case law and established economic theory are considered most likely to influence the effects of the conduct. This conduct examination is often accompanied by a description of some basic market magnitudes such as the level of sales, that are an input to stage 2 and can be also considered as been part of that stage. We can refer to this as the “conduct characterisation screen”.

Stage 1: Detailed contextualisation of the market(s) and, most importantly, establishing that there is Significant Market Power (SMP, lack of contestability, or Dominance). We can refer to this as the “market contextualisation and SMP screen”.

Stage 2: Establishing that there is potential for significant exclusionary impact, or, more generally, a competition lessening effect (by enhancing ability to exercise market power or incentives to coordinate behaviour achieving collusive outcomes). This can be manifested through the exit of a rival or rivals or through the marginalisation of rivals (so that they cannot exploit economies of scale and/or network effects) or through the exclusion of potential rivals.

---

26 Recommendation of group of US experts (Baker et. al., 2020; p.1) to Joint Judiciary Committee.

27 See footnote 5 for details.

28 If, as in the tying cases, there are more than one markets to consider, market power must be established in the tying and the tied market.
entrants or through concerted practices. We can refer to this as the “enhanced ability to exercise market power screen”.

**Stage 3**: Establishing that there is potential for consumer welfare loss before accounting for efficiencies. Salop (2017) provides an extensive discussion of how for many of the practices usually considered under Abuse of Dominance (AoD), exclusionary potential may or may not be associated with consumer welfare harm\(^{29}\). Ideally, consumer welfare should be evaluated in terms of effects on prices, on output, on consumer choice (product variety), on quality and on innovation\(^{30}\). We can refer to this as the “potential consumer welfare loss due to anticompetitive effects screen”.

**Stage 4**: Establishing lack of potential for significant efficiencies that can benefit consumers, specifically, establishing that *efficiencies are not sufficiently significant to outweigh the anticompetitive effect of the conduct*. We can refer to this as the “efficiencies and balancing screen”.

Depending on the screens, \(i = 0, 1, \ldots, 4\), examined we can then distinguish the following legal standards.

1. **Strict Per Se (SPS) LS**, is the LS under which the liability decision relies purely on the initial characterisation of the conduct (in stage 0) and the presumption that this generates about its welfare impact.

2. **Modified Per Se LS (MPS LS)**: under this, a liability decision relies just on the information from stages 0 and 1 and the presumption that this generates about its welfare impact.

3. **Truncated Effects Based I LS (TEB I LS)**: under this, a liability decision relies on the information from stages 0, 1 and 2 and the presumption that this generates about its welfare impact. The US Quick Look LS can be considered as an intermediate LS between MPS and TEB I involving a “quick look” on the efficiency defense\(^{31}\).

4. **Truncated Effects Based II LS (TEB II LS)**: under this, a liability decision relies on the information from stages 0, 1, 2 and 3 and the presumption that this generates about its final welfare impact.

5. **Full Effects Based (or rule of reason) LS (FEB LS)**: under this, a liability decision relies on the information from all assessment stages 0 – 4 and a balancing between anticompetitive and efficiency effects to determine the net effect on consumer welfare.

---

\(^{29}\) As noted we are assuming that the substantive or liability standard is one of consumer welfare. With a total welfare standard an additional investigation stage would be added. See also Katsoulacos (2019a).

\(^{30}\) Concentrating on consumer choice may mean reaching decisions on the basis of effects on “competitors”, the exclusion of which may reduce consumer choice. This would be wrong however since there can well be an increase in consumer welfare even with less consumer choice.

\(^{31}\) Under this, a liability decision relies just on the information from stages 0 and 1 and sometimes on the effect to competitors assessed in stage 3, on the basis of which anticompetitive effect is inferred. This term is used essentially in discussions of US enforcement and it signifies that the Court reviews also (has a quick-look) on the efficiency defense presented by defendants (see Harrington, 2020; Hovenkamp, considers this LS as problematic and argues that it has rarely been used, 2018 p. 122-131).
We note that LSs (i) – (iv) are all what can be termed presumption-based LSs, in the sense that they all rely on some presumption about the outcome of subsequent assessment(s), were one or more subsequent assessments made. Only in case (v) the liability decision relies on case-specific information from all assessment steps (0 – 4). So the distinguishing characteristic of this LS is that there is no reliance on presumptions when the liability decision is made.

**Presumption of legality and illegality**

Clearly, for all presumption-based LSs there can be either a presumption of illegality (that is, a presumption that the conduct is on average harmful) or a presumption of legality (that is, a presumption that the conduct is on average benign). To clarify, consider stage 0: in this stage the LS can be that of Strict (or, for simplicity, let us just say, omitting the word “strict”) Per Se Illegality if, just on the basis of the information collected in this stage, the conduct is considered presumptively illegal; or, the LS can be that of Per Se Legality if, just on the basis of the information collected in this stage, the conduct is considered presumptively legal.

To determine this, in stage 0, following the CA’s characterisation of the conduct as being, by virtue of its specific formal features, of a particular type, the CA can draw on knowledge of other cases involving this type of conduct, of relevant economic theory and evidence, and the information collected from the complainants and the firm(s) involved in the specific case, in order to come to a view that a fraction $\gamma$, $0 < \gamma < 1$ of such cases are genuinely harmful to consumer welfare, with (average) harm $H > 0$, while the

---

32 E.g. tying of products, engaging in exclusive dealing contracts, offering quantity discounts or fidelity rebates, refusing to deal with a rival firm etc. In each type, the formal characteristics of different cases are likely, of course, to be different.

33 This is what Hylton and Salinger (1999) call the “base rate” probability (p. 60). At this point, we neglect subscript 0 on parameter $\gamma$, indicating that the value of $\gamma$ depends only on stage 0 information. We extend and generalise below.

34 We assume throughout here that the substantive (or liability) standard is that of consumer welfare. This would seem, according to a number of commentators, to be the most appropriate assumption for North America: “In US since the end of 1970s, the Courts have accepted the view that antitrust law is a “consumer welfare prescription” (Jones and Kovacic, 2017; also, Hyman and Kovacic, 2013). But it is worth noting that recently there have been quite a few voices that have argued that this should change, and the emphasis should return to the protection of the competitive process (e.g. Werden & Froeb, 2018 and Wu, 2018). Indeed, Werden (2014) claims that, “commentators either have merely asserted that a welfare standard must be applied or mistakenly claimed that the Supreme Court has endorsed a welfare standard”. In the EU, the weaker substantive standard concerning the impact on competitors or to the «competitive process» has been favored by Courts (see for discussion and references, Katsoulacos, 2019a), though not necessarily the DGCOMP. In developing countries other, public interest objectives are also very important. This will tend to strengthen the argument that effects-based LSs aiming to assess the welfare impact of conduct are not appropriate. For a recent discussion putting forward arguments in favour of the consumer welfare liability standard see Melamed and Petit (2019). They argue that “both the general and platform-specific assaults on the CW standard are misguided, that the CW standard is capable of addressing the economic concerns that critics have raised, and that the proposed alternatives would make things worse—not better”. For a strong critique see Khan (2018). For an extensive review see ABA Report (2020).
remaining fraction are genuinely benign, with (average) benefit $B > 0$. Given this, if the average harm across all cases is $\bar{h}$, the conduct is considered presumptively illegal (PI) if $\bar{h} > 0$ and is considered presumptively legal (PL) if $\bar{h} < 0$. Clearly, knowledge about the values of these parameters need not be very precise in the sense that what a CA actually needs to determine is just whether on average the conduct can be presumed to be harmful or benign. This essentially involves agencies or courts “creating presumptions through experience, to guide their factual investigations and decision making” (Breckner and Salop, 1999; also Gavil and Salop, 2020). CAs or courts have “initial information on the likelihood and magnitude of benefits and harms...(representing) preliminary presumptions for the entire class of similar (conducts) before gathering additional case-specific information”.

Additional remarks on Per Se LSs

Under a Per Se LS, the totality of all conducts, for which a precondition for welfare harm is considered satisfied, are included in a category (e.g. all conducts undertaken by firms considered to be “dominant”, or all conducts with potential for significant exclusionary impact) and the totality of conducts for which the precondition is not considered satisfied in another category, and then the conducts in each category are treated in exactly the same way: those in the first category, are banned; while all conducts in the second category are acquitted. It is important to note that a procedure for reaching liability in this way has two distinct aspects:

1. The first aspect is that there is no attempt to undertake additional investigations that could allow, through the examination of additional harm-conducive preconditions, a finer discrimination between harmful and benign conducts, rather than relying purely on a discrimination based on whether e.g. conducts are undertaken under dominance or not. Thus the Per Se illegality treatment of all conducts in a category, e.g. the category of conducts undertaken under dominance, relies on a presumption that in all additional investigations, following the stage 1 investigation, the preconditions would be satisfied, leading to a finding of welfare harm. In that sense, the term Presumption-Based LS can be used for or instead of, the term Per Se LS. The term Rule of reason or (full) effects-based refers to the LS under which decisions rely on the outcome of the investigation of all the stages – so all the preconditions for welfare harm are examined – and there is no reliance on any presumptions.

2. The second fundamental aspect of a Per Se LS is that, having decided to reach liability decisions following, for example, just the investigation of the precondition in stage 1 and without undertaking all other potential assessment steps, all conducts are treated in

\[ \bar{h} = \gamma H - (1 - \gamma)B. \]

In an adversarial system of enforcement, such as that of the US, estimates of the values of these parameters will be provided by the defendants and the plaintiffs.

Easterbrook (1984) emphasized the importance of presumptions in antitrust inquiries and thought that the open-ended rule-of-reason approach is often impractical - he advocated a more structured rule-of-reason inquiry when a Per Se rule is not used which may be considered closer to the concept of the rule of reason used here. For a recent very useful discussion in the context of applying decision theory see also Gavil and Salop (2020).
exactly the same (non-discriminating) way: banned when the precondition is considered satisfied and acquitted in the other case. Thus, Per Se neglects the fact that, the signals generated by an investigation, given that a precondition is considered satisfied, about whether or not the conduct is or is not harmful are subject to error: depending on the exact characteristics of the conduct and the exact characteristics of, for example, “dominance” a stronger or a weaker signal of harm may be generated, given that some (one or more) of the preconditions for harm may not be examined (in subsequent stages). Specifically, we will say that, if the investigation identifies the firm as “dominant”, this allows us to identify the specific conduct as harmful when it is indeed harmful on average in a fraction \( p_{H,1} < 1 \) of the cases examined and to identify the conduct as benign when it is indeed benign on average in a fraction \( p_{B,1} < 1 \) of the cases. In other words, \( p_{H,1} \) (resp. \( p_{B,1} \)) can be considered as the probability that a genuinely harmful conduct by a dominant firm will be banned (resp. a genuinely benign conduct by a dominant firm acquitted) if a decision is made following the investigation in step 1. This is very important in determining decision error costs from taking decisions in any given specific stage.

What we wish to stress here is that even though we may not wish to undertake a full analysis involving all the investigative/assessment steps (aspect 1 of Per Se LS), understanding that not all conduct for which precondition \( i < N=4 \) is satisfied are harmful, we may not ban all of them, rather we may discriminate between them, banning a fraction \( p_{H,i} \) of them and acquitting a fraction \( p_{B,i} \). Of course, if also screen \( i = N \) is examined and it is satisfied the conduct is certainly harmful so we ban with certainty.

To capture these distinctions, Katsoulacos and Ulph (2022) as in this paper, distinguish between presumption-based LSs and Rule of Reason. For the former we distinguish between the sub-categories of (a) discriminating presumption-based LSs and (b) Per Se, or, non-discriminating presumption-based LSs. To avoid confusion we will use the term Strict Per Se when we refer to the case where liability decisions rely purely on the categorisation involving the initial characterisation of the conducts in stage 0.

---

38 A firm may be assigned to the category of firms that are considered dominant with a 50% market share and with a 90% market share. Often the signal that the conduct of the latter will be harmful will be much stronger.

39 If the last assessment step (N) is undertaken (which implies that all previous ones have been undertaken and all the preconditions in the previous steps were satisfied), and precondition N is also considered satisfied, then the conduct is considered to be certainly harmful (and can be banned without errors). So, in stage N either the precondition N is considered satisfied and the conduct is banned with certainty or precondition N is not considered satisfied and then the conduct is acquitted with certainty. Thus, in stage N, having completed all steps of analysis needed in order to establish welfare harm, the decision does not need any more to rely on presumptions about harm.

