

Legal standards and economic analysis in antitrust enforcement: an empirical investigation for the case of Russia¹

October 2018

Avdasheva S.², Golovanova S.³, Katsoulacos Y.⁴

In this paper we propose a methodology to measure empirically the extent of economic analysis used and the legal standards (LSs) adopted by Competition Authorities (CAs) and then apply it to the appealed investigations of the Russian competition authority, the Federal Antitrust Service (FAS). We show that on average, economic analysis still plays a very modest role in the investigations, as FAS applies close to *Per Se* legal standards even when assessing conducts for which effects-based LSs would be more suitable. There is no discernible evolution towards a more effect-based approach during the period 2008-2015. Further, the choice of LSs for specific conducts can create considerable legal uncertainty for firms about how these conducts will be assessed by the CA. Overall, our empirical findings indicate *low quality of enforcement*. We also examine how LSs influence the probability that FAS decisions are annulled under appeal, as well as their effect on litigation costs and on the probability to appeal. Our results are consistent with recent arguments, according to which, the higher disputability of decisions as a result of increasing the extent of economic analysis under effects-based, increases the annulment rates of decisions under appeal. Also, our results show increasing enforcement costs and higher appeal rates when LSs move closer to effects-based.

Key words: antitrust, economic analysis, legal standard, Per Se, effect-based enforcement.

¹ We are grateful, for financial support, to the Higher School of Economics Basic Research program and for the comments from participants of the 13th CRESSE Conference, 29th June – 1st July 2018. Of course, all errors and ambiguities remain our own responsibility.

² Professor of Economics, National Research University Higher School of Economics. Email: avdash@hse.ru

³ Professor of Economics, National Research University Higher School of Economics-Nizhny Novgorod. Email: svtgolovanova@hse.ru

⁴ Professor of Economics, Athens University of Economics and Business. Email: yanniskatsoulacos@gmail.com

1. Introduction: factors that influence the choice of LSs – a brief review

The issue of the choice of legal standards and of the appropriate role and the extent to which economic analysis and evidence should be applied in the enforcement of Competition Law (CL) has always been and remains very controversial⁵. We can distinguish two strands in the recent literature. The first, examines the question of what is the *optimal* choice of legal standards, and hence of the role of economics in CL enforcement, assuming an welfarist substantive standard and using a *minimization-of-costs of decisions errors* framework and, more recently, a more general *maximization-of-welfare* framework (that incorporates the former). The main factors that then need to be taken into account and have been discussed quite extensively in this literature can be summarized as follows:

- the cost of decision errors (of Type I and Type II) under the alternative standards;
- the deterrence or indirect (or incentive) effects of the standards;
- whether the standard generates legal uncertainty;
- other enforcement costs (including the administrative costs of enforcement and the costs to firms of self-assessing their actions or of reducing legal uncertainty).

In a series of papers, Katsoulacos and Ulph (2009, 2011, 2015 and 2016) have attempted, by using a maximization-of-welfare framework to provide answers on how the factors above affect the choice of the (optimal) legal standard⁶ and hence, indirectly, about the appropriate role and extent of economic analysis in CL enforcement. Their analyses, point quite strongly to the view that for a range of conducts, which now are understood not to be strongly presumptively illegal⁷ and for which the developments in economic theory and modeling in the last 20 or so years improved significantly the discriminating quality of the assessment⁸, moving to assessment with effects-based standards will improve welfare due to a reduction in the costs of decision errors and an improvement in deterrence effects⁹. But, as is widely recognized, the legal standards actually adopted in many countries and, most importantly in the EU and its member states, remain close to Per Se (and the extent of economic analysis applied by the vast majority of CAs today remains

⁵ For discussions and empirical information concerning the use and usefulness of economics in competition law enforcement see Baker (2003), Gavil (2008), Neven (2006), Schinkel (2008) and Lianos (2012).

⁶ Extensive references and reviews of the literature related to these issues are contained in these papers. See also J Padilla (2011), page 435.

