

# **The role of economics and the quality of antitrust enforcement in UK: an empirical analysis of the CMA/OFT antitrust cases<sup>1</sup>**

(Short title: “Antitrust enforcement in UK: the role of economics”)

Vasiliki Bageri<sup>2</sup> and Yannis Katsoulacos<sup>3</sup>

## **Abstract**

In this paper we empirically investigate the role of economic analysis and evidence used by the Competition and Markets Authority (CMA) (and the Office of Fair Trading, before CMA) in reaching antitrust infringement decisions, from 2000 to 2020. The study uses a novel methodology to construct indicators that measure the extent to which economic analysis is utilised and its deviation from the optimal level, aiming to understand its role in CMA's decision-making process and the quality of the Authority's enforcement. Our analysis offers insights into how economic analysis shapes antitrust enforcement outcomes and compares in detail the quality of enforcement, across different conduct categories and overall, between CMA and DGCOMP. We find that CMA's enforcement record exceeds in quality that of DGCOMP and of other EU and BRICS authorities.

**Keywords:** antitrust, economic analysis, legal standards, Per Se, effects-based, enforcement quality

**JEL:** L4, K21

---

<sup>1</sup> We are grateful for comments from participants of the CRESSE Conference, July 2023 and Shandong University Conference, September 2023, and of seminars at the Toulouse School of Economics, CITY University, London, ESSEC Business School, OECD Competition Division, Dept. of Economics, Universidad de la Republica, Uruguay, Higher School of Economics, Moscow, Shanghai University of Finance and Economics in which the methodology adopted in this paper has been presented. We are particularly grateful for comments by Svetlana Avdasheva, Antonio Copabianco, Jacques Cremer, Pablo Ibanez Colomo, Xenia Dassiou, Svetlana Golovanova, Antonio Gomez, Frederic Jenny, Massimo Motta, Lin Ping, Eduardo Ribeiro and Patrick Rey. We are also grateful for financial support to the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the “1st Call for H.F.R.I. Research Projects to support Faculty Members & Researchers and the Procurement of High-and the procurement of high-cost research equipment grant” (Project Number: 1314).

<sup>2</sup> Assist. Prof. of Economics, Panteion University of Social and Political Sciences, Athens.

<sup>3</sup> Y. Katsoulacos, Emeritus Professor of Economics, Athens University of Economics and Business; Affiliated Chair Professor, Jiangxi University of Finance and Economics, China (corresponding author: e-mail: yanniskatsoulacos@gmail.com).

## 1. Introduction

Following a long period during which most business practices associated with horizontal or vertical agreements and potentially exclusionary behaviour by firms with significant market power, were treated under a Per Se Illegality legal rule, the assessment standard in the United States, under the influence of the Chicago School, shifted drastically from the late 1970s onwards, starting to treat most practices other than hard-core horizontal agreements as presumptively legal (that is, in economists' terms, as on average benign) under the Easterbrook Hypothesis (1984). This shift was associated with an increased emphasis on using a net consumer harm liability standard<sup>4</sup>. The proponents of this shift argued that it was necessary for effective enforcement not to be based on vague, subjective political and policy goals that second guessed the market and ignored consumer benefits (see for recent brief reviews Katsoulacos and Ulph, 2022; Katsoulacos, 2023, 2023a; Fox, 2023 and Kovacic, 2021). This approach has led to a very high rate of acquittals<sup>5</sup> (and has been particularly criticized with respect to enforcement in the high-tech digital markets).

This has not been the case in the EU which would typically treat potentially exclusionary practices and vertical restraints in the past and currently as presumptively illegal. While in the US, the dominant view has been, at least since the 1980s, that a primary objective in antitrust enforcement is to limit false convictions rather than false acquittals – a view that has only started to be strongly criticized recently<sup>6</sup> – this view has never been shared in the EC and EU. Further, the shift to the economics-based approach came much later and it has influenced enforcement very slowly. Indeed, the European Commission's move towards a more economics-based approach for practices by dominant firms has not been formalized until 2008 with the publication of the Commission's Guidance Paper (2008) and the European Courts have only very recently endorsed this approach<sup>7</sup>.

Finally, the substantive or liability standard in EC and EU has been more that of “disadvantaging rivals” than “consumer welfare” in the EU<sup>8</sup> - and this naturally makes the choice of LS what we

---

<sup>4</sup> Though as J Kanter (2022) notes “the legislated goals of the antitrust laws are clear – Congress sought to protect and the competitive process”.

<sup>5</sup> For empirical evidence see two articles in 1999 and in 2009 by Michael Carrier. He reports the results of rule of reason (r of r) assessments from 1977 – 1999 and then from 1999 – 2009 (717 rule of reason cases in total) in US. He confirms the use of the burden shifting framework in these cases (arguing that this proves that RoR is far less amorphous than is sometimes claimed). He also concludes that “the courts almost never balance. And, they dispose of 97% of r of r cases on the grounds that the plaintiff cannot show anticompetitive effect”.

<sup>6</sup> See in particular the detailed and very careful arguments of Baker (2015) and Hovenkamp (2021). Also Hovenkamp and Scott Morton (2019, p. 19-20).

<sup>7</sup> See Witt Anne (2019). See also the recent (2023) European Court of Justice judgment in the Unilever Italia case.

<sup>8</sup> This was certainly true prior to the EC's switch to a more economics-based approach in mid-2000s, but also for some years, following the publication of EC's Guidance Paper on art. 102, in 2008, during which there was an inconsistency between policy documents and competition law enforcement standard. ‘Consumer welfare’ was often cited in policy documents but it was not the standard adopted in EU competition law enforcement. Interestingly, Giulio Federico, Fiona Scott Morton, Carl Shapiro (2020) argue, in the context of assessing effects on innovation in favor of adopting the disadvantaging rivals” or “distorting the competitive process” SS.

have called Truncated Effects-Based LS<sup>9</sup> that is based, in order to reach liability decisions, on just evidence of exclusion of (or disadvantaging) rivals<sup>10</sup>. It is worth noting that Lina Khan (2018) and Tim Wu (2018) as well as other neo-Brandeisians have been proposing the adoption of a non-welfarist substantive (liability) standard in the US too<sup>11</sup>, while the movement away from object-based identified in Katsoulacos and Makri (2020) in the EC indicate that in the latter increased emphasis is placed on consumer harm, as also noted by Caffara (2023). These developments suggest that enforcement procedures may be converging in recent years on the two sides of the Atlantic with respect to conducts that are considered on both sides as presumptively illegal.

Yet, even in cases where competition practice shows an inclination for economic inputs and dedicated empirical economic work, there remains a strong concern with the significant increase in enforcement costs, the increase in legal uncertainty and the potential underenforcement that a rigorous economic approach to competition law enforcement implies<sup>12</sup>.

In this paper, we empirically investigate the role of economics in antitrust enforcement for the case of one of the oldest and, widely considered as one of the, most sophisticated CAs in Europe, UK's CMA. We achieve that by constructing and measuring indicators capturing the quantity of economic evidence used by the agency in reaching infringement decisions in the period 2000-2020. Moreover, by measuring the deviation of CMA's overall enforcement performance from the optimum we capture the *quality of enforcement*. Subsequently, we construct an indicator that measures the "total" economic analysis undertaken in the assessment of the cases "irrespective of burden of proof", i.e. an indicator that also incorporates the *efficiency arguments*.

For the measurement of the indicators we rely on a database that we constructed of antitrust infringement decisions made by the CMA, using an updated version of the methodology originally developed by Katsoulacos, Avdasheva and Golovanova (2018) and revised by Katsoulacos and Ulph (2022) and Katsoulacos (2023, 2023a). We also utilize the dataset in order to identify the legal standards (LSs) adopted by the agency when assessing different conducts.

The dataset distinguishes decisions according to the main conduct type categories associated with the enforcement of Competition Law, incorporates the more recent enforcement categories of commitments and settlements and traces out the Commission's sanctioning practices. It contains 54 such decisions reached in the period 2000–2020. It also contains information about the decisions

---

<sup>9</sup> See for example, Katsoulacos (2023) and, for further explanation, below.

<sup>10</sup> For a an extensive analysis on the relationship between legal and substantive standards, see also Katsoulacos (2019a) and Bageri et.al. (2023).

<sup>11</sup> For detailed criticisms of these views by strong adherents to the consumer welfare standard, see Melamed A.D and N. Petit (2019) and Manne Geoffrey (2020).