40 To complete this discussion, we note that the term Per Se is commonly and rather loosely deserved for the case in which the liability decision is based only on the initial characterisation of the conducts in stage 0. However, in EU the often similarly treated term object-based LS is deserved to categorise and reach decision on conducts on the basis of the initial characterisation and also the initial market contextualisation associated with stage 1. Further, in formal terms no conduct is strictly Per Se Illegal, in the sense that all (including hard-core cartels) are rebuttable under article 101 (3). The closest to a (strict) Per Se LS is that used in US to treat hard-core horizontal cartels, though, as noted by Harrington (2020), in the US too there are always defenses in practice, so “in practice, there does not seem to be much difference between the US and the EU with regard to explicit agreements” (p. 10).
Discriminating vs. non-discriminating presumption-based (PB) LSs

Under a Presumption-based (henceforth, PB) $LS_i$, $i = 1, \ldots, N-1$, liability decisions rely on information from investigations up to step $i$, $0 < i < N-1$, and no further case-specific analysis/investigation is undertaken. From the information already collected by the $i$th and previous investigations a presumption is formed about whether or not the preconditions for welfare harm in the stages after $i$ will be satisfied or not. Thus, when banning a presumptively illegal (PI) conduct in stage $i$, it is presumed that the preconditions after $i$, that must be examined in order to establish harm to welfare, will be satisfied. As noted above, two approaches for reaching a liability decision under PB LSs can be distinguished (for a PI conduct):

PB Non-Discriminating (or Per Se) $LS_i$, henceforth referred to as $PBNDS_i$: under this all PI conducts for which precondition $i$ is considered to be satisfied are banned, though it is recognised that precondition $i$ is identified with errors and that some of the conducts for which the precondition are satisfied are not harmful.

PB Discriminating $LS_i$, henceforth referred to as $PBD_i$: under this, liability decisions are made on the basis of the understanding that if an assessment step shows a precondition $i = 1, \ldots, N-1$ as satisfied, this allows us to identify the specific conduct as harmful when it is indeed harmful on average in a fraction $p_{H,i} < 1$ of the cases and to identify the conduct as benign when it is indeed benign on average in a fraction $p_{B,i} < 1$ of the cases (further details on these probabilities are provided below).

We note here that the choice between a discriminating and a non-discriminating LS makes sense when $\gamma_i < 1$ – the probability that the conduct is considered harmful is less than unity, which will apply in stages $i = 1, \ldots, N-1$. Then, as already mentioned, one can either ignore this and treat (as under Per Se) all conducts for which precondition $i$ is satisfied as one category for which a uniform “ban” decision will be applied (ALL these conducts being presumed harmful), or one can try to discriminate between harmful and benign conduct for which precondition $i$ is satisfied, taking into account that there will be a range of different circumstances under which precondition $i$ will be considered to be satisfied, and this allows us to identify truly harmful conduct in a fraction $p_{H,i} < 1$ of the cases and to identify truly the conduct as benign in a fraction $p_{B,i} < 1$ of the cases.

Under a Rule of Reason or (full) Effects-Based LS, henceforth referred to as $LS_N$, the liability decision relies on case-specific information from all assessment steps $i = 1, \ldots, N$. Given that there are no more investigation steps after step $N$, if the precondition $N$ is considered satisfied for a conduct (and given that all the preconditions examined in the previous steps are also satisfied), the conduct is considered to be certainly harmful i.e. $\gamma_N = 1$, and there are no (false conviction) errors in banning it. If the precondition is considered not satisfied then (as with previous preconditions) it is acquitted with certainty. This means that when, as under RoR (i.e. under $LS_N$) all assessment steps are taken, there are no errors in step $N$ in identifying harmful conducts among those for which precondition $N$ is considered satisfied. That is, $41$

41 We simplify the discussion focusing on PI conducts unless otherwise stated.
applying a discriminatory approach is irrelevant here. But there are still errors in stage N, though not in the decisions regarding harm. Preconditon N may be mistakenly considered not to be satisfied when it is (leading to a false acquittal), and it may be mistakenly considered to be satisfied when it is not (leading to a false conviction).

4. Determining the optimal error-minimising LS for different conducts and markets

4.1 Introduction

In order to determine what is the optimal legal standard for a specific conduct undertaken in a given market we need to determine the investigation stage at which the cost of decision errors (net of enforcement cost) is minimised for this conduct and market. We start by noting that below we measure the DEC associated with each one of the assessment stages i = 0,...,4, as the error costs that would result if liability decisions (to condemn or to acquit) were taken in that stage. So, $DEC_i$, i = 0,...,4, measures the DEC that would result if liability decisions are made on the basis of the information collected up to and including stage i. Of course, investigating the precondition associated with stage i presupposes that the precondition associated with previous steps have been investigated and are considered to be satisfied. Measuring DEC in this way allows us to determine whether an additional assessment step should be undertaken (because it would lower DEC) as well as the optimal LS. Specifically:

(i) if $DEC_1 < DEC_0$, it is optimal to take step 1; otherwise optimal LS is $LS_0$
(ii) if $DEC_2 < DEC_1$, it is optimal to take step 2; otherwise optimal LS is $LS_1$
(iii) if $DEC_3 < DEC_2$, it is optimal to take step 3; otherwise optimal LS is $LS_2$
(iv) if $DEC_4 < DEC_3$, it is optimal to take step 4; otherwise optimal LS is $LS_3$

So, the RoR ($LS_4$) will be the optimal LS if:

$DEC_4 < DEC_3 < DEC_2 < DEC_1 < DEC_0$

Decision errors emerge because:

- it is not possible to determine with certainty whether a precondition for harm to welfare is or is not satisfied;
- when a precondition is considered to be satisfied, unless all the previous investigations have been undertaken, it is not possible to determine with certainty whether the conduct is harmful or benign.

Of course, DEC will be different depending on whether decisions rely or not on presumptions and, when they are, depending on whether or not they are or are not discriminating.

To proceed with modelling DEC, the following parameters must be introduced.

4.2 Parametrization

Having already defined parameters $\gamma$ (or $\gamma_0$), H and B above, we must now also define:

---

42 E.g. digital or non-digital market.
\( y_i, i = 1, \ldots, N \), is the probability that conducts for which precondition \( i \) is genuinely satisfied are genuinely harmful to consumer welfare, given that the first \( i - 1 \) steps have been undertaken and satisfied\(^{43}\). So if, for example \( i = 2 \), \( y_2 \) is the probability that conducts of this general type (as determined in step 0), undertaken by dominant firms (established under \( i = 1 \), that have exclusionary effects (established under \( i = 2 \), lower consumer welfare. Clearly:

\[
y_0 < y_i < y_{i+1} < 1, \quad i = 1, \ldots, N - 1, \quad y_N = 1
\]

(1)

This means that the fraction of genuinely harmful cases in the population of conducts for which precondition \( i \) is satisfied, given that the first \( i - 1 \) steps have been undertaken and satisfied, increases as \( i \) increases. Thus, the fraction of such cases in the population of dominant firms (\( i = 1 \)), will be less than the fraction of such cases for which there is an exclusionary effect (\( i = 2 \), etc.

\( \beta_i < 1, \quad i = 1, \ldots, N \) is the probability that the precondition examined in step \( i \) is genuinely satisfied given all preconditions in the previous assessment steps are satisfied; for example, \( \beta_1 \) is the probability that the conduct is undertaken by a dominant firm, while \( \beta_2 \) is the probability that, when undertaken by dominant firms, this conduct type has exclusionary effects etc.\(^{44}\)

Instead of using the probability that precondition \( i \) is satisfied in the population of conducts for which the previous preconditions are satisfied, it is also useful to define:

\( \hat{\beta}_i < 1, \quad i = 1, \ldots, N \), is the probability that precondition \( i \) is genuinely satisfied in the total population of conducts of the type examined. Clearly:

\[
\hat{\beta}_1 = \beta_1, \quad \hat{\beta}_2 = \beta_1 \beta_2, \quad \hat{\beta}_3 = \beta_1 \beta_2 \beta_3, \quad \hat{\beta}_4 = \beta_1 \beta_2 \beta_3 \beta_4; \quad \hat{\beta}_1 > \hat{\beta}_2 > \hat{\beta}_3 > \hat{\beta}_4
\]

(2)

The extent to which \( \hat{\beta}_i \) falls with \( i \) depends on the type of conduct and market context. If for example, almost all dominant firms’ conduct is very likely to be exclusionary and when exclusionary almost always consumer welfare is reduced then \( \hat{\beta}_1 \approx \hat{\beta}_2 \approx \hat{\beta}_3 \).

Note that:

\[
y_0 = \beta_1 y_1, \quad y_1 = \beta_2 y_2, \quad y_2 = \beta_3 y_3, \quad y_3 = \beta_4 y_4 = \beta_4
\]

(3)

and so:

\[
y_0 = \prod_{i=1}^4 \beta_i = \hat{\beta}_4, \quad y_1 = \prod_{i=2}^4 \beta_i, \quad y_2 = \prod_{i=3}^4 \beta_i, \quad y_3 = \beta_4, \quad y_4 = 1
\]

(4)

and

\(^{43}\) Undertaking step \( i \) implies that the preconditions associated with the previous steps are considered satisfied.

\(^{44}\) We comment on other parts of the paper how the value of these parameters can be assessed. To give an example here, a most important consideration for assessing the value of \( \beta_i \) is whether without significant market power there is no incentive to undertake the conduct. If so (as, for example, in a refusal to deal or an exclusive deal case), \( \beta_1 \) is likely to be very high, otherwise (as for a number of tying practices) it will be low.
\[ \beta_1 y_1 = y_0 = \hat{\beta}_4 < y_1 = \beta_2 y_2 < y_2 = \beta_3 y_3 < y_3 = \beta_4 y_4 = \beta_4 \]  

Finally, it is clear from the above that:

\[ \hat{\beta}_1 (1 - y_1) > \hat{\beta}_2 (1 - y_2) > \hat{\beta}_3 (1 - y_3) > \hat{\beta}_4 (1 - y_4) \]  

Also we define the following probabilities:

\[ \hat{p}_{H,i} < 1, \ i = 1, \ldots, N, \] is the probability that having undertaken investigative step i, the precondition examined in step i is considered satisfied when indeed this is the case. Clearly, \( (1 - \hat{p}_{H,i}) \) is the probability that the precondition examined in step i is erroneously considered as not satisfied and hence the conduct is acquitted.

\[ \hat{p}_{B,i} < 1, \ i = 1, \ldots, N \] is the probability that having undertaken investigative step i, the precondition examined in step i is not considered satisfied when indeed this is the case. Clearly, \( (1 - \hat{p}_{B,i}) \) is the probability that the precondition examined in step i is erroneously considered as satisfied and hence the conduct is banned (if there is no further assessment after step i).

Finally, we repeat here the definition of the following two parameters already defined above:

\[ p_{H,i} < 1, \ i = 1, \ldots, N - 1 \] is the average fraction of the cases examined in which the conduct is identified as harmful when it is indeed harmful, having undertaken investigative step i. Clearly, \( (1 - p_{H,i}) \) is the average fraction of cases examined erroneously considered as benign (and acquitted) for which step i is satisfied and which reduce consumer welfare. In this framework it makes sense to assume that this is less than one for stage 1,…,N-1. In the last stage N, with all assessment steps completed, if the precondition in this stage, as all other preconditions, is considered satisfied and recognised as satisfied, the conduct is certainly assessed as harmful and is banned, i.e. we assume that \( p_{H,N} = 1 \).

\[ p_{B,i} < 1, \ i = 1, \ldots, N \] is the average fraction of the cases examined in which the conduct is identified as benign when it is indeed benign, having undertaken investigative step i. Clearly, \( (1 - p_{B,i}) \) is the average fraction of cases examined erroneously considered as harmful (and banned, if there is no further assessment after step i) for which step i is satisfied but which increase consumer welfare. Clearly, \( p_{B,N} < 1 \) since if precondition in stage N does not hold (so the conduct is benign), this may not be recognised and with positive probability the conduct will be wrongly banned.

All the four last probabilities are assumed to have values between 0 and 1 reflecting the fact that analyses and tests are never perfect and there can be false convictions as well as false acquitals. Further, it is assumed that, the additional investigative steps and tests carried out

---

45 Given that, for example, \( \hat{\beta}_2 (1 - y_2) = \hat{\beta}_1 \beta_2 (1 - y_2) < \hat{\beta}_1 (1 - y_1) \) etc

46 Note that in our framework in makes sense to define this and the next probability for stage 1,…,N-1. In the last stage N, with all assessment steps completed, if the precondition in this stage, as all other preconditions, is considered satisfied the conduct is certainly harmful and banned whilst if the precondition is not considered satisfied the conduct is certainly acquitted.
have some *discriminatory power* so the probability of banning a harmful conduct is greater than the probability of banning a benign one and so:

\[ p_{H,i} > 1 - p_{B,i} \]  \hspace{1cm} (7)

As can be seen, this is equivalent to assuming that the probability of acquitting a genuinely benign conduct is higher than the probability of acquitting a genuinely harmful conduct.