⁷ But which up to the 1990s were widely considered as strongly presumptively illegal. See for more details Katsoulacos, Avdasheva and Golovanova (2017a).

⁸ That is, the ability of the assessment to discriminate accurately between harmful and benign conducts.

⁹ Which are likely to more than compensate for higher administrative costs and legal uncertainty. As Jones and Kovacic (2017, p 7) note «many jurisdictions apply a rule of per se illegality, or virtual per se illegality, against some horizontal agreements such a price fixing. The extent to which such a rule should be expanded beyond this.....is much more controversial and contested». As they indicate (p. 16) nowadays in US, vertical restraints, mergers and single-firm exclusionary behavior are not assessed by per se. Also Blair and Sokol (2012).

low)¹⁰ for cases in which effects-based LSs would be considered more appropriate from the point of view of welfare maximisation. As Geradin and Petit (2010, p. 31)) note, the assessment of abuse of dominance cases in EU has relied on «old, formalistic legal appraisal standards, and (has shown) a reluctance to endorse a modern economic approach»¹¹.

This implies that the arguments concerning decision errors, deterrence effects (as well as legal uncertainty and administrative costs), are not the only, or even the most important, in considering the choice of legal standards. In practice, other factors must be important. In some, more recent, papers these are at the center of the analysis (Katsoulacos, 2018a,b and Katsoulacos, Avdasheva and Golovanova, 2018). One is related to the objectives of CAs, specifically the reputational concerns of those deciding the enforcement procedures, that are affected by the judicial review of the CA's decisions. As a result of these concerns, CAs will make their choice taking into account what they anticipate to be the Courts' choice of legal standard. Given this, it must also be recognized that the choice of legal standard is dependent on the *substantive standard* adopted by Courts. While in academic discussions this is usually assumed to be welfarist (liability requiring a showing of adverse effects on welfare¹²), in practice this is often not the case. For example, the substantive standard may be just to “protect the economic freedom of market participants”, or, the pursuit of a “system of undistorted competition” (Wils, 2014), without obligation to show adverse effects on consumer welfare or efficiency (Rey and Venit, 2015) – which would imply that any conduct that puts one or more competitors at a disadvantage would be considered unlawful¹³, irrespective of the ultimate consequences of the conduct for welfare¹⁴. The link between substantive standards and the choice of legal standards has been discussed recently and it has been

¹⁰ There are exceptions to this, such as US or Canada (see Hovenkamp, 2017, especially onwards from p. 43), but the statement does reflect accurately the reality in vast majority of other jurisdictions. The statement does not concern hard-core horizontal collusion for which all arguments favor a Per Se legal standard. Thus, the type of practices that we will have in mind are the other business conduct for which there is no universally accepted choice of legal standards i.e. unilateral conduct by dominant firms, vertical restraints and concerted practices

¹¹ In the meantime, the importance of effects-based standards and relying on the predictions of sound economic analysis has been stressed by OECD not just in the context of developed countries but equally and perhaps more importantly in developing ones. For example, in its recent report evaluating the Russian competition authority, that has in the last few years become the largest competition authority in the world, the OECD (2013) makes as its top recommendation that the authority must “improve the quality of economic analysis and its application to competition enforcement throughout the competition authority and in support of improved judicial decisions”.

¹² Consumer or total welfare – see also below.

¹³ The meaning of “preserving undistorted competition” was actually made clear by the EU General Court which, upholding in its entirety the Commission's Decision on *Intel*, argued that making it more difficult for a rival to compete “in itself suffices for a finding of infringement”.