<sup>12</sup> See Edward Cavanagh (2023). He notes that "the fact is that plaintiffs rarely win antitrust cases under the r of r. Not that they are never successful. They have been occasionally victorious but the victories amount to a few grains of sand amid a beach dominated be judgements for the defendants. Also, Whish (2023) Keynote Lecture 17<sup>th</sup> CRESSE Conference, July, 2023 (in Keynote Presentations in [www.cresse.info](http://www.cresse.info)).

appealed and about whether or not the decisions were finally annulled by the Competition Appeal Tribunal or the Court of Appeal.

Our main objectives have been to:

- (i) Examine the LSs adopted in the assessment of different conduct categories by CMA.
- (ii) Identify the extent and type of economic analysis utilised by the agency to reach decisions, their evolution over time, and the extent to which they diverge from their optimal level (see below).
- (iii) Propose and then measure, indicators of the “quality of enforcement of competition law” (meaning the *quality of assessment* of the antitrust decisions reached) by CMA.
- (iv) Identify the consistency with which specific blocks of economic analysis are used for each conduct type and hence the extent to which CMA’s choices contribute to legal certainty.
- (v) Construct an indicator that incorporates all the economic arguments utilised in the assessment of a decision including those that are claimed by the defendants (efficiency analysis).
- (vi) Compare the above indicators with those obtained for EC (Katsoulacos and Makri, 2020; 2023).

The structure of the paper is as follows. First, our antitrust decisions database is used to identify the LSs adopted in the assessment of the infringement decisions. We distinguish among 8 LSs each one corresponding to a specific level of economic analysis applied in the assessment procedure (including efficiency arguments), as explained in the next section. Then we construct Cumulative Economic Analysis Indicators (CEAI) associated with each decision i.e. indicators that measure the quantity of economic analysis used by the agency in examining anticompetitive effects. We distinguish among four CEAI depending on the extent and type of economic analysis utilised. Decisions are categorised into four main conduct types or groups: hard-core horizontal agreements (G1), other horizontal agreements and concerted practices (G2), vertical agreements and restraints (G3) and abuse of dominance practices (G4). Next, we develop a number of indices. We measure the Weighted Cumulative Economic Analysis Indicators (WACEAI) adopted for each conduct group (the weights being the share of each CEAI used in assessing decisions for each conduct), and the degree of concentration of the CEAI of each conduct<sup>13</sup>. The higher the concentration on specific CEAI when assessing specific conducts the greater the certainty with which it can be anticipated that a specific amount of economic analysis will be used in the future for a conduct.

We also present *indices of the quality of enforcement* measuring the extent of deviation of WACEAI from its optimal level for each conduct group relative to the maximum theoretical deviation, as well as an overall index of quality of enforcement by CMA across all conduct groups.

Finally, we calculate an indicator that captures the Total Economic Evidence (TEE) considered on average during the assessment of decisions in a given conduct group, irrespective of the burden of proof, i.e. taking into account also efficiency defence arguments by the defendants. Therefore,

---

<sup>13</sup> The same indicators were constructed for the EC in Katsoulacos and Makri (2020; 2023).

TEE is measured as the weighted average of the LSs adopted in the assessment of the decisions of a particular conduct group.

Our results demonstrate that economic analysis played a key role in investigations conducted by the Competition and Markets Authority (CMA) during the studied period. The research suggests that CMA's enforcement quality, evaluated by comparing actual outcomes to socially optimal outcomes using the CEAI index, is exceptionally high with an overall score of 0.91. These results surpass or significantly exceed those of DGCOMP, assessed using the same methodology (and even more so, the results obtained in other EU countries or the BRICS).

## **2. A methodology for identifying the extent of economic analysis and legal standards in antitrust enforcement**

The methodology outlined below, following Katsoulacos (2023), makes use of an old idea, specifically, that there is a *continuum* of legal standards. The idea originated in the writings of legal scholars (Phil Areeda, Herbert Hovenkamp and Bill Kovacic, in particular) and judges (Judge Stevens), some time ago. It was present in the first edition of the “Antitrust Law treatise” (as mentioned in Areeda and Hovenkamp, 2017). The idea’s articulation closest to our approach is in Jones and Kovacic (2017). Other references include Gavil (2008 and 2012), Kovacic (2021) and Italianer (2013). The latter mentions that the US Supreme Court has explicitly recognised that “the categories of analysis cannot be 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”. It is worth mentioning here that the use of the idea of a continuum of LSs also clarifies significantly the relationship between the notions of legal and of substantive standards (see Katsoulacos, 2019a).

Our approach can be summarised in the following way. We define the continuum of LSs as: 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 objective of each of the steps 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 / lack of contestability, potential for exclusion or maintaining market power, potential for consumer harm and potential for efficiencies. Note that, in principle, decision error costs across steps or stages can be derived and compared in order to determine the optimal number of stages, that defines the error-minimising LS.

Given that in the literature a number of different terms are used, often very confusingly, in discussions of legal rules and the standard of proof, we should stress to avoid confusion that here, we use the term “legal standard” to represent the “quantity” of evidence, as measured by the necessary number and type of screens needed to be examined for proof of liability. It is also sometimes referred to in the literature as the *evidentiary standard of proof*. This is a very important

dimension of the notion of the “*standard of proof*”<sup>14</sup>. However, the latter entails *also* the dimension of the “*level of certainty necessary for proof*”, if, as assumed here, the objective is to minimise the cost of decision errors. Often authors refer to the standard of proof without clarifying which of the two dimensions they have in mind. We think it is best to distinguish between these two dimensions since to any given level of certainty necessary for proof there may be associated, depending on the case, different legal or evidentiary standards<sup>15</sup>. In Katsoulacos and Ulph (2022) we discuss in more detail how the second dimension of the standard of proof can be incorporated in the analysis of LSs and its influence on the choice of the optimal (error-minimising) LS<sup>16</sup>.

### 3. Potential investigation stages, screens and LSs

A Competition Authority (CA) can assess 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. Given that the ultimate objective (i.e. the SS) is to identify whether the conduct is harmful to consumer welfare<sup>17</sup>, 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<sup>18</sup>:

**S1:** Conduct characterisation screen.

**S2:** (a) Market contextualisation and (b) when relevant, Significant Market Power (SMP) /contestability screen.

**S3:** Potential for significant exclusionary impact or enhanced ability to exercise / maintain market power screen.

**S4:** Potential consumer welfare loss, due to just anticompetitive effects, screen. Examination of potential effects on output, prices, quality, variety and innovation.

**S5:** Efficiencies assessment and balancing screen.

---

<sup>14</sup> See for example Merriam-Webster Legal.

<sup>15</sup> E.g. to a high level of certainty (clear and convincing evidence or preponderance of the evidence) there may be associated different quantities of necessary evidence in different circumstances and for different conduct types, for proof of liability (depending, for example, on what economic theory directs us to presume with regard to what is the strength of the presumption of illegality associated with different quantities of evidence).

<sup>16</sup> For example, when the necessary level of certainty is considered satisfied, additional evidence will not be examined in practice even though this evidence could reduce further the costs of decision errors.

<sup>17</sup> That this is indeed the criterion adopted by CMA has been demonstrated in a parallel empirical study by Bageri, Katsoulacos and Metsiou (2023). In that study we examine the SS adopted by CMA in the decisions analysed in this paper. It is found that in 97% of the decisions the SS has been consumer welfare (in 3% the SS is “disadvantaging rivals”).

<sup>18</sup> We note that in order to focus the analysis we omit the examination of exploitative conduct, which would have affected stage 2.

### *A Typology of Legal Standards*

We treat each of 5 evidentiary screens identically and assign a value of 1 to the “amount of evidence produced by the screen” if the screen is assessed and a value of 0 otherwise. Obviously the total evidence has a maximum of 5 if all screens are assessed. We distinguish 8 Legal Standards to which we assign a value ranging from 1 to 5 depending on the number of the five potential screens that have been utilized which measure the extent of economic analysis used when a case is assessed. We then define the following legal standards:

- (i) Strict Per Se (SPS) LS: with this, the liability decision relies purely on the initial characterisation of the conduct (in stage 1) and on the presumption that this generates about its consumer welfare impact.
- (ii) Object – based or Modified Per Se LS (MPS LS): the liability decision then relies just on the information from stages 1 and 2a or from stages 1, 2a and 2b, about the conduct and the presumption that this generates about its welfare impact.
- (iii) Truncated Effects Based I LS (TEB I LS): the liability decision relies on the information from stages 1, 2 and 3 about the conduct and the presumption that this generates about its welfare impact.
- (iv) Truncated Effects Based II LS (TEB II LS): under this, a liability decision relies on the information from stages 1, 2, 3 and 4 about the conduct and the presumption that this generates about its final welfare impact.
- (v) Full Effects Based (or rule of reason) LS (FEB LS): under this, a liability decision relies on the information from all assessment stages 1, 2, 3, 4 and 5 about the conduct and a balancing between anticompetitive and efficiency effects to determine the net effect on consumer welfare.
- (vi) Quick Look I: a SPS LS where the agency, also, considers efficiency claims.
- (vii) Quick Look II: a TEB I LS where the agency, also, considers efficiency claims.
- (viii) Quick Look III: a TEB II LS where the agency, also, considers efficiency claims.