Also it is assumed that

\[ \hat{p}_{H,i} > 1 - \hat{p}_{B,i} \]  \hspace{1cm} (8)

i.e. the probability that precondition \( i \) is considered to be satisfied when it is, is higher than when it is erroneously considered to be satisfied, when it is not.

It is natural to assume that

\[ p_{H,i} < p_{H,i+1}, i = 1, \ldots, N - 1 \]  \hspace{1cm} (9)

\[ p_{B,i} < p_{B,i+1}, i = 1, \ldots, N - 1 \]  \hspace{1cm} (10)

(9) and (10) say that the ability of the CA to recognise without errors genuinely harmful and genuinely benign conduct increases as the investigative steps and thus the information and evidence about the specific conduct examined increases.

To conclude this sub-section we generalise our definition of *presumptively legal (PL)* or *presumptively illegal (PI)*, for any stage \( i = 1, \ldots, N \): having carried out \( i \) investigative steps, \( i = 1, \ldots, N \), we can say that the conduct is PL or PI depending on whether, respectively, \( \bar{h}_i = \gamma_i H - (1 - \gamma_i) B < 0 \), or \( \bar{h}_i = \gamma_i H - (1 - \gamma_i) B > 0 \), \( \bar{h}_i \) being the average harm of conducts for which precondition \( i \) is satisfied. Of course, \( \bar{h}_N = H \).

**4.3 The potential liability decisions based on the DEC following assessment in stage \( i \)**

A precondition examined in investigation stage \( i \) can be satisfied or not be satisfied: if satisfied the conduct can be harmful or benign; if it is not satisfied the conduct is certainly benign. Thus, having completed investigation \( i \), the following decisions can be made:

- If precondition \( i, i = 1, \ldots, N \), is not considered to be satisfied, acquit the conduct in step \( i \);
- If precondition \( i, i = 1, \ldots, N-1 \), is considered to be satisfied:
  - Decide whether to use \( PBND \ LS_i \) or \( PBD \ LS_i \) by comparing their respective DEC;
  - Decide whether the information and evidence collected up to step \( i \) is sufficient to reach an infringement decision or whether additional analyses and evidence should be sought, by comparing DEC under \( LS_i \) to DEC under \( LS_{i+1} \). If DEC under \( LS_i \) are considered lower than DEC under \( LS_{i+1} \), use \( LS_i \) to reach a liability decision, otherwise move to stage \( i+1 \) and \( LS_{i+1} \). If it is decided to use \( LS_i \), then:
    - Ban all conduct if it is decided to use \( PBND \ LS_i \);
    - Ban or allow the conduct depending on the strength of the harm signal received under a \( PBD \ LS_i \) (measured by the probabilities \( p_{H,i} < 1 \) and \( p_{B,i} < 1 \)).
If precondition \( i = N \), is considered to be satisfied and given that examining precondition \( i = N \) (the last one), implies that all previous preconditions were examined and considered to be satisfied, the conduct is considered certainly harmful, so \( \gamma_N = 1 \). In this case the conduct is either banned with certainty if precondition \( N \) is considered satisfied or is acquitted with certainty if precondition \( N \) is not considered to be satisfied (since then it is considered certainly benign).

The following decision tree provides a succinct presentation of all the decisions that could be reached depending on the circumstances, as reflected in the value of the probabilities described above.

4.4 Determining the DEC from reaching liability decisions in each assessment stage
We start with the following result:

**Lemma 1:** A PL conduct in step \( i = 0, \ldots, N-1 \) can turn into a PI conduct in the next steps while a PI conduct in step \( i = 0, \ldots, N-1 \) will be even more PI in the next assessment step.

**Proof:** true since \( \gamma_i \) is increasing with \( i \).

In most of the cases that are the focus of our investigation (i.e. cases excluding horizontal agreements and, specifically, considering abuse of dominance practices, vertical restraints or concerted practices) there is broad unanimity that conducts are PL in step 0 but may be PI in step 1, i.e. following detailed market contextualisation, when these conducts are undertaken by firms with SMP or dominant firms. If the conduct is considered PL in step 0 and is also PL in step 1 then it is unlikely that it will be condemned or be the subject of further investigation.

**So below we focus on cases where the conduct is PL in step or stage 0, so \( h_0 < 0 \) while it is PI in stage 1, i.e., \( h_1 > 0 \).**

A next question is whether, given these assumptions, we should pursue investigation 1, rather than use a strict Per Se rule and allow all the conducts in step 0. In the latter case, the DEC will be the cost of errors from false acquitals, i.e:

\[
DEC_0 = \gamma_0 H
\]  
(11)

If the assessment in stage 1 lowers DEC then the question becomes whether additional information should be obtained in steps 2, 3 etc.

To compare DECs more generally, we now define the DEC for all the above types of LSs. They are the following.

(a) **DEC of Strict Per Se (SPS) LS**

\[
DEC LS_0 = \gamma_0 H \ 	ext{DEC of False Acquitals if } h_0 < 0 \ \text{(PL conduct) and } DEC LS_0 = (1 - \gamma_0)B = \text{DEC of False Convictions if } h_0 > 0 \ \text{(PI conduct)}^{47}.
\]

(b) **DEC of Presumption-based (PB) LS**

b.1) **DEC under a PB Non-Discriminating LS** : under a PBND LS, \( i = 1, \ldots, N-1 \), with \( h_i > 0 \) (conduct is PI), all conducts that are found to satisfy precondition i are banned, though not all preconditions for establishing liability have been investigated and it is understood that not all conducts that satisfy precondition i are harmful (and hence should not be banned). So, for example, in step 1, under PBND, all conducts in the category in which dominance is considered to be present are banned without any attempt to discriminate between those conducts in this category that are harmful from those that are benign – essentially, with this LS, it is presumed that all the preconditions that must be investigated in steps after step 1 in order to establish harm to welfare are satisfied. DEC are then given by:

\[
DEC PBND LS_i = \beta_i \hat{p}_{H,i} (1 - \gamma_i)B + \hat{p}_i (1 - \hat{p}_{H,i}) \gamma_i H + (1 - \hat{p}_i) (1 - \hat{p}_{B,i}) B
\]  
(12)

The first term on the RHS is the DEC (false convictions) from banning benign (probability \( (1 - \gamma_i) \)) conducts, that would generate welfare benefits B, for which precondition i is genuinely satisfied (probability \( \hat{p}_i \)), given that precondition i is considered to be truly satisfied with

---

probability $\hat{p}_{H,i}$. The second term on RHS is the DEC (false acquittals) from acquitting harmful (probability $y_i$) conducts, and thus incurring welfare loss $H$, for which precondition i is satisfied (with probability $\hat{p}_i$), given that precondition i is mistakenly considered not to be satisfied with probability $(1 - \hat{p}_{H,i})$. Finally, the third term on the RHS is the DEC (false convictions) from banning benign conducts and thus losing welfare benefit $B$, for which precondition i is not satisfied (with probability $(1 - \hat{p}_{B,i})$), given that following the investigation it is mistakenly considered that precondition i is satisfied (with probability $(1 - \hat{p}_{B,i})$).

b.2) DEC under a PB Discriminating LS: under a PBD $LS_i$, $i = 1, \ldots, N - 1$, with $\tilde{h}_i > 0$, not all conducts that are found to satisfy precondition i are banned, which is the difference between this LS and PBND $LS_i$. DEC in this case are:

$$DEC_{PBD} LS_i = \hat{p}_i \hat{p}_{H,i} y_i (1 - p_{H,i}) H + \hat{p}_i \hat{p}_{H,i} (1 - y_i) (1 - p_{B,i}) B + \hat{p}_i (1 - \hat{p}_{H,i}) y_i H + (1 - \hat{p}_i) (1 - \hat{p}_{B,i}) (1 - p_{B,i}) B$$

(13)

There is now an additional false acquittals DEC term from not banning all harmful conducts (the first term on the RHS of equation (13)) – banning with probability $(1 - p_{H,i})$. The false convictions DEC term, second term on the RHS of equation (13) now changes compared to DEC PBND $LS_i$ - first term of (12): false convictions, from banning conducts for which precondition i is satisfied even though they are benign, are now lower, given that now only a fraction of these, $(1 - p_{B,i})$, are banned. The 2nd term on the RHS of DEC PBND $LS_i$ in (12) is the same as the 3rd term in equation (13). The 4th term in (13) shows DEC from false convictions of benign conducts, for which precondition i is not satisfied, but this is not recognised, and so they are banned with probability $(1 - p_{B,i})$.

c) DEC of Rule of Reason or (full) EB LS (LS_N): as explained above under this LS if precondition N is considered satisfied the conduct is banned with certainty given than $y_N = 1$. As noted, this cannot be a discriminating LS, since the latter can be used in order to discriminate between harmful and benign conducts when precondition i is satisfied. And, if precondition N is considered not-satisfied the conduct is acquitted with certainty. Thus, depending on the finding regarding precondition N, conducts are treated in exactly the same way - all those for which the precondition is considered satisfied are banned and all those for which the precondition is considered not satisfied are acquitted. This is as for the case of PBND $LS_i$, $i = 1, \ldots, N - 1$. So from (12) (DEC for PBND $LS_i$) the DEC for the Rule of Reason are:

$$DEC_{LS_N} = \hat{p}_N (1 - \hat{p}_{H,N}) H + (1 - \hat{p}_N) (1 - \hat{p}_{B,N}) (1 - p_{B,N}) B$$

(14)

So now there are DEC from false acquittals from wrongly acquitting conducts for which precondition N is satisfied but this is not recognised and DEC from false convictions from wrongly convicting conducts for which precondition N is not satisfied but, again, this is not recognised.

4.5 Comparisons and results: optimal LSs

We can now prove a number of results about whether additional assessment steps should be undertaken and thus about the determination of optimal LSs. To start with we define the following indicators:
\[ \frac{\gamma_i H}{(1-\gamma_i)B} = s_i > 1 \] is the measure of the “strength of the presumption of illegality” \( s_i \) in stage i, i.e. the strength of the presumption that the conduct is harmful, when the preconditions for welfare harm in i and before i \( (i = 1, ..., N-1) \) are considered satisfied.

\[ \frac{(\hat{p}_{H,i})}{(1-\hat{p}_{B,i})} > 1 \] is the measure of the “discriminatory power” \( \hat{d}_i \) in identifying correctly when preconditions are satisfied.

\[ \frac{p_{H,i}}{(1-p_{B,i})} > 1, \quad \frac{p_{B,i}}{(1-p_{H,i})} > 1 \] these measure the “discriminatory power” \( (d_{H,i}, d_{B,i}) \) of presumption-based discriminatory LSs (PBD LSs) in identifying correctly when the conduct is harmful and when it is benign, given that precondition i is satisfied.

\[ \left( \frac{\hat{p}_{H,i}}{(1-\hat{p}_{B,i})} \right) \] this measures the “degree of prevalence” \( \hat{\omega}_i \) of the preconditions i, i.e. the extent to which the presence of the preconditions is widespread in the market under consideration.

**Proposition 1: Conditions for undertaking step 1, i.e. for not adopting Strict Per Se**

Assume that the conduct is PL in step 0 but PI in step 1, so \( \tilde{h}_0 < 0 \) while \( \tilde{h}_1 > 0 \). It is then optimal to proceed with the step 1 investigation, as this will lower DEC relative to a Strict Per Se treatment of the conduct, when \( \gamma_i \) and hence \( s_i \), as well \( \hat{\omega}_1 \) and \( \hat{d}_{H,1} \) are quite large.