¹⁴ Rey and Venit (2015) note that the effects-based standard *starts* with a showing of a distortion of the competitive process but, in order to assess this distortion and find liability, one “should (also) look at the actual or likely *effects of the conduct*“, on consumer welfare or efficiency (p. 17, italics ours). Note that here we will not try to examine the pros and cons of using “consumer welfare” or “total welfare / efficiency” as the right substantive standard. There is currently quite an intense debate on this issue, with some economists arguing for a total welfare standard, e.g. D. Carlton (2007). For a recent contribution also containing a review of the recent debate see Katsoulacos, Metsiou and Ulph (2016). Also, CAs often take into account the presence of “public interest concerns” as additional liability criteria.

demonstrated (Katsoulacos, 2018a) that adopting non-welfarist substantive standards increases the likelihood that Per Se legal standards are applied and a limited amount of economic analysis and evidence is utilized in investigations of specific conducts.

In contrast to the theoretical analyses, systematic empirical assessments of legal standards have been very limited. There are case-studies based on particular decisions or meta-analysis of a group of decisions, but there is no *statistical representation* of the legal standards applied by competition authorities. The absence of empirical measurement and statistics on legal standards limits our ability to answer important questions. Thus, it makes any international comparisons of legal standards applied in different jurisdictions and judgments on the role of economic analysis speculative. Further, it impedes the analysis of the evolution of legal standards over time and explaining the factors that drive this evolution. Both issues are important for the identification of the deviation of legal standards actually applied in competition cases from their optimal level.

For the purposes of this paper we collected and analysed a unique large dataset of 1133 antitrust infringement decisions reached by the Russian CA between 2008 – 2015, which were appealed to Courts for annulment. Our main objectives have been to use this dataset to examine the following:

1. Measure to what extent economic analysis and evidence is used in the decisions of the CA and how it evolves over time. This can be compared to the “optimal” level of economic analysis, for any given conduct, that would be applied if the appropriate¹⁵ (“optimal”) Legal Standard (LS) for that conduct was adopted. We develop a number of indices. First, we measure the Weighted Average Legal Standard (WALS) adopted for each potentially anticompetitive conduct-type (the weights been the share of each legal standard used in assessing each conduct-type), and the degree of concentration (CONC) for the WALS of each conduct-type. The higher is CONC the greater the concentration on specific LSs when assessing specific conducts and hence the greater the certainty with which it can be anticipated that a specific LS will be used in the future for a conduct. The standard deviation of WALS is another *measure of the uncertainty* that surrounds the anticipated LS for a conduct. Finally we develop *indices of quality of enforcement* measuring the deviation of WALS from its optimal level. The indices are given in Tables, 4, 5 and 6.
2. Examine how changes in the extent of economic analysis or variations in LSs, for any given conduct, is related to (how it affects) the probability that decisions on that conduct are annulled in Courts of Appeal. That is, examine empirically the probability of annulment as a function of the LS adopted. Results on this are contained in Table 7.

¹⁵ As inferred by the recent relevant literature.

3. Examine the effect of movements in LSs from Per Se towards effects-based on litigation costs and the duration of litigation. Results on this are contained in Table 8.
4. Examine the effect of movements in LSs from Per Se towards effects-based on the probability of appeal to higher courts (than just the Court of first instance). Results on this are contained in Table 9.

Our main results are as follows. First, we find, for the competition authority of Russia (FAS), both that it applies on average an insufficient degree of economic analysis, close to the amount under Per Se legal standards, as well as that the legal standards are not applied in a consistent way. In investigations of similar conducts, the legal standards applied vary quite substantially, creating high levels of legal uncertainty for firms that we also measure empirically and compare across different conduct groups. Overall, with the exception of price fixing and market sharing conducts, the quality of enforcement (measured by the deviation of the LSs adopted from the theoretical optimum) has been and remains throughout the period investigated very low.

Next, our results show that when legal standards are low (Strict Per Se – as defined below) higher legal standards tend to decrease the likelihood of decisions being annulled – indicating that adopting very low legal standards by FAS is considered “wrong” by the Courts. Simultaneously, we provide evidence that increasing the extent of economic analysis, as required in cases where the law allows otherwise Per Se violations to be rebutted, leads to an increase in the annulment rate of decisions. Our results are thus consistent with the existence of a U-shaped relationship between LSs and the probability of annulment - higher legal standards decrease and then increase the likelihood of decisions being annulled. This can be supported theoretically as showing that FAS adopts the “wrong” (lower) LSs in many cases when Courts’ standards are close to Modified Per Se or Truncated Effects-Based (MPS or TEB, as defined below) but follow the Courts when the LS is higher than TEB. When LSs are higher than TEB (and closer to full effects-based) the increase in the disputability of decisions increases the probability of annulment¹⁶.