The above discussion is summarised in the Table below.

**Table M1: Identifying legal standards**

Screens examined in assessment	Legal Standard
S1	Strict Per Se (SPS)
S1 and S2	S1 and S2 (a): Object – based (EU) S1 and S2 Modified Per Se (MPS)
S1 and S2 and S3	Truncated Effects Based I (TEB I)

S1 and S2 and S3 and S4	Truncated Effects Based II (TEB II)
S1 and S2 and S3 and S4 and S5	Full Effects Based (FEB) or Rule of Reason
S1 and S5	Quick Look I
S1 and S2 and S5	Quick Look II
S1 and S2 and S3 and S5	Quick Look III

#### 4. Description of the dataset and of the conduct groups

Our dataset consists of all the antitrust infringement decisions by CMA, that is all the decisions that fall under Chapters I and II of the UK Competition Act 1998 (or Articles 101 and 102 of the EC Treaty). These laws prohibit anti-competitive agreements between undertakings (Chapter I) and the abuse of a dominant position by an undertaking (Chapter II). The period under investigation is 2000-2020 during which 54 antitrust decisions were reached (excluding decisions on exploitative conduct).

For a more detailed and targeted analysis of our results we have broken down the decisions mentioned above by conduct group (decision category) as follows:

- Conduct group G1: this consists of violations of Chapter I, which have strong market power-enhancing effects. They include price fixing, bid rigging, boycotts, market sharing and exclusive territories (or a mixture of these);
- Conduct group G2: this group is meant to cover concerted practices – all the decisions that were included here, involved price information exchange cases;
- Conduct group G3: this consists of vertical restraints that, according to conventional wisdom, can have both anticompetitive and welfare-enhancing effects; such as various types of resale price maintenance and other vertical agreements such as exclusive dealing or exclusive territories.
- Conduct group G4: this consists of practices by dominant firms (such as margin squeeze, price discriminations, loyalty rebates, exclusive contracts, tying and bundling and refusals to deal) that may have anticompetitive effects and hence considered abusive under Chapter II, though again economic theory suggests that generally many of these practices can also be motivated by efficiencies that have welfare-enhancing effects.

In Table 1 we present the total number of decisions per conduct group and indicate which of the decisions were appealed and which were subsequently annulled.

**Table 1: Number of decisions, appeals and annulments (by conduct group)**



G1	G2	G3	G4	Total
Total number of decisions (shares)				
27 (0,491)	7 (0,127)	15 (0,273)	5 (0,093)	54
Number of appealed decisions (shares)				
6 (0,222)	2 (0,286)	7 (0,467)	2 (0,400)	17 (0,315)
Number of annulled decisions (shares)				
0 (0,00)	0 (0,00)	1 (0,143)	0 (0,00)	1 (0,018)

As expected, hard-core horizontal agreements (conduct group G1) constitute the majority of the antitrust decisions in UK amounting for around 49%, followed by decisions on vertical restraints (conduct group G3) at around 27.3%. The number of decisions on concerted practices (G2) take up 12.7% of the total infringement decisions and abuse of dominance cases are just below 9.5 % (G4). Appeal rates appear to be higher for group G3 (46.7%), followed by G4 decisions (40%), however, a substantial percentage of group G2 (28.6%) and G1 (22.2%) were also appealed.

Interestingly, while quite a significant amount of infringement decisions is appealed (31.5%) an extremely small fraction of these is finally reversed by the CAT or the Court of Appeal (1.8%). Specifically, for the years examined only one decision has been annulled (a vertical restraint decision), indicating the more or less complete alignment between the CMA and the Competition Appeal Tribunal (CAT) and the Court of Appeal regarding the liability of the decisions. However, we should notice here that *in a large fraction of the appealed cases (i.e. 59%, or 10 out of the 17 appealed cases) the CAT has ordered a reduction of the fine originally set by the agency.*

The results regarding the appeal and annulment rates in UK are significantly different from other countries/jurisdictions. In EC the appeal and annulment rates are 72% and 44%, respectively between 1992 – 2021; in France 57% and 22%; in Greece 90% and 30%. In In Russia the percentages are also much higher than in UK - the annulment rate between 2008 – 2020 reaches 57% (Katsoulacos et.al., 2021). Only in China (2008 – 2020) is the annulment also very low as in UK (Katsoulacos and Zhixue Guo, 2023).

*To conclude this section, in UK, the appeal and annulment rates of infringement antitrust decisions are significantly lower relative to the EC, and other countries (other than China). One important factor for explaining this is the fact that CAT (the main appeal body in UK) is a specialized court (unlike the appeal courts in Europe and most other countries) in which decisions are influenced by high calibre academic and other economists that specialize in competition and the CMA decisions are of high quality in terms of their economic argumentation (as our empirical research indicates - below).*

## 5. Economic analysis and legal standards in CMA’s antitrust enforcement: results of the empirical analysis

### 5.1. The role and type of economic analysis applied

As a first measure of enforcement quality we examine which potential investigation stage is the highest in each decision. According to the data shown in Table 2, the highest level of economic analysis for most decisions (37 out of 54) is S2. This is not surprising given that about 62% of the decisions involve horizontal agreements (G1 and G2) for which it is optimal to consider just screens S1 and S2. However, we also notice the low scoring (11%) attributed to the efficiency analysis screen (S5), where the arguments should be presented by the defendants – we return to this below.

**Table 2: Highest level of economic analysis present in each decision**

	The highest level of economic analysis present in the decision					
	S1	S2	S3	S4	S5	Total
Total number of decisions	1	37	6	4	6	54
Shares	2%	69%	11%	7%	11%	

Continuing with the results of the empirical analysis, Table 3 presents the extent and type of economic analysis applied in the decisions per conduct group. As expected, in all of its cases, the CMA considers the nature and characteristics of the conducts (analysis component S1), while a basic analysis of the market characteristics of the case under investigation (S2) also characterizes essentially all decisions (98%). This could be attributed to the need to calculate the potential fines and to verify whether Articles 101 or 102 TFEU or both should be applied. On the other hand, a *more formal market definition based on more sophisticated economic tests (e.g. SSNIP test, Critical loss analysis) appears in only one case* (Table M2). The information required to undertake such tests may be the preventing factor here, since it usually is unavailable to the CA.

**Table 3. The role and type of economic analysis applied per conduct group.**

Type of Analysis Applied	S1=1	S2=1*	S3=1	S4=1	S5=1
Conduct group					
G1	27	26	2	0	1
% within group	100%	96%	7%	0%	4%
G2	7	7	1	0	1

% within group	100%	100%	14%	0%	14%
G3 (RPM)	13	13	4	1	1
% within group	100%	100%	31%	8%	8%
G3 (non-RPM)	2	2	2	1	2
% within group	100%	100%	100%	50%	100%
G4	5	5	5	4	1
% within group	100%	100%	100%	80%	17%
Total	54	53	14	6	6
	100%	98%	26%	11%	11%

**Table M2: Formal market delineation (e.g. SSNIP test, Critical loss analysis)**

<b>Conduct group</b>	<b>SNIPP test</b>
G1	0
% within group	0%
G2	0
% within group	0%
G3 (RPM)	0
% within group	0%
G3 (non-RPM)	0
% within group	0%
G4	1
% within group	17%
Total	1
% of Total	2%

Regarding the use of economic analysis components S3 and S4, Table 3 also provides useful observations. Starting with the analysis undertaken to identify if the conduct has market power enhancing or maintaining, or exclusionary effects (S3), we notice a significant difference in its usage

between conduct groups. Specifically, this analysis appears in a very small fraction of G1 and G2 cases while in group G4 appears in the total number of the cases. Conduct group G3, as expected, requires two different enforcement paths from the agency. When RPM cases are considered there is a low to intermediate usage of analysis component S3 (31%), while on the other hand when examining cases that are not RPM (e.g. exclusivity agreements), S3 appears in the analysis of all of them. All the above is justifiable by economic theory. G1 conducts (hard core agreements) are, mainly, strict Per Se restrictions and so there is no need to undertake analyses S3. The same holds for G2 conducts where for CMA the information exchange agreements involved future prices as well as for G3 conducts that constitute RPM cases. G3 (non-RPM) and G4 (abuse of dominance) conduct groups include conducts for which we anticipate that the utilization of more economic analysis is needed (examination of S3 and S4), in order to consider that an anticompetitive effect that could lead to consumer welfare loss is likely. This is confirmed by considering the analysis corresponding to potential consumer harm (S4). Specifically, there are no G1 and G2 cases where the agency undertook a consumer harm analysis: consumer harm is presumed from screens 1 and 2); while the fraction for G3 RPM cases is also significantly small (8%). The larger fractions belong to conduct group G3 (non-RPM) (50%) and G4 (80%). Finally *for efficiency defense analyses under screen S5, the results show that defendants are quite reluctant to bring forward efficiency claims (except for G3 (non-RPM) cases). Overall, S5 is present in only 11% of the cases.*