Proof: we need to compare \( DEC \ LS_0 = \gamma_0 H \), given \( \tilde{h}_0 < 0 \), with \( DEC \ PBNB \ LS_1 \) and \( DEC \ PBD \ LS_1 \). For the result to hold at least one of the two latter should be lower than the former. Comparing first with \( DEC \ PBNB \ LS_1 \), we get, given that \( \gamma_0 = \beta_1 \gamma_1 \) and \( \hat{\beta}_1 = \beta_1 \):

\[ DEC \ LS_0 > DEC \ PBNB \ LS_1 \] if:

\[ \gamma_0 H > \hat{\beta}_1 \hat{p}_{H,1} (1 - \gamma_1)B + \hat{\beta}_1 (1 - \hat{p}_{H,1}) \gamma_1 H + (1 - \hat{\beta}_1) (1 - \hat{p}_{B,1})B \\
(15) \]

i.e., if:

\[ \left( \frac{1-\hat{p}_{B,1}}{\hat{p}_{H,1}} \right) \hat{p}_{H,1} < (1 - \gamma_1) \left( \frac{\gamma_1 H}{(1-\gamma_1)B} - 1 \right) \]

\[ 1 < (1 - \gamma_1) \left( \frac{\gamma_1 H}{(1-\gamma_1)B} - 1 \right) \left( \frac{\hat{\beta}_1}{\hat{\beta}_1} \right) \left( \frac{\hat{p}_{H,1}}{(1-\hat{p}_{B,1})} \right) \]

(15')

Or, using the notation of the indicators introduced above:

\[ (1 - \gamma_1) \hat{\omega}_1 \hat{d}_1 (s_i - 1) > 1 \]

(16)

\[ \hat{\omega}_1 \hat{d}_1 [\gamma_1 \frac{H}{B} - (1 - \gamma_1)] > 1 \]

(16)

This can be generalised to the case where the information in stage i the conduct is PL and in stage \( i + 1 \) it is PI. In this case, \( \tilde{h}_i < 0 \) while \( \tilde{h}_{i+1} > 0 \). In stage I we can allow all conducts with cost \( \gamma_i H \) or move to the next stage. We can compare the DEC's in the two stages as above,

---

48 These terms were first used by Katsoulacos and Ulph (2009).
leading again to the equations (15) / (16) with subscripts 0 replaced by i and 1 replaced by i + 1.

The condition (16) says something eminently intuitive: it is more likely to be optimal to take step 1 and then adopt, all other things equal (i.e., before determining whether it will be optimal to take additional steps), a PBND LS:

- the more widespread are dominant firms in the market for the type of conduct examined (the greater \( \hat{\omega}_1 \));
- the more likely that the conduct is harmful given it is undertaken by a dominant firm i.e the higher is \( \gamma_1 \) – hence, the greater the strength of the presumption of illegality \( s_1 \);
- the greater the discriminatory power in distinguishing dominant from non-dominant firms (\( \hat{d}_1 \)).

The interpretation is of course analogous for the general comparison between stages i and i + 1.

Note also that all these parameters / indicators are potentially different when we consider the same conduct types and markets across different economies and jurisdictions in which market structures and other characteristics are different. Thus these indicators can provide the basis for developing empirical hypotheses about why there are differences in LSs adopted for similar type of conduct in different countries.

Next, comparing strict Per Se to undertaking step 1, using, however, a PBD (i.e. discriminating) LS, the DEC under the former will be higher if:

\[
DEC\ LS_0 > DEC\ PBD\ LS_1 \quad i.e.\ if:
\]

\[
y_0^H > \beta_1 \hat{p}_{H,1} \gamma_1 (1 - p_{H,1})^H + \beta_1 \hat{p}_{H,1} (1 - \gamma_1) (1 - p_{B,1})^B + \beta_1 (1 - \hat{p}_{H,1}) \gamma_1^H + (1 - \beta_1) (1 - \hat{p}_{B,1}) (1 - p_{B,1})^B
\]

or, given that \( \gamma_0 = \beta_1 \gamma_1 \) and \( \hat{\beta}_1 = \beta_1 \) and re-arranging, if:

\[
\beta_1 \hat{p}_{H,1} \gamma_1 p_{H,1}^H - \hat{\beta}_1 \hat{p}_{H,1} (1 - \gamma_1) (1 - p_{B,1})^B - (1 - \hat{\beta}_1) (1 - \hat{p}_{B,1}) (1 - p_{B,1})^B > 0
\]

(17)

Or

\[
\hat{\beta}_1 \hat{p}_{H,1} (1 - \gamma_1) \left( \left[ \frac{\gamma_1^H}{(1 - \gamma_1)^B} \right] p_{H,1} - (1 - p_{B,1}) \right) - (1 - \hat{\beta}_1) (1 - \hat{p}_{B,1}) (1 - p_{B,1}) > 0
\]

(18)

Or:

\[
\hat{\omega}_1 \hat{d}_1 [\gamma_1^H p_{H,1} - (1 - \gamma_1)] > 1
\]

(19)

Comparing (19) with expression (16) we see that the expression on the RHS of (19) is greater if \( d_{H,1} > 1 \), as we have assumed, i.e. if there is some discriminatory power in our rule (the rule is better than deciding randomly whether examined conducts are harmful or benign). So we have:
Proposition 2: Comparison of discriminating to non-discriminating PB LSs at stage 1 (relative to the strict Per Se)

The discriminating LS \( PBD \, LS_1 \) is more likely to lower DEC relative to the use of the SPS LS than the non-discriminating LS \( PBND \, LS_1 \).

Next, we can compare PB LSs in any given stage. We can do this for stage \( i, i = 1,\ldots,N \). First we compare PBND LSs to PBD LSs and give a general proof that condition (19’) is a sufficient condition for a discriminating LS to be superior in stage \( i < N \) to a non-discriminating LS.

Proposition 3 (general comparison of the two types of presumption-based LSs) in stage \( i \):

Comparing the DEC of discriminatory and non-discriminatory PB LSs, we find that a sufficient condition for the \( PBD \, LS_i \), to be superior to a \( PBND \, LS_i \, i = 1,\ldots,N-1 \), is that the discriminatory power of \( PBD \, LS_i \, (d_{b,i}) \) following the investigation in step \( i \) is higher than the strength of the presumption of illegality of the conduct for which precondition \( i \) is considered to be satisfied.

Proof: must examine when \( DEC \, PBND \, LS_i > DEC \, PBD \, LS_i \). This will be true if:

\[
\beta_i (1 - \gamma_i) d_{b,i} B + \beta_i (1 - \beta_{H,i}) (1 - \gamma_i) + (1 - \beta_i)(1 - \beta_{B,i}) B > \beta_i (1 - \gamma_i) d_{b,i} B + \beta_i (1 - \beta_{H,i}) (1 - \gamma_i) + (1 - \beta_i)(1 - \beta_{B,i}) B
\]  

(20)

Or if:

\[
\hat{\beta}_i (1 - \gamma_i) d_{b,i} B - \beta_{H,i} (1 - \gamma_i) (1 - p_{B,i}) B + (1 - \beta_i)(1 - \beta_{B,i}) (p_{B,i}) B > \hat{\beta}_i (1 - \beta_{H,i}) (1 - p_{H,i}) H
\]  

(21)

Or if:

\[
\hat{\omega}_i (1 - \gamma_i) d_{b,i} B + (1 - \beta_{B,i}) (p_{B,i}) B > \hat{\omega}_i (1 - \beta_{H,i}) (1 - p_{H,i}) H
\]

Or if:

\[
\hat{\omega}_i (1 - \gamma_i) d_{b,i} B + (1 - \beta_{B,i}) (d_{b,i}) B > 0
\]  

(19’)

Which can also be written as:

\[
\hat{\omega}_i (1 - \gamma_i) (d_{b,i} - s_i) + d_{b,i} > 0
\]

(alternative to 19’)

This condition for discriminatory conducts to lower DEC – called condition for effective discrimination - was first put forward by Katsoulacos and Ulph (2009).49

Next, and most importantly, we examine when an additional assessment step (taking a further step in the analysis of the conduct’s effects) will lower DEC, for stages \( i = 1,\ldots,N-1 \). We focus our discussion on the comparison between discriminatory LSs in stages \( i \) and \( i+1 \), \( i = 1,\ldots,N-1 \) and so compare DEC of \( PBD \, LS_i \) to the DEC of \( PBD \, LS_{i+1} \).

---

49 See also Katsoulacos and Ulph (2015 and 2016).
Taking an additional assessment step will lower the DEC i.e. \( \text{DEC PBD } LS_{i} > \text{DEC PBD } LS_{i+1} \) if:

\[
\hat{\beta}_i \hat{p}_{H,i}y_i (1 - p_{H,i})H + \hat{\beta}_i \hat{p}_{H,i} (1 - y_i)(1 - p_{B,i})B + \hat{\beta}_i (1 - \hat{p}_{H,i})y_i H + (1 - \hat{\beta}_i)(1 - \hat{p}_{B,i}) (1 - p_{B,i})B > \hat{\beta}_{i+1} \hat{p}_{H,i+1}y_{i+1} (1 - p_{H,i+1})H + \hat{\beta}_{i+1} \hat{p}_{H,i+1} (1 - y_{i+1})(1 - p_{B,i+1})B + \hat{\beta}_{i+1} (1 - \hat{p}_{H,i+1})y_{i+1} H + (1 - \hat{\beta}_{i+1})(1 - \hat{p}_{B,i+1}) (1 - p_{B,i+1})B
\]

Consider the DEC from False Acquitals (FA) and False Convictions (FC) at stage \( i \) (the same expressions can be used for stage \( i + 1 \) by replacing subscript “\( i \)” with subscript “\( i+1 \)”). These are given by:

\[
\text{DEC FA}_i = \hat{\beta}_i \hat{p}_{H,i}y_i (1 - p_{H,i})H + \hat{\beta}_i (1 - \hat{p}_{H,i})y_i H \quad (22)
\]

or

\[
\text{DEC FA}_i = \hat{\beta}_i y_i (1 - \hat{p}_{H,i} p_{H,i})H \quad (22')
\]

and

\[
\text{DEC FC}_i = \hat{\beta}_i \hat{p}_{H,i} (1 - y_i)(1 - p_{B,i})B + (1 - \hat{\beta}_i)(1 - \hat{p}_{B,i}) (1 - p_{B,i})B \quad (23)
\]

or

\[
\text{DEC FC}_i = \hat{\beta}_i (1 - p_{B,i})B [\hat{p}_{H,i} (1 - y_i) - (1 - \hat{p}_{B,i})] + (1 - \hat{\beta}_i)(1 - \hat{p}_{B,i}) (1 - p_{B,i})B \quad (23')
\]

We can use Diagram 1 to clarify the above magnitudes. Note that in (22),

\[
\hat{\beta}_i \hat{p}_{H,i}y_i (1 - p_{H,i}) = \text{prob of FA1 (Diagram 1)}, \text{ and }
\]

\[
\hat{\beta}_i (1 - \hat{p}_{H,i})y_i = \text{prob. of FA2 (Diagram 2)}.
\]

In (23):

\[
\hat{\beta}_i \hat{p}_{H,i} (1 - y_i)(1 - p_{B,i}) = \text{prob of FC1 (Diagram 1)}, \text{ and }
\]

\[
(1 - \hat{\beta}_i)(1 - \hat{p}_{B,i})(1 - p_{B,i}) = \text{prob of FC2 (Diagram 1)}.
\]

We start by comparing (22) or (22’) to (23) in order get an exact characterisation of the factors that determine FA and FC and thus to determine under what conditions the Easterbrook (1984) hypothesis that has led to what Hovenkamp (2021) calls “an anti-enforcement bias in antitrust”, namely that expected error costs from FC are higher than from FA, holds\(^\text{50}\). Specifically, we see that:

**Lemma 2**: the Easterbrook (1984) hypothesis is likely to be valid (i.e., \( \text{DEC FC}_i > \text{DEC FA}_i \)), when:

\(^\text{50}\) According to G Manne (2020) “It was Judge Frank Easterbrook (1984) who generalized the approach for antitrust, and offered the clearest exposition of the error-cost approach”. Page 40.
- \( \hat{\beta}_i \), the degree of prevalence of precondition i is small (e.g. for stage 1, that the prevalence of dominant firms is small);
- \( \hat{p}_{Hi,i} \) and \( p_{Hi} \) are large, i.e. we can identify when harmful conduct is indeed harmful with a high degree of accuracy, while \( p_{Bi,i} \) is small, i.e. we cannot identify when benign conduct is indeed benign with a high degree of accuracy;
- \( \gamma_i \) is small, so the likelihood that the specific conduct type investigated is genuinely harmful, is small;
- B is large relative to H.

The hypothesis has recently been the subject of severe criticism. According to Shapiro (2021) “...Easterbrook argued that antitrust courts should err on the side of defendants, because “judicial errors that tolerate baleful practices are self-correcting, while erroneous condemnations are not. Like Bork, Easterbrook achieved his desired result based not on economic theory or empirical evidence, but by making strong and unjustified assumptions”.

Hovenkamp (2021), criticizes particularly the Easterbrook (1984) assumption that the average welfare cost B from a false conviction is likely to be larger than the average welfare cost H from a false acquittal, but we see from the above comparison of the DEC from FC and from FA that even if this were to be true (which may well not be in many cases as noted by Hovenkamp (2021)), there is no obvious reason to expect that \( DEC_{FC_i} > DEC_{FA_i} \).