Finally, we show that an increase in legal standards leads to additional costs of enforcement for competition authorities. Specifically, in terms of the time necessary for litigation to be completed, decisions in which FAS applies analysis close to effects-based take longer than decisions based on *per se* standards. And, higher legal standards increase the appeal rate, inducing greater re-submission of claims to higher courts which is also one of the main predictors of litigation cost.

¹⁶ See discussion below on the argument that increasing LSs close to full effects-bases increases the disputability of decisions and thus leads to an increase in annulment rates (Neven, 2006, Katsoulacos, Makri and Metsiou (2018) and Katsoulacos, 2018a).

Overall, our results are consistent with recent theoretical predictions concerning the choices of a young, reputation maximizing authority, whose reputation (and, hence, utility) is decreasing in the annulment rate of its decisions under appeal and which is operating under increasing enforcement costs as LSs move closer to effects-based (Katsoulacos, 2018a).

The remainder of the paper is organized as follows. The second section describes the empirical approach to measuring legal standards. The third section briefly reviews the context of the Russian competition enforcement that is important in order to understand the origin, structure and limitations of our dataset. The fourth and fifth sections contain the empirical analysis undertaken and a detailed discussion of the results obtained related to objectives 1 – 5 described above. The sixth section concludes.

2. Identifying the extent of economic analysis and legal standards in competition law enforcement: a methodology

2.1 Types of economics analysis applied in competition law enforcement

In order to measure the extent to which economic analysis is used and hence the legal standard that is adopted for the assessment of specific conducts by a CA we propose using the following approach. We identify the important components of economic analysis that are necessary to substantiate, in a specific investigation, how the relevant market has been defined, how market power raising or exclusionary effects have been shown, how efficiency effects have been shown and what is, ultimately, the welfare impact of the conduct. Then we analyze the documents on particular decisions made by the CA and identify whether a component (or one or more of its sub-components) of economic analysis has been undertaken or not. We assign to the variable corresponding to this analysis a value of 1 in case the analysis has been undertaken and a value of 0 otherwise. The methodology (presented in detail in Katsoulacos, Avdasheva and Golovanova, 2018) identifies four broad components or categories of economic analysis (A, B, C and D in Table 1) that must be performed for the investigation to constitute an effect-based (or rule-of-reason) analysis. One or more variables (sub-components) make up every one of these main categories of economic analysis (e.g. sub-components B.1, B.2, C.1 etc; see Table 1). Taking into account the fact that non-exploitative and exploitative conducts require different blocks of economic analysis, different analysis variables are used in each of these two conduct categories. We abstract from differences in the analysis variables that could be included in the assessment of each specific conduct type, among the many of the non-exploitative conducts category, assuming that these are about the same.

Table 1. Types of economic analysis taken into account in the construction of the indicators of the extent of economic analysis (or of the Effects-Based (EB) indicators)¹⁷

	<i>Non-exploitative conducts</i>	<i>Exploitative conducts</i>
A. <i>Conduct analysis</i>	A.1. Discussion of the nature and characteristics of the conduct	
B. <i>Market analysis*</i>	B1. Basic analysis of the market (for instance, this can include calculations of market shares without formal analysis of Relevant Market Definition using SSNIP or a similar approach) OR	
	B2. Formal market delineation using SSNIP (Hypothetical Monopolist test)	
C. <i>Evidence on restrictions of competition/harm imposed</i>	C1. Analysis undertaken in order to identify whether conduct has market power enhancing (e.g. through agreements) or exclusionary effects (e.g. in monopolization practices)	C1. Analysis undertaken to compare price with cost
	C2. Articulation of theory of harm to consumer welfare (without taking into account of efficiencies)	C2. Comparison of the prices of the dominant supplier with the prices in other markets
	C3. Analysis of potential efficiencies	C3. Comparison of the prices of the dominant supplier with the price of competitors
		C4. Excess profitability analysis
D. <i>More effects-based analysis to support robustness of C</i>	D1. Counterfactual analysis undertaken	D1. Analysis of potential justification of exploitative conduct
	D2. Balancing of potential anticompetitive effects of conduct with the efficiencies and determination of the final impact of consumers (or on total welfare)	D2. Analysis of welfare effects of exploitative conduct