**5.2. Identifying the LSs adopted**

Following the results of Table 3, we identify the LSs adopted in the decisions of each conduct group in Table 4. The findings can be apprehended in an identical manner as those of Table 3. *For the vast majority of G1 and G2 cases the LS adopted in the assessment is the object-based. The same holds for G3 (RPM) cases with a slight slide towards TEB I. G3 (non-RPM) and G4 cases tend to be approached much more with a more effects based LS. In G4 cases the predominant LS is the TEB II while in G3 (non-RPM) the FEB LS is adopted in 50% of the cases. A small but, nevertheless, non-negligible number of cases (7,4%), is assessed under what we call Quick Look II and Quick Look III LSs. Quick Look LSs are assessment procedures where the agency did not go through all the screens of economic analysis that examine anticompetitive effects before examining efficiency claims i.e. it made a quick examination of the anti-competitive considerations (assessing some but not all of screens S1 – S4) and then ask defendants to present efficiency defences of the allegedly illegal conduct.*

**Table 4: Number of decisions per conduct group in which different Legal Standards were adopted**

Conduct group	Legal Standard
---------------	----------------

	SPS	Object-based or MPS	TEB I	TEB II	FEB	Quick Look I	Quick Look II	Quick Look III	Total
G1	1	23	2	0	0	0	1	0	27
% within group	3.7%	85.2%	7.4%	0.0%	0.0%	0.0%	3.7%	0.0%	
G2	0	6	0	0	0	0	0	1	7
% within group	0.0%	85.7%	0.0%	0.0%	0.0%	0.0%	0.0%	14.3%	
G3 (RPM)	0	8	3	1	0	0	1	0	13
% within group	0.0%	61.5%	23.1%	7.7%	0.0%	0.0%	7.7%	0.0%	
G3 (non-RPM)	0	0	0	0	1	0	0	1	2
% within group	0.0%	0.0%	0.0%	0.0%	50.0%	0.0%	0.0%	50.0%	
G4	0	0	1	3	1	0	0	0	5
% within group	0.0%	0.0%	20.0%	60.0%	20.0%	0.0%	0.0%	0.0%	
<b>Total</b>	<b>1</b>	<b>37</b>	<b>6</b>	<b>4</b>	<b>2</b>	<b>0</b>	<b>2</b>	<b>2</b>	<b>54</b>
<b>% of Total Decisions</b>	<b>1.9%</b>	<b>68.5%</b>	<b>11.1%</b>	<b>7.4%</b>	<b>3.7%</b>	<b>0.0%</b>	<b>3.7%</b>	<b>3.7%</b>	

### 5.3. Measuring the Cumulative Economic Analysis Indicator (CEAI), the Weighted Average Cumulative Economic Analysis Indicator (WACEAI) and indicators of legal certainty and of the Quality (Q) of enforcement per conduct group

An alternative way of measuring the role of economic analysis that focuses on the analysis undertaken just by the agency is through CEAI, an indicator of the total economic evidence examined in assessing conducts in a given category *by the agency*, in order to prove that it is anticompetitive. It is essential to note that for the construction of this index we do not take into account efficiency defense analyses, i.e. economic analyses where the burden of proof lies with the defendants. The value of CEAI depends on the number of screens assessed as shown below:

Screens examined in assessment	Value of CEAI corresponding to screens examined
S1	1
S1 and S2	2
S1 and S2 and S3	3
S1 and S2 and S3 and S4	4

Table 5 shows the number and share of decisions that correspond to each CEAI, presented by conduct group. The information provided here is in complete alignment with the findings in Table 3. The most decisions in conduct groups G1, G2 and G3 (RPM) correspond to low CEAI, while the CEAI in most G3 (non-RPM) and G4 conduct group cases are on average higher, as we would expect given that the extent to which economic analysis has to be utilized, in order to show an anticompetitive effect, according to economic theory, is greater.

In this table we also present an indicator that captures the frequency that efficiency gains have been claimed by the defendants *and* subsequently examined by the agency. We call this indicator *Average Efficiencies Indicator (AEfI)* and we calculate it as the fraction of decisions in a conduct group in which efficiency arguments were used. The results show that the more economic analysis we expect that it should be used in a case (conduct groups G4 and G3, non-RPM), in order to demonstrate anticompetitive effects, the higher the probability that screen S5 would be examined in the assessment.

**Table 5: Number and share of decisions that correspond to each CEAI**

Conduct group		CEAI				No of Cases	Average Efficiencies Indicator (AEfI)
		1	2	3	4		
G1	No of decisions	1	24	2	0	27	0.04
	Share of decisions for each CEAI	3.7%	88.9%	7.4%	0.0%		
G2	No of decisions (Info exchange of prices)	0	6	1	0	7	0.14

	Share of decisions for each CEAI	0.0%	85.7%	14.3%	0.0%		
G3	No of decisions (RPM)	0	9	3	1	13	0.08
	Share of decisions for each CEAI	0.0%	69.2%	23.1%	7.7%		
	No of decisions (non-RPM)	0	0	1	1	2	1.00
	Share of decisions for each CEAI	0.0%	0.0%	50.0%	50.0%		
G4	No of decisions	0	0	1	4	5	0.20
	Share of decisions for each CEAI	0.0%	0.0%	20.0%	80.0%		

Table 6 below presents some of the most crucial estimates and indicators for the analysis undertaken in this paper, calculated using the information in the tables above.

- The Weighted Average Cumulative Economic Analysis Indicator (WACEAI), which is the sum of the values of CEAI, as shown in Table 5, weighted by the share of decisions that correspond to each CEAI in a given conduct group.

$$WACEAI_j = \sum_i CEAI_{ji} \times s_{ji} , i = 1,2,3,4, j = 1,2,3,4$$

where  $s_{ji}$  are the shares of the CEAI in each particular conduct group  $j$ .

The values of WACEAI are interpreted as follows: the higher the WACEAI value the higher the extent of economic analysis utilised by the agency to reach decisions.

- The Concentration Index, which captures the concentration of CEAI and is calculated as  $HHI = \sum_i s_i^2$  - where  $s_i$  are the shares of the CEAI in each particular conduct group,  $\frac{1}{4} < Concentration\ index < 1$ . This indicator, measures the extent to which there is consistency in the use of economic evidence creating *legal certainty*.
- The *Quality of enforcement (Q) indicator*, which measures the divergence of the WACEAI from a theoretically “optimal” CEAI. In our analysis we assume that the “optimal” CEAI would be equal to: “2” for conduct group G1, G2 (for CMA, information exchange involving future pricing) and G3 (RPM) (since contextualizing the market (economic analysis S2) is considered important even for b-object restrictions; “3” for G2 (information exchange not involving future pricing) and “4” for conduct groups G3 (non-RPM) and G4. Q is calculated as the difference between the maximum possible deviation and the (absolute value of the) actual deviation of WACEAI from the theoretically optimal CEAI. If we denote the quality of enforcement for conduct group  $j$ ,  $j = 1,2,3,4$  as  $Q_j$ , then:

$$Q_j = 3 - ABS(WACEAI_j - \widehat{CEAI}_j) \text{ where } \widehat{CEAI}_j \text{ is the optimal CEAI for conduct group } j.$$

We, also, normalize these values, i.e. we rescale them so that they end up ranging between 0 (minimum quality) and 1 (maximum quality). The normalization, is achieved by expressing  $Q$ , relative to the maximum deviation, which is equal to 3, since the maximum CEAI is 4 and the minimum actual CEAI is 1.

- Finally, the Weighted Average Enforcement Quality (WAEQ) index, which expresses the *overall quality of enforcement* and is the only indicator calculated across all conduct groups. The WAEQ ( $0 \leq WAEQ \leq 3$ ) is estimated by weighting the value of the quality indicator  $Q$  of each conduct group with respect to the share of the decisions of that group relative to the total number of the decisions in the sample.

$$WAEQ = \sum_j Q_j \times s_j, \quad j = 1,2,3,4$$

where  $s_j$  is the share of the decisions in each conduct group  $j$  relative to the total number of the decisions in the sample.