Next, and very importantly, we compare DEC for successive stages, using expressions (22) / (22’) and (23). we get:

**Proposition 4: determinants of relative DECs of taking additional assessment steps, for stages \( i = 1, \ldots, N-1 \):**

Here we examine the effect on DECs of taking additional assessment steps due to the changes in the probabilities \( \gamma, \hat{\beta}, p_H \) and \( p_B, \hat{p}_H \) and \( \hat{p}_B \).

(i) In stage \( i+1 \), the probability that the conduct is harmful (\( \gamma \)), if the precondition assessed in \( i+1 \) is satisfied, is higher than in stage \( i \). This implies, *ceteris paribus*, that FCs fall (as the prob that the conduct is benign falls) and FAs rise. FAs rise by \( \hat{\beta}H[\hat{p}_H(1 - p_H) + (1 - \hat{p}_H)] \). FCs fall by \( \hat{\beta} \hat{p}_H(1 - p_B)B \). Thus, **DECs will fall if**

\[
\frac{H}{B} < \frac{\hat{p}_H(1 - p_B)}{\hat{p}_H(1 - p_H) + (1 - \hat{p}_H)} \quad \text{or} \quad \frac{H}{B} < \frac{\hat{p}_H(1 - p_B)}{(1 - \hat{p}_H p_H)} = \sigma \tag{24},
\]

Or, multiplying both sides by \((\gamma / 1 - \gamma)\) and using the notation above, if:

\[
S = \frac{\gamma H}{(1 - \gamma)B} < \frac{pr_{FC1}}{pr_{FA1+FA2}} = \frac{\gamma \hat{p}_H(1 - p_B) \hat{\beta}}{(1 - \gamma)[\hat{p}_H(1 - p_H) + (1 - \hat{p}_H)]} = \frac{\gamma}{(1 - \gamma)} \sigma \tag{24’}
\]

where \( \sigma = \frac{\hat{p}_H(1 - p_B)}{[\hat{p}_H(1 - p_H) + (1 - \hat{p}_H)]} = \frac{\hat{p}_H(1 - p_B)}{(1 - \hat{p}_H p_H)} \tag{25} \)

Comments: DECs will fall after the assessment of an additional precondition, as a result of the increase in \( \gamma \), if the strength of the presumption of illegality (s) or \( (H/B) \) is lower than \( \sigma \) where the latter measures the *probability of FCs relative to that of FAs*. Thus:
• As intuition suggests, the lower the strength of the presumption of illegality the more likely that DEC will fall by undertaking further investigations.

• Given the presumption of illegality, it will be more likely that DEC will fall by undertaking further investigations the higher is $\alpha$, i.e. the probability of FCs relative to that of FAs. The probability of FCs is higher, the lower the discriminatory power of economic analysis in the assessment of benign conducts (the lower is $p_B$) and the probability of FAs is higher, the lower the discriminatory power of economic analysis in the assessment of harmful conducts (the lower is $p_H$). Thus, when benign conducts are relatively difficult to identify accurately ($p_B$ is low) while harmful conducts are relatively easy to identify accurately ($p_H$ is high), $\alpha$ will be high, (24) will be more likely to hold and the increase in $\gamma$ after an additional assessment is more likely to lower DECs. The intuition is that as the conduct is PI, i.e. $H$ is large relative to $B$ and FAs are large relative to FCs, DECs are more likely to fall as a result of the increase in $\gamma$, by a further investigation, if the probability of false acquittals is small relatively to the probability of false convictions.

• The higher is $\hat{p}_H$, i.e. the prob that the precondition examined is correctly identified as holding, the more likely that condition (24) is satisfied and therefore that DECs will fall by a further investigation.

The probability ($\hat{\beta}_i$) that a screen is satisfied given that previous screens are satisfied, which becomes lower for additional screens. From (22) and (23) we see that, ceteris paribus, this unambiguously reduces costs of FAs and certainly also reduces costs of FCs if $\hat{p}_{H,i}(1 - \gamma_i) - (1 - \hat{p}_{B,i}) > 0$. In relation to this we note that assuming that the probability of identifying correctly that condition $i$ holds is higher than the probability of identifying incorrectly that the condition does not hold (the minimum level of the ability to discriminate), means that $\hat{p}_{H,i} + \hat{p}_{B,i} > 1$. Even so, the condition above will not hold if $\gamma_i$ is very high.

Looking at (22’) and (23) however we note that the impact on DEC depends also on the combined effect of $\gamma_i$ and $\hat{\beta}_i$. We note that the probability $\hat{\beta}_i \gamma$, that the precondition examined holds but the conduct is harmful, may fall or rise, so from (22’) the effect on FAs is ambiguous – it depends on conduct and market context. In many cases $\hat{\beta}_i \gamma$ may rise. This will be the case when in stage $i+1$, the probability $\hat{\beta}_i$ does not fall much as will be the case when, between for example stages 1 and 2, the probability does not fall much because in most cases where dominance is identified the conduct by the dominant firms will have explosionary effects51. Also, relative to stage $i$, in stage $i+1$, the probability $\hat{\beta}_i (1 - \gamma)$, that the precondition examined holds but the conduct is not harmful falls, since $\hat{\beta}_i$ falls, $\gamma$ rises and $\hat{p}_{H,i} > 1 - \hat{p}_{B,i}$. It follows from (23) that FCs fall. Thus, there will be a fall in FCs and there may be an increase in FAs. The fall in FCs will be $\hat{p}_H(1 - p_B)B$ while if there is an increase in FAs the increase will be $(1 - p_H \hat{p}_H)H$, i.e. DECs will again fall if condition (24) holds.

51 Or, between 2 and 3, the probability does not fall much because in most cases where exclusion is identified the conduct leads to welfare harm – if efficiencies are not accounted for.
The parameters $p_H$ and $p_B$ will increase as additional investigative steps are undertaken, reflecting the fact that the accuracy of the estimate about whether conduct is harmful or benign improves as more analyses and tests are applied in assessing additional preconditions / screens. As is clear from (22) and (23) this lowers FAs and FCs and hence reduces DECs.

The parameters $\hat{p}_H$ and $\hat{p}_B$ in stages i and i+1 measure the accuracy of estimates of whether or not different preconditions examined in these stages hold and this may rise or fall. E.g. the prob of identifying correctly if dominance exists or not in stage 1 may be higher or lower than the prob of identifying correctly if there is exclusion in stage 2, given that dominance is identified in stage 1. Similarly for stage 3, etc.

$\hat{p}_B$ affects only FCs and an increase (decrease) in $\hat{p}_B$ in successive stages reduces (increases) FCs and hence reduces (increases) DECs. $\hat{p}_H$ affects both FCs and FAs. From (22’) and (23), an increase reduces FAs and increases FCs. However, we note that the increase in FCs is $\hat{p}(1 - \gamma)(1 - p_B)B$ and the decrease in FAs is $\hat{p}_H\gamma p_HH$. Thus DECs will decrease if

$$s > \frac{(1-p_B)}{p_H} = \frac{1}{d_H}$$

which holds for a presumptively illegal conduct ($s > 1$) and since $d_H > 1$. Thus if $\hat{p}_H$ and $\hat{p}_B$ increase between stages i and i+1 this will unambiguously reduce DECs.

Even if $\hat{p}_H$ and $\hat{p}_B$ decrease between stages i and i+1 we note that what matters for FAs is what happens to $\hat{p}_H p_H$ and for FCs what happens to $(1 - \hat{p}_B)(1 - p_B)$ and $p_H$ and $p_B$ will increase. Thus for as long as discriminating between harmful and benign conduct improves significantly between stages i and i+1 this will reduce DECs from the additional investigation for as long as $\hat{p}_H$ and $\hat{p}_B$ also increase or, if they decrease, this is compensated by the increase in $p_H$ and $p_B$ so that $\hat{p}_H p_H$ and $(1 - \hat{p}_B)(1 - p_B)$ decrease.

One last result concerns the comparison between the RoR LS, that involves undertaking also the last assessment stage (stage 4 in our context), with using a LS that stops short of this last stage in the investigation. Let us consider in particular whether it is optimal to undertake this last step – which involves comparing DEC of stage 3 to DEC of Stage 4, assuming that in stage 3 a PBD LS is used – the comparison with PBND LS for stage 3 is then obvious given Proposition 3.

From equations (22) and (23) for $i = 3$ and equation (14), repeated here for stage $i = 4 = N$:

$$DEC\;LS_4 = \hat{p}_H(1 - \hat{p}_{HA})H + (1 - \hat{p}_H)(1 - \hat{p}_{BA})(1 - p_{BA})B$$

we can compare DEC from FAs and DEC from FCs for these two LSs. We get:

**Proposition 5:**
(a) it is sufficient (though not necessary) condition for the last assessment step to lower the DEC from FAs that the probability of identifying the precondition in step 4 ($\hat{p}_{H,4}$) is greater or not much smaller than it is for step 3.

(b) However, the last step can decrease or increase FCs, so the overall effect on DEC is ambiguous. Using sets of reasonable parameter values we provide two examples in the Appendix in which the overall effect is that DEC decrease by undertaking the last step.

Proof: to see (a) note that

$$DEC \ FA_3 = \hat{\beta}_3 Y_3 (1 - \hat{p}_{H,3} p_{H,3}) H$$  \hspace{1cm} (28)

and

$$DEC \ FA_4 = \hat{\beta}_4 (1 - \hat{p}_{H,A}) H$$  \hspace{1cm} (29)

Thus, given that $\gamma_3 = \beta_4$ and so $\hat{\beta}_3 Y_3 = \hat{\beta}_4$

$$DEC \ FA_4 < DEC \ FA_3 \ if \ p_{H,3} \hat{p}_{H,3} < \hat{p}_{H,A}$$  \hspace{1cm} (30)

For (b) we need to compare:

$$DEC \ FC_3 = \hat{\beta}_3 \hat{p}_{H,3} (1 - \gamma_3) (1 - p_{B,3}) B + (1 - \hat{\beta}_3)(1 - \hat{p}_{B,3})(1 - p_{B,3})B$$  \hspace{1cm} (31)

with

$$DEC \ FC_4 = (1 - \hat{\beta}_4)(1 - \hat{p}_{B,A})(1 - p_{B,A})B$$  \hspace{1cm} (32)

Thus there are two conflicting effects. $DEC \ FC_3$ tend to be greater than $DEC \ FC_4$ because:

- Of the positive first term on the RHS of (31): before the investigation in stage 4 is undertaken, there is no certainty whether the conduct is harmful, even though all preconditions were satisfied in the previous stages, so $\gamma_3 < 1$. On the other hand, $\gamma_4 = 1$;
- Also, as noted above, given that $p_{B,3} < p_{B,A}$ and the assumption in (a), we may well have

$$ (1 - \hat{p}_{B,3})(1 - p_{B,3}) > (1 - \hat{p}_{B,A})(1 - p_{B,A}) $$  \hspace{1cm} (33)

On the other hand, $DEC \ FC_3$ tend to be smaller than $DEC \ FC_4$ because $\hat{\beta}_3 > \hat{\beta}_4$. For many conducts and markets in which preconditions 2, and 3 are satisfied, efficiency considerations in stage 4 will not be sufficient to outweigh the anticompetitive effects of the conducts and so the probability $\hat{\beta}_4$ will be very close to the probability $\hat{\beta}_3$. In these cases $DEC \ FC_3 > DEC \ FC_4$ and a full-effects-based or rule of reason should be adopted.

To get a better feeling of how DEC and its components (the cost from FA and from FC) compare, in the Appendix we make some comparisons using numerical examples with reasonable parameter values.
5. Differences in the presumption of illegality vs. legality and in error-minimising LSs across countries

5.1 Introductory remarks

As we mentioned in the introduction, at present in many jurisdictions (including those of North America and of the EU) only very rarely will a general conduct type identified in stage 0 be characterised as Presumptively Illegal, without any additional contextualisation of the circumstances under which the specific conduct is undertaken, the exception been that of horizontal hard-core cartels in US\textsuperscript{52}. Indeed, for abuse of dominance practices and many vertical restraints and concerted practices this is not the case and the general conduct types examined under these enforcement categories are currently characterised as PL (in stage 0). Of course, liability decisions on such conducts are never or very rarely taken using a (Strict) Per Se LS: at least some case specific investigations are first undertaken. At a minimum, this is in order to contextualise market conditions and to establish whether there is significant extant market power (step 1 of the investigative steps defined above). Having undertaken this step, if it is determined that the firms involved have SMP (or are dominant), the question then becomes whether the general type of conduct, \textit{when undertaken by dominant firms}, is considered PI (on average harmful) or PL (on average benign) and what is considered to be the strength of this presumption.