Note that the value (1 or 0) of an analysis variable (e.g. of B.2 or C.2 etc) is based on a judgment whether the relevant analysis has been undertaken or not and it says nothing about the correctness or “quality” of the analysis or of the data used. In other words, the value of an analysis variable indicates *whether the competition authority, in the particular case, has tried to address the specific question associated with that analysis variable.*

¹⁷ For details see Katsoulacos et.al. (2018).

2.2. Effect-based scores and types of legal standards

The first thing one can do with the data from Table 1 is, of course, to construct *effects-based scores (EBS hereafter)* calculated as the sum of the analysis variables presented in Table 1 – with a minimum of 1¹⁸ and a maximum of 7. The question is, is it reasonable for undertaking empirical analysis to use data that aggregate scores over many different conducts (e.g. all the non-exploitative conducts)?

The answer is that a straight aggregation of scores across different conduct types will *not* provide indicators which we can use to undertake meaningful empirical analysis of the extent of economic analysis and type of legal standards adopted. To see this note that, such aggregate indicators cannot be used to measure meaningfully whether economic analysis is used “optimally” – since optimal LSs can only be defined at the level of each conduct. Also, such aggregate indicators cannot be used to make comparisons between different countries and over time – since the level of the aggregate indicator will depend on the composition of conduct types that will be different for different countries and will change over time, so, for example, an EBS indicator score of, say, 3.2 for both Russia and Canada certainly does NOT mean that the extent of economic analysis relative to some optimal level is the same in Russia and Canada given that the composition of conduct types may well be completely different between the two countries¹⁹. And, such aggregate indicators cannot be used to examine how changes in the economic analysis, if measured by changes in the value of the aggregate indicator, affect the annulment rate, since the latter is expected to be influenced by what “type” of economic analysis²⁰ is utilized and how this changes, while a given value of the aggregate indicator cannot reflect what “types” of analysis are utilized and, when the value of the indicator changes, what “type” of economic analysis is responsible for the change in the indicator’s value.

Empirical researchers can respond in two ways to the above difficulties for undertaking empirical analysis. One way is to increase the available data for *each* conduct type, e.g. by putting together different countries, and constructing indicators for *each* conduct type²¹, using a table (that may be very similar to Table 1) which identifies the analysis variables for the specific conduct type. This is not an approach without difficulties, one of which is the collection of the data in different countries. At this point we are collecting data on antitrust decisions in 7 countries and we

¹⁸ There must always be some discussion at least of the nature and the characteristics of the conduct.

¹⁹ E.g. in Canada there may be proportionally many more decisions on conduct types for which the appropriate LS is Per Se or close to Per Se.

²⁰ For example, different types of economic analysis can lead to a score 3 and different ways of increasing economic analysis can increase the score from 3 to 4 but the implications of each case for the rate of annulment may not be the same.

²¹ The amount of data for each conduct type, e.g. bundling, among those in the non-exploitative conducts category, is quite small for any one country for undertaking empirical analysis.

hope that in the future we will be able to aggregate data from these different countries and thus extend our sample and to follow up on this approach.

A second way²², that we report in this paper, in which we concentrate on the data of one country (Russia), is, rather than just use the EB scores described above, to *use EB scores that result from aggregation across conduct types but for which, when aggregating, we make sure that we assign the same score to different decisions only when the same amount and the same “type” of economic analysis is undertaken*. We follow this procedure below for all non-exploitative conducts together (these include horizontal and vertical agreements and exclusionary conduct²³).