Table 6 leads to a number of very useful observations. Firstly, the values of WACEAI for conduct groups G1 (2.04), G2 (2.14) and G4 (3.80) are very close to their “optimal” CEAI, as defined above. Group G3 (RPM) slightly diverges from its optimal CEAI (2), while the rest of the G3 cases (non-RPM) diverge even more from their own optimal with WACEAI equal to 3.50.

Another important remark extracted from Table 6 concerns the *concentration of CEAI*. For conduct groups G1 and G2 the concentration is high, indicating a very high degree of uniformity in the approach for assessing these conducts, as we should expect given that these are essentially conducts that restricted by-object. For G4 the concentration is also quite high (0,68) indicating uniformity in the assessment approach with some deviations that can be justified by the differences in the sub-groups of conducts in G4. For the other conduct group, G3, the concentration is quite low though, as shown below (Table 8), there is no evidence that there was any significant shift over the years in CMA’s approach in assessing such cases. Not much weight should be placed on the result regarding G3 (non-RPM) given the very small number of data (just 2 cases). For G3 (RPM) we can see that for the great majority of the cases (about 70%, Table 5), the CMA relies on just screens 1 and 2 to reach decisions, so the by-object restriction LS is the dominant LS; though, in a quite high fraction of cases (23%) it investigated also S3, in order to ascertain anticompetitive effects and in 7% of the cases it also investigated S4 in order to ascertain that these effects were likely to lead to welfare harm.

Next, regarding the quality of enforcement indicator,  $Q$ , the findings indicate high quality across all conduct groups. Especially, for conduct groups G1, G2 and G4 the quality is very high and very close to the maximum level (0.99, 0.95 and 0.93 respectively), while conduct group G3 follows closely with overall quality well above 0.80.

As expected, from the results just described, *the overall WAEQ of the CMA can be considered as very high (0.91). Indeed, it is significantly higher than that of the DGCOMP that has been examined using the same methodology* (Katsoulacos and Makri, 2023) – see bellow.



**Table 6: The WACEAI and indicators of the quality of enforcement and of legal certainty by conduct group**

Conduct group (share)		WACEAI	CEAI with highest share	Concentration index: Max. 1	Quality (Q) of enforcement (Optimal CEAI): Max. 3	Value of Q relative to max. dev. of 3: Max. 1
G1	0.50	2.04	2	0.80	2.96 (2)	0.99
G2	0.13	2.14	2	0.76	2.86 (2)	0.95
G3 (RPM)	0.24	2.38	2	0.54	2.62 (2)	0.87
G3 (non-RPM)	0.04	3.50	3,4	0.50	2.50 (4)	0.83
G4	0.09	3.80	4	0.68	2.80 (4)	0.93
WAEQ (Weighted Average Enforcement Quality (WAEQ) of agency): Max. 3					2.74	
WAEQ relative to maximum: Max. 1					0.91	

#### 5.4. Index of Total Economic Evidence (TEE)

In order to capture the total economic evidence considered on average during the assessment of decisions in a given conduct group, irrespective of the burden of proof, i.e., taking also into account the extent to which efficiency arguments were presented, we construct the Total Economic Evidence (TEE) indicator.

For the calculation of the TEE index we treat each of the 5 evidentiary screens identically and assign a value of 1 to the “amount of evidence produced by the screen” if the screen is assessed and a value of 0 otherwise. Obviously the total evidence has a maximum of 5 if all screens are assessed. Subsequently, to each of the 8 LSs we assign a value of the total economic evidence considered by the LS, ranging from 1 to 5 for SPS to FEB, a value of 2 for QL I, of 3 for QL II and of 4 for QL III.

Therefore, the TEE index for each conduct group will be the weighted average of the LSs adopted in the assessment of the decisions of that particular conduct group. We calculate the index as:

$$TEE_j = \sum_i share_{ji} \times LS_i, \quad i = 1 \dots 8, \quad j = 1,2,3,4$$

where  $share_{ji}$  is the fraction of the number of decisions that correspond to each LS "i" relative to the total number of decisions in each conduct group "j" and  $LS_i$  is the legal standard (i.e. SPS, MPS, TEB I, TEB II, FEB, Quick Look I, Quick Look II, Quick Look III).

**Table 7: Index of Total Economic Evidence (TEE) considered during assessment, irrespective of burden of proof**

Conduct group	TEE irrespective of burden of proof: Max. 5	TEE irrespective of burden of proof, relative to maximum: Max. 1
G1	2.07	0.42
G2	2.29	0.46
G3 (RPM)	2.46	0.49
G3 (non-RPM)	4.50	0.90
G4	4.00	0.80

The results in Table 7 are exactly as expected. For conduct groups G1, G2 and G3 (RPM), the cases involve low TEE, as they should, given that in these cases anticompetitive effects leading to consumer welfare loss can be presumed following, in the vast number of cases, assessment of just screens S1 and S2; while groups G3 (non-RPM) and G4 have a significantly larger value of TEE as more screens need to be assessed for each specific case in order to show an anticompetitive effect, as suggested by economic theory.

### 5.5. The evolution of the CEAI scores over time

As mentioned above in the last few decades the views emerging from the theoretical economic literature regarding the extent of economic analysis in assessing different conducts has been changing. This section presents intertemporal descriptive statistics of the CEAI indicator in the investigation of the aforementioned conduct groups for 4 consecutive periods of 5 years each, from 2000 to 2020. Table 8 shows how the WACEAI for each conduct group changes across the 4 time periods.

**Table 8: Evolution of WACEAI by conduct group**

Conduct Group	2000 - 2004	2005 - 2009	2010 - 2014	2015 - 2020
G1: WACEAI	1.75	2.00	2.00	2.14
No of decisions: 27	4	6	3	14
G2: WACEAI	n/a	n/a	2.00	2.25
No of decisions: 7	0	0	3	4
G3 (RPM): WACEAI	2.40	n/a	2.00	2.50
No of decisions: 13	5	0	2	6
G3 (non-RPM): WACEAI	n/a	n/a	4.00	3.00
No of decisions: 2	0	0	1	1
G4: WACEAI	4.00	4.00	3.50	n/a
No of decisions: 5	2	1	2	0

*Table 8 shows that the WACEAI of all conduct groups show substantial uniformity across the years. This indicates that OFT and then CMA had achieved a very high quality of antitrust enforcement already in 2000 and this has been to very large maintained in the following 2 decades.*

## **6. Comparative Analysis: CMA (UK) – DGCOMP (EU)**

In this section we present a comparative analysis, on the role of economics and the quality of antitrust enforcement, between one of the most prominent CAs, EUs DGCOMP and one of the oldest and most sophisticated CAs in Europe, UK’s CMA. To achieve this objective we will utilize the dataset created for the purposes of this paper and the dataset and results currently available in Katsoulacos and Makri (2020; 2023). In other words we will compare and contrast the indicators capturing the extent and type of economics used in reaching infringement decisions as well as the *quality of enforcement* that we measured for the purposes of this paper with EUs counterparts measured in Katsoulacos and Makri (2020; 2023). For convenience, we present the main results concerning DGCOMP in the Annex.

The two datasets differ in terms of data volume and the period under investigation. The number of cases examined by DGCOMP is significantly larger as could be expected considering the difference in the size of the two jurisdictions. The differences in the time periods reviewed can be attributed to the insufficient number of older decisions publicly available by the CMA, especially from the OFT period. Nevertheless, very useful conclusions can be drawn when comparing the results generated by the analysis of the datasets. We should note here that since the CMA dataset

includes only information exchange cases that involve information on prices, the comparison for conduct group G2 cases will be limited only to those cases.

Starting with the examination of the extent and type of economic analysis applied in the decisions per conduct group and comparing Tables 3 (CMA) and 3.A (DGCOMP, Annex) we observe a number of similarities between the two sets of data. Specifically, analysis components S1 and S2 are present in almost all decisions, indicating that apart from the determination of the nature and characteristics of the conducts *both agencies consider market contextualization an important part of their assessment irrespective of the type of conduct assessed.*

Moving on to the rest of the economic analysis components (S3, S4 and S5) we identify a number of differences between CMA and DGCOMP, although the main tendency for each conduct group is quite similar. More specifically, for conduct group G1 it is clear that the economic analysis used is mainly limited to economic components S1 and S2. Regarding the other economic components we notice slightly different findings. These differences are more noteworthy when we examine the so called efficiency defense analysis component (S5), 4% in CMA, 15% in DGCOMP. *The above indicate that either the percentage of cases where the defendants invoke Article 101(3)/Section 9 is much larger in EU than in UK or that it is much more frequent for DGCOMP (than for CMA) to undertake a comprehensive economic analysis in order to address the efficiency arguments put forward by the defendants.*

Continuing the analysis for the other conduct groups, we observe that the *two CAs converge regarding their approach for cases G2 and G3 (RPM)* except in the use of economic analysis component S5, where DGCOMP shows higher figures.