There are still significant differences in the answer to this question: the presumption formed and hence the view taken on whether a conduct is PI or PL as well as the view on the strength of the presumption will depend on the jurisdiction in which it is examined\textsuperscript{53}. Also it might depend on the type of market and this is considered in detail in the next section for the category of the “big tech digital markets”.

From our model above, we note that in stage 1 (which assesses the degree of market power / contestability) conduct will be presumptively illegal if $H$ is large relative to $B$ and/or $\gamma_1$ (the probability that when undertaken by a dominant firm the conduct will be harmful) is quite large and it will be presumptively legal if $H$ is small relative to $B$ and/or $\gamma_1$ is quite small. It is important to note that if the conduct is presumptively legal in stage 1 and the Modified Per Se legality LS is chosen, all conducts will be permitted and the cost of decision errors will be just the \textit{costs of false acquittals}, $\gamma_1 H$. If the conduct is presumptively illegal and the Modified Per Se Illegality LS is chosen, all conducts will be banned and the cost of decision errors will be just the \textit{costs of false convictions} $(1 - \gamma_1) B$\textsuperscript{54}. Thus false convictions are large relative to false acquittals if $H$ is small relative to $B$ and/or $\gamma_1$ is quite small.

More generally, we provided an exact characterisation of all these and all other influencing factors and thus can determine under what conditions the Easterbrook (1984) hypothesis holds (Lemma 2 above). As already noted above, the Easterbrook hypothesis has

\textsuperscript{52} See also discussion above.

\textsuperscript{53} We would like to thank David Evans for bringing to our attention recently his 2009 article in which there is a thoughtful discussion of many of the issues we analyse here.

\textsuperscript{54} In the simplest case where the CA does not try to discriminate between harmful and benign conduct undertaken by dominant firms.
led to what Hovenkamp (2021) calls “an anti-enforcement bias in antitrust”, namely that expected error costs from false convictions are higher than from false acquitals. We showed that, the smaller the probability that a screen is satisfied, which in stage 1 means the smaller the prevalence of dominant firms (or, more correctly, the higher the probability of markets’ contestability), the higher the relative value of false convictions. Also, when we can identify that harmful conduct is indeed harmful with a high degree of accuracy (that depends on the probability of identifying correctly if the screen is satisfied and on the probability of identifying that, given this, the conduct is harmful when it is), but cannot identify benign conduct with a high degree of accuracy, the higher will be the relative value of false convictions.

The perception about the value of these parameters explains why, as mentioned above, there are very significant differences in the answer to the question of whether a specific conduct should be considered presumptively illegal or legal and what is the relative value of false convictions and false acquitals across different jurisdictions. Hovenkamp (2021), criticizes particularly the Easterbrook (1984) assumption that B is likely to be larger than H, but we see from the list of factors just mentioned that even if this were to be true there is no obvious reason to expect that the decision error costs of false convictions are higher than those of false acquitals.

Of course, in a jurisdiction in which the dominant economic ideology places greater trust on the markets’ ability to self-correct, that tends to significantly lower the value of H (dominant US model in the last 40 years) and puts great weight on the incentive effects of false convictions\(^{55}\) (raising the value of B), it is much more likely to characterize a conduct as PL and to consider false convictions more costly than false acquitals than in a jurisdiction that does not place as much trust on markets’ ability to self-correct, de-emphasizes incentive effects and places trust in the governments’ ability to improve outcomes through intervention (EU model). In the latter, it is much more likely to characterize a conduct as PI and to consider false convictions less costly than false acquitals. This is, of course, a very important consideration in explaining the different enforcement approaches in US and EU mentioned above. Indeed, Anu Bradford et.al. (2019) attribute to this difference in economic ideology the emerging “Global Dominance of European Competition Law Over American Antitrust Law”.

Comparison of developed and developing countries

Needless to say, in developing countries\(^{56}\) the ability of markets’ to self correct will usually be even more limited than in developed economies, as entry barriers and other market failures will be higher, and the CA decisions are unlikely to have significant adverse incentive

\(^{55}\) Adverse deterrence effects or “chilling” effects, also mentioned above. Another important factor is the significance attributed to the potential efficiencies generated by a conduct. As an example, Hylton and Salinger consider that for the case of tying “false acquittal costs are likely to be small relative to false convictions when there are (1) market constraints on the firm’s conduct, (2) strategies other than tying that the firm could use to gain the same advantage in the market, or (3) no clear incentive to use tying in order to harm consumers. On the other hand, false conviction costs are likely to be relatively large when (1) there are substantial potential efficiencies associated with tying and (2) tying is an important competitive instrument”.

\(^{56}\) See footnote 6 above for a precise definition of this term in our context.
effects\textsuperscript{57}. The likelihood that the conduct undertaken by a dominant firm is genuinely harmful, which as we have seen is a very significant factor in determining whether conduct is presumptively illegal and the relative size of error costs, is likely to be much higher in developing countries. Thus, in the latter, \textit{the presumption is much more likely to be that of illegality and false acquitals more costly than false convictions}. To illustrate, consider a potentially anticompetitive conduct (such as some predatory pricing or rebate schemes) that on their own would not be able to limit entry into markets but can do so (and will be used to do so) when some other entry barriers or market failures are present. In this case, in countries/jurisdictions with low other entry barriers or market failures, might not even engage actively in enforcement against such practices, while in jurisdictions with many other entry barriers the value of enforcement will be very high. Further, in the former countries, if CAs do remain involved in enforcing competition law for such conduct, they should be using effects-based assessment, but (the administratively less costly) Per Se LSs should be used in jurisdictions in which other barriers are relatively high (prevalence of contestability is low), since in the latter the presumption of illegality of these conducts – i.e., presumption that they create harm - is much higher\textsuperscript{58} with cost of false convictions more likely to be lower than the cost of false acquitals.

Using our model to examine the factors that determine whether an additional assessment stage (and, hence, a movement from Per Se towards effects-based) will lower decision errors, also shows that \textit{in developing countries this movement is less likely to be justified for any given conduct than in developed countries}. Proposition 4 above, considers all the factors that need to be taken into account. In the discussion below we assume that the conduct type is presumptively illegal in stage 1, i.e., when the conduct is undertaken by firms with significant market power. Hence: $\tilde{h}_i = \gamma_i H - (1 - \gamma_i)B > 0, i \geq 1$. We consider each factor in turn taking into account Proposition 4 and (22) or (22') and (23).

(i) The first factor, $\gamma_i$, is considered in Proposition 4 (i). As noted above this is likely to be higher in developing countries and this, as well as the fact that $(H/B)$ is likely to be higher, makes the strength of the presumption of illegality higher. Thus, from (24), an additional investigation is more likely to increase DECs than would be the case in a developing country. Note that in (24), it is not clear whether the value of $\sigma$ will be higher or lower in a developing country since, relative to a developed country, both the probability of FAs and that of FCs are likely to be higher. However, $\tilde{p}_H$, i.e., the prob that the precondition examined is correctly identified as holding, is likely to be lower in developing countries make it the more likely that

\textsuperscript{57} Hovenkamp’s (2021) remark that “Firms are pretty good at inventing around legal rules” and that “Courts can also invent around their own previous ruling, construing them more broadly or narrowly as perspective changes” is even more likely to be true in developing countries.

\textsuperscript{58} See also Fox and Gal (2014) for a closely related discussion concerning the need for enforcing competition law in developing countries. Their discussion reminds us that different jurisdictions are characterised by different degrees to which competition is workable in products and services markets, \textit{in the absence of CL enforcement}. As we stressed in the introduction, the degree to which competition is workable, while it depends both on the anticompetitive conduct of firms, which enforcement seeks to eliminate, also, and perhaps primarily, depends on the more general conditions mentioned in footnote 5. Also see Gal (2004) and Bageri and Katsoulacos (2020).
condition (24) is not satisfied and therefore that DECs will rise by a further investigation.

(ii) The probability \( \hat{\beta}_i \) that a screen is satisfied given that previous screens are satisfied, which becomes lower for additional screens. From (22) and (23) we see that this unambiguously reduces costs of FAs and certainly also reduces costs of FCs if \( \hat{p}_{H,i}(1 - \gamma_i) - (1 - \hat{p}_{B,i}) > 0 \). This is less likely to hold in developing countries since, as noted, \( \gamma_i \) will be higher and \( \hat{p}_{H,i}, \hat{p}_{B,i} \) are likely to be smaller. Thus again DECs are less likely to fall in developing than in developed countries with additional investigations.

Furher, the reduction in the value of the probability \( \hat{\beta} \) is likely to be smaller or much smaller in developing countries since the percentage size of the reduction is smaller the greater the probability that in the next screen assessment the screen will be satisfied, and this probability will be greater in developing countries. Thus, using an example, in order to be concrete, the fact that the probability that in the additional screen the conduct has exclusionary impact is higher in developing countries implies that the percentage reduction of this probability relative to the probability of significant market power (the previous screen) is smaller in these countries, which tends to make the reduction in the costs of false convictions smaller.

(iii) Then, there are four probabilities that essentially determine the improvement in the discriminatory power of the assessment after an extra screen is examined. Specifically, the discriminatory power of the assessment depends on the following four probabilities. First, there are the probabilities that a harmful conduct is correctly identified as harmful \( \hat{p}_H \) and the probability that a benign conduct is correctly identified as benign \( \hat{p}_B \), once a screen is examined. It is expected that these two probabilities increase as more assessments are made (screens examined). The increase in the former probability lowers the cost of false acquittals while the increase in the latter lowers the costs of false convictions. There are a significant number of factors that tend to make the level and the increase in these probabilities, if additional assessments are made, much smaller in developing than in developed countries. These factors are limited experience, a “short” case-law history on which to rely, limited skills and resources and limited data in terms of availability and quality. This lack in the discriminatory power of the assessments is one of the most important factors making LSs closer to Per Se than effects-based the error minimising choice in developing countries.

(iv) Then, there are the probabilities that a screen is correctly identified as holding when it holds \( \hat{p}_H \) and is correctly identified as not holding when it does not hold \( \hat{p}_B \). It is not obvious how these probabilities will change for additional screens. One important consideration is that additional screens are more likely to require more sophisticated, but also more ambiguous in its predictions economic analysis, and this leads to an increase in the difficulty of correctly identifying the screen. Thus, to give an example, it is likely to be less difficult to identify correctly exclusionary impact than a reduction
in consumer welfare – thus the probability of identifying correctly the second screen will be lower than the probability of identifying correctly the third screen.

A decrease in the first probability (correctly identifying a screen as holding when it holds) increases decision error costs from FAs and the decrease in this probability is likely to be more pronounced in developing countries, making even larger the increase in these costs. While the decrease of this probability tends, as a first effect, to reduce costs from FCs, the latter will tend to increase from a decrease in the second probability (correctly identifying a screen as not holding when it does not hold) and this increase is likely to outweigh the first effect. Again, the fact that the decrease in the second probability is likely to be more pronounced in developing countries, makes even larger the increase in costs from false convictions in these countries.

In short, the examination of the factors affecting the DECs of additional investigations implies quite strongly that there is a reduced need to apply effects-based in developing than developed countries for error minimising reasons.

6. Should there be a switch to more presumption bases LSs in the case of the big tech digital platforms?

To start with, we would like to stress that the purpose of the following decision error-minimising analysis is to contribute to the discussion of whether there should be some general modification to the legal standards or tests applied in assessing conduct types that are potentially abusive, in dominance cases when these involve big tech digital platforms. That is, we agree with Cremer et.al (2019) that “competition law should not try to work with the error cost framework in a case by case basis”59.

Differences between US and EU reflected in the treatment of big tech platforms

There are a number of differences. To start with, US tends to treat many more practices, even when undertaken by firms with significant market power, as presumptively legal (rather than illegal), that is, on average benign, than is the case in the EU, that would treat these practices when undertaken by dominant firms as presumptively illegal. Next and related to this, in US, the dominant view has been that the primary objective in antitrust enforcement is to limit false convictions rather than false acquittals – a view that has only started to been strongly criticised recently60 – and a view that is not held in the EC. As an outcome of these views in US there is a strong tendency to treat what are considered presumptively legal conducts using a Per Se Legality LS (or Modified Per Se Legality, or Quick Look), an approach that leads to a high rate of acquittals and has been particularly criticised with respect to


60 See in particular the detailed and very careful arguments of Baker (2015) and Hovenkamp (2021) to which we have already referred above. Also Hovenkamp and Scot Morton (2019, p. 19-20.)
enforcement in the high-tech digital markets\textsuperscript{61}. If, on the other hand, in US a practice is considered \textit{presumptively illegal}, then at least in the last two decades, there is a strong tendency to rely on extensive use of economic analysis and evidence in case-specific investigations, that is, to rely much more on the \textit{rule of reason (or full effects-based approach)}.