According to this procedure, we use the analysis variables that describe the different steps of economic analysis that are utilized in antitrust investigations, ordered, as in Table 1 above, in a sequence that represents what most economists would recognize as successively increased application of economic analysis. That is, Table 1 describes *additional blocks of analysis applied, as we move from a Pure Per Se to a Full Effects-based LS*. This is very useful when we come to map the extent of economic analysis applied in a specific case to the legal standard adopted in that case. However, while we consider the *order of statements* above to reflect a common (or “natural”) order in which economic analysis is applied as we move from “low” (Per Se) to “high” (effects-based) legal standards, this order cannot be considered as unique for the assessment of *all* conducts in practice. Indeed, the statements described in Table 1 distinguish between what is a reasonably good set for all conducts *other than exploitative* and another set of statements, given for exploitative conducts²⁴.

Given these remarks, the aggregate EB indicator that will be used below is obtained by constructing the following *Sets of EB analysis (SEB)* which, hereafter, we will also term *Legal Standard Indicators (LSI)*, using the statements in Table 1:

S1: this contains all the infringement decisions in the sample in which we find “1” scores *just* for the A statement (for all other statements score is “0”).

²² Which is, indeed, complementary to also using data from different countries together.

²³ There are significant common elements in the assessment of these conducts to justify using a unified methodology for constructing EB-indicators. Of course we could distinguish (additionally) between two sub-categories of anticompetitive agreements and exclusionary conduct (and, can disaggregate even further) and construct EB-indicators for each of these more disaggregated conduct categories. As already noted, the main disadvantage of disaggregating further is that disaggregation leads to smaller samples with which to undertake statistical work.

²⁴ However, note that this does not necessarily imply, when in the text of a decision we find some analysis of a higher level (in the sequence), that lower level analyses *has been* included and has also been explicitly described in the decision text. This is particularly important with regard to the statements B relating to the Contextual Analysis of the Market and the Firms. Examination of a large data set of the Russian Authority decisions indicate that decisions may for example provide a description that aims to provide a “theory of harm” (under statements C.2 below) even though there is NO explicit analysis or description of the market characteristics under B. We believe, however, that an analysis putting forward a theory of harm even if it is not preceded by an *explicit* description of the market in the decision text, will be based on developing some understanding of market characteristics and conditions.

S2: this contains all the decisions in which we find “1” scores for the A statement *and* for the B statement (for all other statements score is “0”).

S3: this contains all the decisions in which we find “1” scores for the A statement and for the B statement and for the C1 statement (for all other statements score is “0”).

S4: this contains all the decisions in which we find “1” scores for the A statement and for the B statement and for the C1 statement and for the C2 statement (for all other statements score is “0”).

S5: this contains all the decisions in which we find “1” scores for the A statement and for the B statement and for the C1 statement and for the C2 statement and for the C3 statement (for all other statements score is “0”).

S6: this contains all the decisions in which we find “1” scores for the A statement and for the B statement and for the C1 statement and for the C2 statement and for the C3 statement and for the D1 statement (for all other statements score is “0”).

S7: this contains all the decisions in which we find “1” scores for the A statement and for the B statement and for the C1 statement and for the C2 statement and for the C3 statement and for the D1 statement and for the D2 statement.

Thus, by construction, our (new) aggregate EB-indicator with a value of 1 is represented by the set of decisions S1, that is, 1 is the value of the indicator when, in decisions, only block of analysis A is undertaken; our aggregate EB-indicator with a value of 2 is represented by the set of decisions S2, that is, 2 is the value of the indicator when, in decisions, only block of analysis A and B is undertaken; our aggregate EB-indicator with a value of 3 is represented by the set of decisions S3, that is, 3 is the value of the indicator when, in decisions, only block of analysis A, B and C1 is undertaken; etc.