*The assessment procedures for abuse of dominance cases (G4) show more uniformity between the two CAs. On the other hand, the G3 (non-RPM) cases seem to be treated very differently. We observe a tendency on behalf of DGCOMP to limit the extent of economic analysis that it applies to the assessment of these cases to components S1 and S2 even if the cases under consideration clearly require, according to economic theory, the utilization of more screens during the assessment<sup>19</sup>.*

Continuing with the comparison of the two CAs we examine one of the main objectives of this paper which is the identification of the LSs adopted in the assessment of the infringement decisions. Considering the data presented in Table 4 (CMA) and Table 4.A (DGCOMP, Annex) we notice that although the overall picture seems rather similar between the two CAs (i.e. when we compare the number of decisions where a certain LS is used as a percentage of the total decisions) with both CAs showing a significant use of object based LS followed by TEB I, the results alter when we examine each conduct group separately. For DGCOMP the predominant LS for all conduct groups in review, except G4, is the object based LS and even for G4 the predominant LS is TEB I. On the other hand, *CMA shows a tendency towards a more effects based approach when groups G3*

---

<sup>19</sup> See discussion in Katsoulacos and Makri (2000) for discussion regarding potential explanations.

(non-RPM) and G4 are considered. However notice that, with regard to G4 (abuse of dominance), while CMA’s enforcement quality score is higher overall, *if we concentrate on just decisions with a FEB LS, the DGCOMP scores higher than CMA especially if we concentrate on the latter’s decisions of high-tech digital markets* (Bagari et. al., 2023). This is consistent with our finding above that efficiency defense arguments play a more significant role in DGCOMP than in CMA decisions.

Next in our comparison is the indicator that holds a fundamental role in our analysis, i.e. the Cumulative Economic Analysis Indicator (CEAI). Tables 5 (CMA) and 5.A (DGCOMP, Annex) show the number and share of decisions that correspond to each CEAI, presented by conduct group. Bellow we present the data in the tables in chart forms in order to visually indicate the differences between the two CAs. The charts clearly show a shift to the right for the CMA curves when conduct groups G3 (non-RPM) and G4 are concerned while the curves for all other conduct groups peak at CEAI=2. The findings are very consistent with the optimal values suggested above.

On the other hand, the corresponding DGCOMP curves peak at CEAI=2 for conduct groups G1, G2, G3 (non-RPM), the G3 (RPM) curve shows an abnormal variance (although this may be due to the small sample) and G4 curve peaks at CEAI=3. The above show a considerable deviation from the optimal values for DGCOMP (except G2), with the greatest deviation being that of conduct group G3 (non-RPM).

Chart 1: Shares of decisions that correspond to each CEAI – CMA

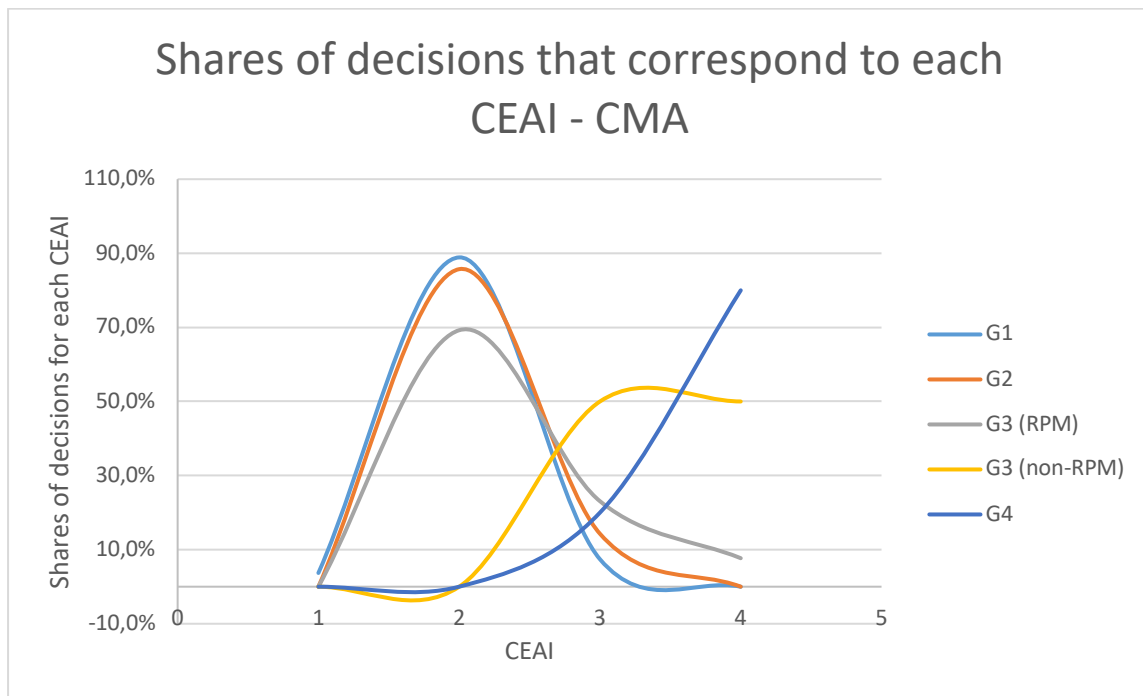
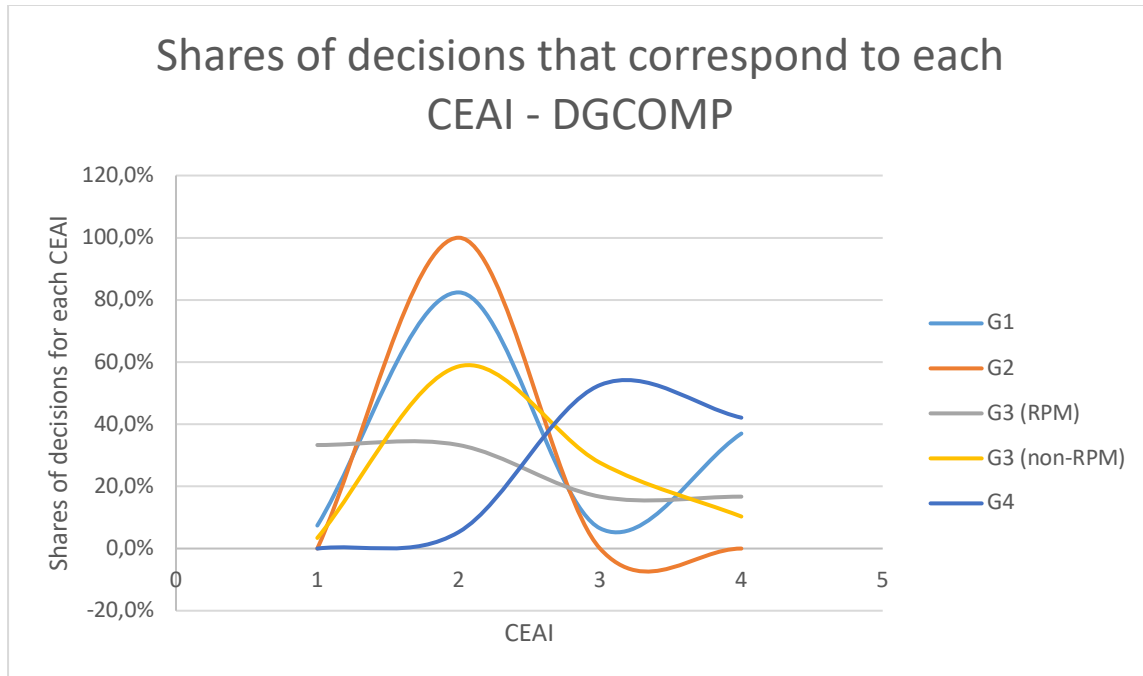


Chart 2: Shares of decisions that correspond to each CEAI – DGCOMP.



The overall picture of the extent of economic analysis utilized in each conduct group shown on the above charts can be, also, considered as a picture of the WACEAI of each conduct group.