This is not the case in EU\textsuperscript{62} which has treated the practices considered \textit{presumptively illegal} until recently by relying more on \textit{object-based} or on intermediate LSs rather than full \textit{effect-based} and it is certainly not true in other jurisdictions\textsuperscript{63}. This difference in the treatment of presumptively illegal conduct is due to two factors. The strength of the presumption of illegality and the weight placed on the cost of false acquittals is higher in EU which leads to a choice of presumption-based LSs rather than effects-based. Another significant factor is that the substantive or liability standard has been more that of “disadvantaging rivals” or “distorting the competitive process” than “consumer welfare” in EU\textsuperscript{64} and this naturally makes the choice of LS what we have called Truncated Effects-Based\textsuperscript{65}.

Federico, Scott Morton and Shapiro (2020) argue, in the context of assessing effects on innovation, that adopting in US the disanvantaging rivals liability standard would be appropriate:

“...the quantum of evidence required to conclude that a dominant firm’s conduct has harmed innovation and thereby violated the antitrust laws is a critical element of competition policy. A more assertive antitrust regime will find antitrust violations in cases where the challenged conduct disrupts the competitive process by impeding the

\textsuperscript{61} As the Stigler Commiirere Report (2019) notes “there have been few antitrust challenges to exclusionary conduct since the government’s 1998 case against Microsoft, and courts have in several instances been hostile to such cases and have imposed daunting proof requirements on plaintiffs. Apparent under-enforcement is in part due to courts’ reliance on so-called Chicago School assumptions that do not have a sound theoretical or empirical basis”, p.64Baker et. all (2020) mention that “Both the Federal Trade Commission and the Justice Department have at times abetted the judicial retrenchment in antitrust law, particularly as it applies to the conduct of high-tech platforms, by declining to challenge (or in some cases even investigate) nearly all of the large number of platform acquisitions of arguably nascent competitors, and declining to challenge platform conduct that has been the subject of enforcement actions by sophisticated competition agencies abroad. Without regard to the merits of any individual decision, this systematic pattern of enforcement avoidance suggests that until now, the agencies have been too cautious in their enforcement posture toward Internet platforms” (page 13).

\textsuperscript{62} Both at the level of the EU Commission (EC) and that of Member States. The difference is thought to be particularly pronounced in abuse of dominance cases but also many vertical restraints. For an excellent overview of the application of economics in a century of antitrust enforcement in US see Kovacic and Shapiro (2000). As Gavil (2008) notes, after the \textit{Sylvania} decision in US “the Court systematically went about the task of dismantling many of the per se rules….., and increasingly turned to modern economic theory to inform its interpretation and application of the Sherman Act”. See also Hovenkamp (2018) for a thorough and very thoughtful review on the rule of reason LS. Neven (2006) reviews the situation in EU, identifying low levels of economic analysis, especially in abuse of dominance cases. Geradin & Petit (2010) note that under a presumption of illegality, the assessment of such cases in the EU has relied on “old, formalistic legal appraisal standards, and (has shown) a reluctance to endorse a modern economic approach”. See also, Gual and Mas (2011), Papandropoulos (2010), Marsden (2010), Wilson (2014), Rey and Venit (2015), Peepcorn (2015) and for a recent extensive review Ibanez Colomo (2016). But see also Katsoulacos and Makri (2020) that show that there has been a systematic and \textit{substantial move towards effects-based} in the DGCOMP decisions in the last two decades.

\textsuperscript{63} See the empirical findings of Katsoulacos, Avdasheva, Benetatou, Golovanova, Makri (2020) covering France, Greece and Russia as well as the EC.

\textsuperscript{64} Federico, Scott Morton and Shapiro ( ) argue, in the context of assessing effects on innovation to adopt the disanvantaging

\textsuperscript{65} For a detailed analysis of the relationship between legal and substantive standards see Katsoulacos (2019a).
incentive or ability of rivals to innovate. This is justified because economists understand that when the incentives for a certain activity (such as investing in R&D) decline, profit-maximizing firms will predictably engage in less of that activity. Thus, if one has persuasive evidence of reduced incentives to develop new products, one can reliably conclude that innovation will be lessened. A more timid antitrust policy would require evidence showing that rivals actually reduced their R&D on particular projects as a result of the challenged conduct, and that this reduction harmed customers because certain specific products were not developed. Proving those elements would be highly impossible in many dynamic markets. If the evidentiary burden of proof is set too high, antitrust enforcement will be ineffective in dynamic, innovation industries’.

Further:

‘Exclusion of a disruptive entrant inherently harms the competitive process, even if that disruptive entrant is (currently) less efficient than the dominant firm. Indeed, that pattern tends to be the norm in industries subject to significant economies of scale (e.g., due to network effects and/or learning by doing). Disrupters that are less efficient at the outset than the established dominant firm can still pose a grave competitive threat to the incumbent, because they have countervailing characteristics that appeal to consumers, or their efficiencies will improve as they gain experience and scale. Regardless of an entrant’s current level of efficiency, the competitive process requires they not be squashed by conduct that does not constitute competition on merit, which can include conduct that would not make economic sense if not for its exclusionary impact on competitors’.

On the other hand, a detailed empirical analysis of the extent and type of economic analysis applied in the assessment of abuse of dominance cases by DGCOMP (rather than the economic analysis utilised in the assessment of the appealed decisions of these cases by the EU Courts), Katsoulacos and Makri (2020) show that there has been a systematic and substantial move towards effects-based in the DGCOMP decisions in the last two decades.

We have argued that whether or not a conduct can be characterised as PI or PL at investigation stage i depends on whether $\tilde{h}_i = \gamma_i H - (1 - \gamma_i)B$ is positive or negative, respectively. For all the reasons, mentioned by Hovenkamp (2021), there are no good reasons for claiming that H is lower than B and, at least in markets in which firms have been shown to have significant market power (in stage 1), that $\gamma_i$ is not sufficiently high to make $\tilde{h}_i > 0$ and hence the conduct PI, in the high-tech digital markets. This indeed has been the presumption adopted by CAs in EU and throughout the world with the notable exception of US until a few years ago.

Further, as noted under Lemma 2 above, there is no reason to think that the Easterbrook (1984) hypothesis is likely to be valid (i.e., that $DEC FC_i > DEC FA_i$), as we have
commented also in Section 4. Specifically for the digital markets the Stigler Committee Report (2019) argues:

“While some markets may self-correct (making the cost of false acquitals small), the findings of this report suggest that rapid self-correction in markets dominated by large digital platforms is unlikely, and that harms to economic welfare from the exercise of market power in such markets are substantial. As discussed above, entrants find it difficult to overcome the high barriers to take on digital platform incumbents. Economies of scale, economies of scope, network effects, and negligible marginal cost all work together to make entry difficult in existing markets. Moreover, while monopoly profits are a lure to competitors, incumbents can use those very profits to entrench themselves and protect their position. No matter how dynamic the technology, an entrant will not unseat a monopolist if the monopolist is permitted to buy the dynamic entrant for a share of monopoly profits. Both parties gain from such a transaction—and the public loses. The result is less entry than a more competitive environment would create. Less entry into digital markets means fewer choices for consumers, stunted development of alternative paths of innovation, higher prices, and lower quality. Self-correction is not a realistic expectation in this environment—indeed, the available evidence suggests it has not happened—and public policy should not rely exclusively on it. Effective antitrust enforcement and regulation must take account of this reality. If there is a force toward self-correction, it may require active promotion to succeed, and in this way public intervention can be complementary rather than antagonistic to market forces. Indeed, the other reports that have addressed this problem around the world have accepted that policy changes are necessary in order to avoid stagnant and harmful digital markets. We now turn to what policy options exist, and which of these the government might adopt.”

For an opposing viewpoint we cite G Manne (2020) who even argues that:

“The concern with avoiding Type I errors (false convictions) is even more significant in the enforcement of antitrust in the digital economy because the “twin problems of likelihood and costs of erroneous antitrust enforcement are magnified in the face of innovation. Because erroneous interventions against innovation and the business models used to deploy it threaten to deter subsequent innovation and the deployment of innovation in novel settings, both the likelihood and social cost of false positives are increased in digital and other innovative markets. Thus the avoidance of error costs in these markets also raises the related question of the proper implementation of dynamic analysis in antitrust.”

Further, “..it is now clear that the market self-correction claim ignores the fact that the exclusionary conduct itself raises artificial barriers to entry”. Salop (2021).

Page 60.

Pages 3 – 4, Manne (p. 35). See also his analysis in Section D claiming that the Easterbrook hypothesis is right and providing a critique of Hovenkamp and Scott Morton (2020).
Once it is determined whether the conduct is PL or PI when carried out by dominant firms, the main issue then is whether, on decision-error minimising grounds, we should rely in order to reach decisions on an effects-based approach that relies on investigations also in stage 3 and 4 or we should rely on presumptions formed following the investigation in some early stage. In particular, it has been argued that we should rely on presumptions formed in stage 2: once exclusion is shown then we can presume liability and not undertake additional investigations (to assess harm to consumer welfare, efficiencies and balancing), though we may allow the defendants to rebut. As Cremer et al (2019) argue (p. 4):

“... in the context of highly concentrated markets characterised by strong network effects and high barriers to entry (i.e. not easily corrected by markets themselves), one may want to err on the side of disallowing potentially anti-competitive conducts, and impose on the incumbent the burden of proof for showing the pro-competitiveness of its conduct. This may be true especially where dominant platforms try to expand into neighbouring markets, thereby growing into digital ecosystems, which become ever more difficult for users to leave. In such cases, there may be, for example, a presumption in favour of a duty to ensure interoperability. Such a presumption may also be justified where dominant platforms control specific competitively relevant sets of user or aggregated data that competitors cannot reproduce”⁶⁹.

The argument by Fedrico, Scott Morton and Shapiro (2020), outlined in the quote above is in the same spirit, though the frame their argument as proposal to replace the consumer welfare substantive standard by a distorting the competitive process substantive standard. Khan (2020) is more explicit, arguing that the consumer surplus standard has crippled antitrust enforcement by prompting “a shift away from per se rules and towards rule of reason analysis”.

Our model can be used to answer the question: do conditions in the hi tech digital markets make it more likely that decision error minimisation is satisfied if, following the stage 2 investigation and showing that conduct is exclusionary, the CA uses a presumption-based discriminatory LS and bans the conduct or should it proceed with additional investigations 3 and 4 (i.e. adopt full effects – based)? We need to examine the effect on DECs of taking these additional assessment steps due to the changes in the probabilities \( \gamma, \hat{\beta}, p_H, p_B, \hat{p}_H \) and \( \hat{p}_B \).

One first remark is that for big-tech digital markets plaintiffs may forego investigation in stage 3 and consider whether DECs would be lowered by analysing defendants’ efficiency arguments and undertaking a balancing test in stage 4. One strong reason for this is that the justification for a stage 3 investigations is that as Salop (2017)⁷⁰ had argued

---

⁶⁹ Page 4, authors’ emphasis.
⁷⁰ Page 378.
“harm to competition and consumers from (raising rival costs\(^{71}\)) foreclosure conduct is not inevitable.... Competitors may not be significantly disadvantaged. In the case of input foreclosure, they may be able to substitute to alternative cost-effective inputs. In the case of customer foreclosure, they may have a sufficient number of alternative customers or distributors to remain a strong competitive constraint. In addition, even if certain targeted rivals are disadvantaged, there may be sufficient competition from other non-excluded competitors or substitute products to prevent the defendant from raising prices in the output market. Thus, ........ the plaintiff must prove “power over price” (i.e., probable harm to competition as well as “raising rivals’ costs” (i.e., harm to competitors))."

It is not unreasonable to argue that these considerations are unlikely to be of much relevance in the case of the big-tech digital markets. As already noted above, firms in these markets are generally characterised by extreme economies of scale and network externalities (creating very highly concentrated markets), the role of data (discriminatory access to data and lack of data portability), lack of interoperability, all of which act as a significant barrier to entry, often lack of multi-homing, and very significant economies of scope (that explain the emergence and growth of ecosystems and give incumbents a strong incumbency advantage). All of these may give dominant digital firms strong incentives to engage in anticompetitive behaviour\(^{72}\). Under these circumstances, it is very likely that exclusion (or foreclosure to use Salop’s terminology) will imply harm to competition and consumers (especially taking into account all aspects that are highly relevant to consumer welfare such as variety, access to new innovative products or services, privacy and quality) unless there are significant efficiencies that would be lost by forbidding this behaviour. Even if Salop’s qualifying factors may be of relevance some cases, it would be best in general for digital markets to integrate stages 3 and 4 investigations and consider whether there will be a cost of errors improvement for these markets to adopt rules that rely on assessment of all the stages, concentrating after stage 2 on stage 4. There may be substantial efficiencies: after all economies of scale and network could well imply that from a static efficiency perspective, concentration is desirable\(^{73}\). On the other hand, the incumbency advantage could prevent a superior platform from overtaking an inferior one – this will depend “on a number of factors including the possibility of multi-homing and....data portability as well as data and protocol interoperability”\(^{74}\).