To summarize, the sets of decisions $S_i, i = 1, \dots, 7$ described above and the corresponding value of the aggregate EB-indicator for each set are:

S1: {A} – aggregate LSI - of value 1.

S2: {A, B} - aggregate LSI - of value 2.

S3: {A, B, C1} – aggregate LSI - of value 3.

S4: {A, B, C1, C2} – aggregate LSI - of value 4.

S5: {A, B, C1, C2, C3} – aggregate LSI - of value 5.

S6: {A, B, C1, C2, C3, D1} – aggregate LSI - of value 6.

S7: {A, B, C1, C2, C3, D1, D2} – aggregate LSI - of value 7.

Now, by comparing the different sets of decisions, $S_i, i = 1, \dots, 7$, we can identify the effects of *additional* economic analysis. For example, by comparing decisions in S2 with decisions in S3 we can identify the effect of adding the block of analysis C1; by comparing decisions in S3 with decisions in S4 we can identify the effect of adding the block of analysis C2. We are also able

to identify the frequency with which the CA applies the analysis associated with each one of the sets in assessing different conduct types and, hence, infer the extent to which the CA favors a certain legal standard for the different conduct types (see below).

We can then map the SEB (or LSI) to specific legal standards as also described in detail in Katsoulacos et.al (2018) and shown in Table 2. A brief description of the LSs follows.

Under the *Strict Per Se (SPS)* LS the CA makes decisions on the basis only of the purely formal characteristics of the conduct under investigation, relying on strong presumptions about the implications of the general class of conducts to which the specific conduct belongs for welfare. Alternatively, one can say that with SPS the CA makes inferences about effects (on welfare) from the formal characteristics of the conduct and some basic analysis of the market.

The *Modified Per Se (MPS)* LS can be considered as a Per Se rule *subject to* a Significant Market Power requirement or, more generally, as supplementing Per Se by undertaking analysis of market characteristics as, for example, in assessing conducts under abuse of dominance or in an information exchange agreement or in a concerted practice for which there is no strong hard evidence of collusion. Alternatively, one can say that with MPS the CA makes inferences about effects (on welfare) from the formal characteristics of the conduct, detailed analysis of market characteristics and, depending on the type of conduct, the implications of these for incentives for achieving sustainable collusion and/or for the assessment of market power.

Truncated Effects Based (TEB) is a higher LS, under which, decisions about whether or not there is liability in the case of a specific conduct are reached by establishing that the characteristics of the specific conduct and of the market in which it is undertaken are such that it belongs to a class of conducts that distort the competitive process by *disadvantaging rivals* (i.e. through *exclusionary effects*, widely defined) *or by enhancing market power* (as in a concerted practice case) and, assuming an welfarist substantive standard, by establishing that the conditions present are such that a strong presumption can be made of adverse welfare effects. Alternatively, one can say that with a TEB LS the CA decides that there is liability by inferring adverse welfare effects from the potential of the conduct to distort the competitive process by *disadvantaging rivals* (i.e. through *exclusionary effects*, widely defined) *or by enhancing market power* (as in a concerted practice case).

Finally, *Full Effects Based (FEB)* represents the LS under which a finding of liability relies on all potential anticompetitive (exclusionary or market power enhancing) *and* also all potential pro-competitive effects of the *specific* conduct being assessed and compared²⁵ and a showing of adverse effects on *welfare* (consumer or total) of this specific conduct be then established.

²⁵ In summary and simplifying somewhat, under (strict) Per Se only conduct characteristics are examined and assessed, under MPS these are examined as well as market characteristics, under TEB additional analysis establishing

Table 2. Mapping *SEB* to legal standards

	<i>Presumed components of economic analysis applied in assessment</i>	<i>SEB (or LSI)</i>	<i>Legal Standards</i>
1	A	S1	Strict Per Se (SPS)
2	A and B	S2	Modified Per Se (MPS)
3	A, B, C1	S3	Truncated Effects-Based (TEB)
4	A, B, C1, C2	S4	Intermediate between Truncated Effects-Based and Full Effects-Based
	A, B, C1, C2, C3	S5	
	A, B, C1, C2, C3, D1	S6	
5	A, B, C1, C2, C3, D1, D2	S7	Full Effects-Based (FEB)