We proceed with the comparison of the indicator that measures the quality of enforcement. The table below presents the differences in the quality of enforcement indicator (Q) per conduct group for each of the two CAs under review. We observe that for groups G2 and G3 (RPM) DGCOMP shows slightly higher figures, however the advantage is negligible and it could be argued that the two CAs display similar levels of quality, as is the case, also, in group G1. On the other hand CMA is undoubtedly superior in terms of quality for conduct groups G3 (non-RPM) and G4 (42% and 15.4% difference respectively). The above lead us to the conclusion that DGCOMP has a quality disadvantage in cases where more economic analysis needs to be utilized.

Finally, comparing the overall quality of enforcement (WAEQ) we notice that the difference between the two CAs is quite substantial (5.5%), clearly a product of the significant differences in quality for groups G3 (non-RPM) and G4, as mentioned.

**Table 10: Comparing the indicators of the quality of enforcement by conduct group**

Conduct group	Quality (Q) [% difference CMA to DGCOMP]
G1	0.7%
G2	-4.9%
G3 (RPM)	-8.0%

G3 (non-RPM)	42.0%
G4	15.4%
WAEQ (relative to maximum) [% difference CMA to DGCOMP]	5.5%

A key factor in order to provide a comprehensive comparative analysis between the two CAs is the examination of the TEE indicator i.e. the indicator that depicts the total economic evidence considered during the assessment of the cases irrespective of the burden of proof. The two agencies display very similar figures as regards conduct groups G1, G2 and G3 (RPM), (around 0.45), supporting the view that these conduct groups do not require the utilization of all evidentiary screens. On the contrary, the figures for conduct groups G3 (non-RPM) and G4 are very divergent. CMA displays rather high values of TEE (0.90 for G3 (non-RPM) and 0.80 for G4) which reflects a more extensive utilization of economic analysis when compared to the significantly lower values of DGCOMP (0.57 and 0.73).

Finally, with regard to the evolution of the CEAI scores over time. Tables 8 (CMA) and 8.A (for DGCOMP, in Annex) depict this evolution by showing how the WACEAI for each conduct group changes across specific time periods. However, the breakdown of the decisions across different time periods leaves us with a small number of observations for each time period (especially for CMA) and the periods under review differ significantly between the two CAs. Nonetheless, the data do not show a significant shift over the years, especially for CMA since the sample covers a quite recent time period. For DFCOMP a significant increase in the value of CEAI for group G4 is noticeable, indicating a significant move towards more effects-based procedures over the years, though in the latest period (2017 – 2021) this move does not seem to be sustained.

## 7. Conclusions

It is widely accepted for a long time that an economics-based approach should play an important role in the enforcement of competition law when assessing abuse of dominance and vertical restraint conducts. In this article we have investigated whether CMA’s enforcement record is consistent with this approach.

To do this we created a dataset of all the antitrust infringement decisions reached by CMA in the period 2000 – 2020, identifying the economic analysis components or screens assessed (as these were presented above) present in each of these decisions. We also traced the outcome, for all the decisions that were appealed, of the judicial review process – up to the final decision of the Appeal Courts.

In our analysis we propose and empirically compute a number of indicators, which could constitute important tools for the empirical measurement of the role of economic analysis and type of LSs in competition law enforcement. In our main results we have shown that, on average, in the period under investigation, economic analysis has played a significant role in CMA’s investigations, also compared to other jurisdictions. Excluding, of course, cases that had as their **object** the prevention, restriction or distortion of competition.

Our findings suggest that the *overall quality of enforcement* by CMA, measured by the divergence of the actual from the socially optimal extent of economic analysis is very high (with the overall quality amounting to 0.91). These figures are higher or significantly higher than those of the DGCOMP that has been examined using the same methodology (Katsoulacos and Makri, 2020; 2023).

## References

Areeda P.E & Herbert Hovenkamp (2017). “Antitrust Law: An Analysis of Antitrust Principles and their Application”. Wolters Kluwer, 4th ed. The first edition of the Treatise was by P. Areeda and D. F. Turner.

Available at SSRN: <https://ssrn.com/abstract=3733662> or <http://dx.doi.org/10.2139/ssrn.3733662>.

Bageri V., Y Katsoulacos, E Metsiou (2023) “The dominance of the Consumer Welfare criterion in antitrust enforcement: an empirical analysis of UK and EC “, *Competition Policy International Antitrust Chronicle*, December 2023.

Baker, J. B. (2015). “Taking the error out of ‘error cost’ analysis: What's wrong with antitrust's right”. *Antitrust L.J.*, **80**, 6.

C. Caffarra (2023), ““Consumer Welfare Is Dead”: What Do We Do Instead?—A Perspective from Europe”, Promarket.

Carrier Michael A. (1999). “The Real Rule of Reason: Bridging the Disconnect”, *BYU L. REV.* 1265, 1267–68.

Carrier Michael A. (2009). “The Rule of Reason: An Empirical Update for the 21st Century”, 16 *GEO. MASON. L. REV.* 827, 828.

Cavanagh, Edward D. (2015). “Impediments to Renewed and Reinvigorated Antitrust Enforcement”. 14 *Wm. & Mary Bus. L. Rev.* 583 (2023).

Fox, E (2023). “Simple Reules”, in “Antitrust and the Digital Economy: Presumptions, Legal Standards and Key Challenges”, Concurrences Antitrust Publications, edited by Katsoulacos Y.

Gavil A.I. (2008), “Burden of Proof in U.S. Antitrust Law” in *Issues in Competition Law and Policy* 1 125, ABA Section of Antitrust Law.

Gavil A. I. (2012). “Moving Beyond Caricature and Characterization: The Modern Rule of Reason in Practice”. 85 *Southern Cal. L. Rev.* 733, 759.



Giulio Federico, Fiona Scott Morton, Carl Shapiro (2020). “Antitrust and Innovation: Welcoming and Protecting Disruption” (Innovation Policy and the Economy 20, Edited by Josh Lerner and Scott Stern, The University of Chicago Press Chicago and London.

Hovenkamp H. J. and F. Scott Morton (2020). “Framing the Chicago School of Antitrust Analysis”, *U. Pa. L. Rev.*, Vol. 168, No. 7, pp. 1843–1878.

Hovenkamp, H. J. (2021). “Antitrust Error Costs”. *Competition Policy International*.  
<https://doi.org/10.2139/ssrn.3853282>

Italianer A. (2013). “Competitor Agreements under EU Competition Law” (40th Annual Conference on International Antitrust Law and Policy, Fordham Competition Law Institute New York).

Jones A. and W. Kovacic (2017). “Identifying Anticompetitive Agreements in the US and the EU: Developing a Coherent Antitrust Analytical Framework”, *Antitrust Bulletin*.

J Kanter (2022), Remarks at New York City Bar Association’s Milton Handler Lecture” (May 18, 2022).

Katsoulacos Y., Avdasheva S. and Golovanova S. (2018). “A Methodology for Empirically Measuring the Extent of Economic Analysis and Evidence and Identifying the Legal Standards in Competition Law Enforcement” in *Festschrift in Honor of Frederic Jenny*, *Concurrences Review*.

Katsoulacos Y., S. Avdasheva, K. Benetatou, S. Golovanova, and G. Makri (2019) “Comparing the Role of Economics/Effects-Based in Antitrust Enforcement and Its Relation to the Judicial Review in the EC to Other Countries”, *Journal of European Competition Law & Practice*, Vol. 12, No. 2.

Katsoulacos, Y. (2019a). “On the concepts of legal standards and substantive standards (and how the latter influences the choice of the former)”. *Journal of Antitrust Enforcement*, 7, 365–385. <https://doi.org/10.1093/jaenfo/jnz011>

Katsoulacos, Y., & Makri, G. (2020). “The role of economics and the type of legal standards in antitrust enforcement by the EC: An empirical investigation”. *Journal of Antitrust Enforcement*, 9, 457–504.

Katsoulacos, Y. & Ulph, D. (2022). “Choosing legal rules or standards in antitrust enforcement: A proposal for extending and facilitating the use of the decision-theoretic approach”. *Managerial and Decision Economics*. 44. 10.1002/mde.3787.

Katsoulacos Y. (2023) “Could a move to more presumption-based Legal Standards be justified, for the big-tech digital markets, on error-cost minimizing grounds?”, in *Antitrust and the Digital Economy: Presumptions, Legal Standards and Key Challenges*”, Concurrences Antitrust Publications, edited by Katsoulacos Y.

Katsoulacos Y. (2023a) “Accelerating antitrust enforcement through the use of more presumption-based Legal Standards for the big-tech digital markets”, *Concurrences Review*, 2003 – 2.

Katsoulacos, Y. & Makri G. (2023). “The role of economics and the type of legal standards in antitrust enforcement by the EC: An empirical investigation revisited”. Mimeo WP available on request.

Katsoulacos, Y. & Zhixue Guo. (2023). “The role of economics and the type of legal standards in antitrust enforcement by China: An empirical investigation”. Mimeo DP available on request.