Given this, below we analyse whether in general we should rely on a presumption-based standard that relies on showing exclusion – assessment stages 0 - 2, or a full effects based.

A starting point is whether condition (24) is more likely not to hold in the big tech digital markets

\[
\frac{H}{B} \leq \hat{p}_H(1- \hat{p}_B) = \sigma \text{ (condition (24) rewritten)}
\]

\(^{71}\) Not predatory foreclosure.

\(^{72}\) For an excellent discussion of all these, on which we have relied here, see the Report by Cremer et al (2019).

\(^{73}\) Cremer et.al. (2019) p.23. They refer to the work of Weyl and Whitle (2014) according to which “competition will lead to too little concentration”.

\(^{74}\) Cremer et.al. (2019) p.23.
i.e. it is more likely that $\frac{H}{B} > \frac{\hat{p}_H(1-p_B)}{(1-\hat{p}_{HH})} = \sigma$.

Then, additional investigations will increase DECs as $\gamma_i$ increases. This would indeed be the case if there was reason to believe that (H/B) is on average higher in such markets, that $\hat{p}_H$ is lower, that $p_B$ is higher and that $p_H$ is lower. It is likely that this will hold for H/B (as already explained above) and for $\hat{p}_H$ (the likelihood of identifying correctly exclusion is likely to be higher for the big tech platforms). There are no strong reasons however, for arguing that $p_B$ is higher and that $p_H$ is lower in the big-tech digital than in other markets in stage 2, when the investigation concerns identifying whether exclusionary effects are harmful or benign, indeed the opposite may well be the case. Thus, the different influencing factors pull in opposite directions and we cannot conclude that condition (24) is more likely not to hold, making more likely an increase in DECs, in the big tech digital markets.

The probability ($\hat{\beta}_i$) that a screen is satisfied given that previous screens are satisfied, which becomes lower for additional screens. From (22) and (23) we see that this unambiguously reduces costs of FAs and certainly also reduces costs of FCs if $\hat{p}_{H,i}(1 - \gamma_i) - (1 - \hat{p}_{B,i}) > 0$. Though $\hat{p}_{H,i} + \hat{p}_{B,i} > 1$ this condition will not hold if $\gamma_i$ is very high. However, as we noted above, the impact on DECs also depends on the combined effects of $\hat{\beta}$ and $\gamma$. As we showed, the probability $\hat{\beta}(1 - \gamma)$, that the precondition examined holds but the conduct is not harmful falls, and this reduces FCs. However the probability $\hat{\beta}\gamma$, that the precondition examined holds but the conduct is harmful, may fall or rise so, from (22'), the effect on FAs is ambiguous. In many cases $\hat{\beta}\gamma$ may rise as when in stage i+1, the probability $\hat{\beta}$ does not fall much as will be true when, between for example stages 2 and 4, the probability does not fall much because in cases where exclusion is identified the conduct by the dominant firms will be very likely to have consumer welfare reducing effects, after accounting for efficiencies. It is not obvious that this will be true in the big-tech digital markets in which efficiencies may be significant. This means that the effect on FAs is ambiguous and hence the overall effect on DECs is again ambiguous, given that the probability of FAs is weighted by H while that of FCs by B, so when the probability of FAs goes down this this will outweigh an equal increase in the probability of FCs.

The parameters $p_H$ and $p_B$ will increase as additional investigative steps are undertaken and additional preconditions are found to be satisfied, reflecting the fact that the accuracy of the estimate about whether conduct is harmful or benign improves as more analyses and tests are applied in assessing additional preconditions / screens. As is clear from (22) and (23) this lowers FAs and FCs and hence reduces DECs.

With regard to the probabilities $\hat{p}_H$ and $\hat{p}_B$, we have showed above that given that (26) holds, DECs will be unambiguously be reduced if $\hat{p}_H$ and $\hat{p}_B$ increase. However, these probabilities, that measure the accuracy of estimates of whether or not different preconditions examined hold may rise or fall. These probabilities are indeed likely to fall and be very low in stage 4 in the big tech digital markets, as many commentators have argued, since it is likely to be difficult to identify correctly that precondition 4 is satisfied (i.e. that there are not important efficiencies) when it is satisfied and, especially, it is likely to be very difficult to identify
correctly that precondition 4 is not satisfied (i.e. that there are very important efficiencies that outweigh anticompetitive effects) when it is not satisfied.

Of course, even though \( \hat{p}_H \) and \( \hat{p}_B \) decrease we note that what matters for FAs is what happens to \( \hat{p}_H p_H \) and what matters for FCs is what happens to \((1 − p_B) \hat{p}_H \) and \((1 − \hat{p}_B)(1 − p_B)\), where \( p_H \) and \( p_B \) will increase. However, (30) is less likely to hold in the markets of interest here than in other markets, which implies an increase in DECs of FAs is more likely in the former markets than in the latter. Further, (33) may not hold in the big tech digital markets because \( \hat{p}_{B,3} > \hat{p}_{B,4} \), given also that the first term on the RHS of (31) is likely to be very small as it involves the product of 4 probabilities each less than one, one of which (\( \hat{p}_{3} \)) is itself the product of 3 probabilities. This creates a tendency for the DECs of FCs to rise with assessment in stage 4. To conclude, the DECs of full effects based assessment is likely to be higher in the big tech digital markets than in other markets, making full effects based less likely to be optimal, as a result of the decrease in the proababilities \( \hat{p}_H \) and \( \hat{p}_B \).

Overall conclusion: the effects of the change in the probabilities \( \gamma \) and \( \hat{\beta} \), is ambiguous; the effect of the change in \( p_H \) and \( p_B \) reduce DECs, while the effect of the change in \( \hat{p}_H \) and \( \hat{p}_B \) is likely to increase DECs. Thus, it is impossible to say a priori whether moving to rule of reason will increase or reduce the DECs relative to a presumption-based LS based on evidence of exclusion for the big-tech digital platforms, though it is certain that this move will increase considerably enforcement costs and make litigation more protracted.

7. Concluding remarks

We have reviewed the literature, especially all the recent developments, in the choice of LSs, and provided a detailed presentation of a new methodology for defining legal standards along the continuum of legal standards (Arreda and Hovenkamp, 2017; Jones and Kovacic, 2017; Katsoulacos et al., 2020 and 2021; Katsoulacos and Ulph, 2022), depending on the screens or preconditions assessed at different stages of the continuum. We also provided a detailed formal examination of how all the pertinent factors that could influence error minimisation interact to determine the optimal error-minimising LSs for different conducts and markets. And, we used this framework to examine how the choice of error minimising LSs depends on the context in which different conducts are undertaken. Specifically, how this choice is affected when comparing developing jurisdiction/countries to developed countries/jurisdictions; and, how the choice is affected when the conduct is undertaken in digital multi-sided platforms. This paper suggests that cost of decision error principles can be used to provide a useful and practical framework for analysing the choice of legal standards for specific categories of conduct in competition law enforcement.

We have shown that in developing countries/jurisdictions the error-minimising choices are more likely to be closer to Per Se than to effects-based, than in developed countries/jurisdictions. The choice is more ambiguous when we compare the big-tech digital platforms as a category to other markets.
Appendix: numerical examples

Assume that dominant firms make up 10% of the population (so to undertake the conduct it is not necessary to possess significant market power), cases with exclusionary effects by dominant firms are 40%, cases where exclusionary effects are associated with consumer welfare loss if no account is taken of efficiencies are 70% and cases with no significant efficiencies to outweigh the loss in consumer welfare are 50%. That is:

\[ \beta_1 = 0.1, \beta_2 = 0.4, \beta_3 = 0.7, \beta_4 = 0.5 \]

So:

\[ \hat{\beta}_1 = 0.1, \hat{\beta}_2 = 0.1 \times 0.4 = 0.04, \hat{\beta}_3 = 0.1 \times 0.4 \times 0.7 = 0.028, \]
\[ \hat{\beta}_4 = 0.1 \times 0.4 \times 0.7 \times 0.5 = 0.014; \]

And so:

\[ (1 - \beta_1) = 0.9, (1 - \beta_2) = 0.96, (1 - \beta_3) = 0.972, (1 - \beta_4) = 0.986 \]

\[ y_3 = \beta_4 = 0.5 \text{ and } y_0 = \hat{\beta}_4 = 0.014 \text{ so less than 1.5% of cases of the conduct type are harmful. This here we focus on a conduct type for which the expected fraction of harmful cases is very small, which does not favour the use of effects-based} \]

Assume also that, discriminating power is not very high:

\[ \hat{p}_{H,3} = \hat{p}_{B,3} = p_{H,3} = p_{B,3} = 0.6 = \hat{p}_{B,A} = \hat{p}_{H,A} \]

So

\[ (1 - \hat{p}_{B,3})(1 - p_{B,3}) = (1 - \hat{p}_{B,A})(1 - p_{B,A}) \quad \text{(condition 33 as equality)} \]

So from (28), (29) and (30), (32):

\[ DEC FA_3 = 0.00896H \]
\[ DEC FA_4 = 0.0056H \]

So there is a 37.5% reduction in the DEC of FAs from the last step.

\[ DEC FC_3 = (0.00336 + 0.1552)B = 0.15888B \]
\[ DEC FC_4 = 0.15776B \]

So there is a very decrease in the DEC of FCs from the last step.

In order to compare overall DECs, we need to assign values to H and B. To do that we note that in order for the conduct to be PL in stage 0 and PI in stage 1, given the values of the parameters above, the value of H must be higher than 6.17B. Let us assume that H = 7B. Then:

\[ DEC_3 = 0.00896 \times (7B) + 0.15888B = 0.2216B \]
\[ DEC_4 = 0.0056 \times (7B) + 0.15776B = 0.18696B \]
So the last investigative step reduces DEC by about 11.2%. While the last step is preferred in terms of minimising DECs, given the additional implementation cost of the last step it may seem optimal not to undertake this step. If CAs use a rule that DECs must decrease by at least 10% to justify the extra enforcement cost of taking an additional step then in example above the CA should undertake the step from stage 3 to stage 4.

Comment – another scenario:

For some conducts a more reasonable set of parameter values would be:

\[ \beta_1 = 0.5, \beta_2 = 0.7, \beta_3 = 0.9, \beta_4 = 0.8 \]

(or even higher values of \( \beta_1 \).

In this case:

\[ \gamma_3 = \beta_4 = 0.8, \gamma_1 = 0.504 \]

and \( \gamma_0 = \hat{\beta}_4 = 0.252 \)

In this case, the conduct is PL in stage 0 and PI in stage 1 with approx. equal H and B. Assume that the values of the discriminating parameters are as before.

Then:

\[ \text{DEC } FA_3 = 0.16128H \]
\[ \text{DEC } FA_4 = 0.1008H \]

So there is the same 37.5% reduction in the DEC of FAs from the last step. This is as we expect since \( \hat{\beta}_3 \gamma_3 = \hat{\beta}_4 \) and the discriminating parameters are the same. Now:

\[ \text{DEC } FC_3 = (0.01512+0.115648)B = 0.2130768B \]
\[ \text{DEC } FC_4 = 0.11872B \]

So there is a 44.3% decrease in the DEC of FCs from the last step.

We compare overall DECs assuming that H = B (having H been much larger or even larger than B may appear unreasonable to some that think that false convictions are the important errors). Then:

\[ \text{DEC}_3 = 0.16128B + 0.2130768B = 0.37435B \]
\[ \text{DEC}_4 = 0.1008B + 0.11872B = 0.21952B \]

So the last investigative step reduces DEC significantly by about 41.3%.

Increasing the value of the discriminating parameters

It is easily seen using the above examples that such increases lead to significant reductions in DECs when an additional investigative step is undertaken.
References

ABA REPORT OF THE TASK FORCE ON THE FUTURE OF COMPETITION LAW STANDARDS, 2020
Hovenkamp, Herbert J. and Fiona Scot Morton (2019). “Framing the Chicago School of Antitrust Analysis”. DP