Khan Lina (2018). “The Ideological Roots of America's Market Power Problem”, *127 Yale L. J. F.* 960.

Kovacic W. (2021). “The Future Adaptation of the Per Se Rule of Illegality in US Antitrust Enforcement”. *Columbia Business Law Review*.

Manne Geoffrey (2020) “Error Costs in Digital Markets”. *The Global Antitrust Institute Report on the Digital Economy*.

Melamed A.D and N. Petit (2019). “The Misguided Assault on the Consumer Welfare Standard in the Age of Platform Markets”. *Review of Industrial Organisation*, 54

Richard Whish & David Bailey (2023). “Private enforcement of competition law: its role and development in the EU,” Chapters, in: Barry J. Rodger & Miguel S. Ferro & Francisco Marcos (ed.), *Research Handbook on Private Enforcement of Competition Law in the EU*, chapter 1, pages 2-27, Edward Elgar Publishing.

Witt Anne C. (2019). “The European Court of Justice and the More Economic Approach to EU Competition Law – is the Tide Turning?” *The Antitrust Bulletin*, Vol.64 (2), 172 – 213.

Wu T. (2018). “After consumer welfare, now what? The ‘protection of competition’ standard in practice”, *CPI*.

## **Annex: Results for DGCOMP (Katsoulacos and Makri, 2020; 2023)**

Table 1.A: The Dataset: Number of decisions, appeals and annulments (by conduct group)  
– DGCOMP: 1992 – 2021

G1	G2	G3	G4	Total
Total number of decisions (shares)				
108 (0,574)	7 (0,037)	35 (0,186)	38 (0,202)	188
Number of appealed decisions (shares)				
87 (0,81)	2 (0,29)	22 (0,63)	23 (0,61)	134 (0,71)
Number of annulled decisions (shares)				
40 (0,46)	0 (0,00)	11 (0,50)	8 (0,35)	59 (0,44)

Table 2.A: Highest level of economic analysis present in each decision

	The highest level of economic analysis present in the decision					Total
	S1	S2	S3	S4	S5	
Number of decisions	10	102	26	8	42	188
Shares	5,3%	54,3%	13,8%	4,3%	22,3%	

Table 3.A: Economic analysis applied per conduct group

Type of Analysis Applied \ Conduct group	S1=1	S2=1*	S3=1	S4=1	S5=1
G1	108	100	11	4	16
% within group	100,0%	92,6%	10,2%	3,7%	14,8%
G2 (price info)	5	5	0	0	2
% within group	100,0%	100,0%	0,0%	0,0%	40,0%
G2 (other info)	2	2	1	0	2
% within group	100,0%	100,0%	50,0%	0,0%	100,0%
G3 (RPM)	6	4	2	1	0
% within group	100,0%	66,7%	33,3%	16,7%	0,0%
G3 (NON-RPM)	29	28	11	3	11
% within group	100,0%	96,6%	37,9%	10,3%	37,9%
G4	38	38	36	16	11

% within group	100,0%	100,0%	94,7%	42,1%		28,9%
Total	188	177	61	24		42
	100,0%	94,1%	32,4%	12,8%		22,3%

Table 4.A: Number of decisions per conduct group in which different Legal Standards were adopted

Conduct group (share)	Legal Standard								Total
	SPS	Object-Based or MPS	TEB I	TEB II	FEB	Quick Look I	Quick Look II	Quick Look III	
G1	7	82	3	0	4	1	7	4	<b>108</b>
% within group	6,5%	75,9%	2,8%	0,0%	3,7%	0,9%	6,5%	3,7%	<b>57,4%</b>
G2 (price info)	0	3	0	0	0	0	2	0	<b>5</b>
% within group	0,0%	60,0%	0,0%	0,0%	0,0%	0,0%	40,0%	0,0%	<b>2,7%</b>
G2 (other info)	0	0	0	0	0	0	1	1	<b>2</b>
% within group	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	50,0%	50,0%	<b>1,1%</b>
G3 (RPM)	2	2	1	1	0	0	0	0	<b>6</b>
% within group	33,3%	33,3%	16,7%	16,7%	0,0%	0,0%	0,0%	0,0%	<b>3,2%</b>
G3 (NON-RPM)	1	13	4	0	3	0	4	4	<b>29</b>
% within group	3,4%	44,8%	13,8%	0,0%	10,3%	0,0%	13,8%	13,8%	<b>15,4%</b>
G4	0	2	18	7	9	0	0	2	<b>38</b>
% within group	0,0%	5,3%	47,4%	18,4%	23,7%	0,0%	0,0%	5,3%	<b>20,2%</b>
<b>Total</b>	<b>10</b>	<b>102</b>	<b>26</b>	<b>8</b>	<b>16</b>	<b>1</b>	<b>14</b>	<b>11</b>	<b>188</b>
<b>% of Total Decisions</b>	<b>5,3%</b>	<b>54,3%</b>	<b>13,8%</b>	<b>4,3%</b>	<b>8,5%</b>	<b>0,5%</b>	<b>7,4%</b>	<b>5,9%</b>	<b>100,0%</b>

Table 5.A Number and share of decisions that correspond to each CEAI

Conduct group		CEAI				No of Cases	Average Efficiencies Indicator (AEI)
		1	2	3	4		
G1	No of decisions	8	89	7	4	108	0,15

	Share of decisions for each CEAI	7,4%	82,4%	6,5%	3,7%			
G2	No of decisions (price info)	0	5	0	0	5	0,40	
	Share of decisions for each CEAI	0,0%	100,0%	0,0%	0,0%			
	No of decisions (other info)	0	1	1	0	2	1,00	
	Share of decisions for each CEAI	0,0%	50,0%	50,0%	0,0%			
G3	No of decisions (RPM)	2	2	1	1	6	0,00	
	Share of decisions for each CEAI	33,3%	33,3%	16,7%	16,7%			
	No of decisions (NON-RPM)	1	17	8	3	29	0,38	
	Share of decisions for each CEAI	3,4%	58,6%	27,6%	10,3%			
G4	No of decisions	0	2	20	16	38	0,29	
	Share of decisions for each CEAI	0,0%	5,3%	52,6%	42,1%			

Table 6.A: The WACEAI and indicators of the quality of enforcement and of legal certainty by conduct group

Conduct group (share)		WACEAI	CEAI with highest share	Concentration index: Max. 1	Quality (Q) of enforcement (Optimal CEAI)	Value of Q relative to max. dev. of 3
G1	0,57	2,06	2	0,69	2,94 (2)	0,98
G2 (price info)	0,03	2,00	2	1,00	3,00 (2)	1,00
G2 (other info)	0,01	2,50	2, 3	0,50	2,50 (3)	0,83
G3 (RPM)	0,03	2,17	1, 2	0,28	2,83 (2)	0,94
G3 (NON-RPM)	0,15	2,45	2	0,43	1,45 (4)	0,48
G4	0,20	3,37	3	0,46	2,37 (4)	0,79
WAEQ (Weighted Average Enforcement Quality (WAEQ) of agency)					2,59	
WAEQ relative to maximum					0,86	

Table 7.A: Index of Total Economic Evidence (TEE) considered during assessment, irrespective of burden of proof

Conduct group	TEE irrespective of burden of proof: Max. 5	TEE irrespective of burden of proof, relative to maximum: Max. 1
G1	2,21	0,44
G2 (price info)	2,40	0,48
G2 (other info)	3,50	0,70
G3 (RPM)	2,17	0,43
G3 (NON-RPM)	2,83	0,57
G4	3,66	0,73

Table 8.A: Evolution of WACEAI by conduct group

Conduct Group	1992 - 1996	1997 - 2001	2002 - 2006	2007 - 2011	2012 - 2016	2017 - 2021
G1: WACEAI	2,20	2,00	1,93	2,27	1,86	2,14
No of decisions: 108	15	16	30	26	14	7
G2 (price info): WACEAI	2,00	n/a	n/a	2,00	2,00	2,00
No of decisions: 5	2	0	0	1	1	1
G2 (other info): WACEAI	3,00	2,00	n/a	n/a	n/a	n/a
No of decisions: 2	1	1	0	0	0	0
G3 (RPM): WACEAI	4,00	2,00	2,00	n/a	1,00	n/a
No of decisions: 6	1	2	2	0	1	0
G3 (NON-RPM): WACEAI	2,31	2,75	2,00	3,00	n/a	3,00
No of decisions: 29	13	8	5	1	0	2
G4: WACEAI	2,75	3,23	3,29	4,00	4,00	3,43
No of decisions: 38	4	13	7	3	4	7