3. The Russian context of competition law enforcement and description of dataset

3.1 The Russian context of competition law enforcement

Russian competition law enforcement is recognized worldwide for its large scale. This has a number of dimensions. First, the national competition authority, the Federal Antitrust Service (FAS), has an unprecedented set of responsibilities, including antitrust enforcement, control over competition in public procurement and procurement of state-owned and regulated companies, tariff and sector-specific access regulation, unfair competition and advertising etc. Second, FAS is organized as a nexus of regional offices with relatively independent powers. There are 84 regional offices, responsible for enforcement within the borders of the Russian Federation (regions), and the Central Office, responsible for enforcement over the national market. In total, there are more than 3500 officers in FAS (about 1/3 of them work in the Central Office). Third, the number of FAS decisions is extremely high in all the areas of responsibility. For instance, in 2016 only FAS analyzed about 44,600 complaints on violation of antitrust law and made 6,900 infringement decisions, considered more than 20,000 cases on violation of advertising law and made 6,700

exclusionary or market power enhancing effects is undertaken and under FEB the above are supplemented by additional analysis and evidence to establish the net effect of the specific conduct on some measure of welfare taking into account potential efficiencies.

infringement decisions. In the area of procurement FAS undertook about 10000 raids to monitor compliance with public procurement rules, examined about 48000 procurement procedures and found 16,500 cases of non-compliance. In the area of procurement of state-owned and regulated companies FAS reviewed 15000 complaints and found 6000 cases of non-compliance. Many observers in the expert and business community consider that the scale of activity is disproportional as compared to the quality of outcomes achieved. One indicator of this is the limited conclusiveness of the evidence applied in the antitrust infringement decisions. A large number of decisions made by FAS coexists with a large number of claims to annul infringement decisions. For instance, in 2016 there were 7000 claims to annul decisions in the commercial courts of first instance²⁶.

For FAS's antitrust decisions, insufficient economic analysis is thought as an important cause of weakness of the enforcement. In 2013, the Organization for Economic Co-operation and Development (OECD) reviewed Russian competition enforcement and highlighted the necessity to improve the quality of economic analysis in competition enforcement as a primary recommendation for future priorities of FAS. The empirical analysis of this paper supports the conclusion of the OECD review.

The insufficiency of economic analysis identified by the OECD review can be explained in various ways. For the the early years of enforcement, an explanation is that competition legislation arose before modern economic education in Russia. Immediately after adoption of the first competition law there were no competition officers in Russia who had taken courses in modern Microeconomics, Industrial organization and Institutional Economics. The first students that graduated under modern standards of economic education were able to be hired by competition authorities only in the beginning of 2000s. However in the interim period the standard approach to substantiate infringement decisions was already in place. Therefore, path-dependence alone could be an explanation of the continuing lack of economic analysis that we also identify in our empirical analysis. A more important and relevant explanation is related to the incentives of FAS management, which in turn depends on the criteria of performance measurement. The latter relies on a showing of a large number of decisions with low rate of annulments. Under limited budget FAS is expected therefore to be motivated to select the group of cases that do not require deep economic analysis, in order to avoid the cost of evidence collection, and also in order to limit the possible counter-evidence under judicial review that can result in annulment of its decisions. Some preliminary evidence (Avdasheva et al, 2016; Avdasheva et al, 2018) supports this explanation showing that FAS has been focusing on generating a very large number of decisions which, on

²⁶ Most of these decisions are not related to antitrust.

closer inspection, cannot be considered as “proper antitrust” decisions. There are three important features of *Non-Proper Antitrust Decisions* (compared to Proper Antitrust Decisions):

- (a) harm without competition concerns is considered sufficient proof of a violation;
- (b) no causal interrelation between market power and harm is required; and
- (c) conduct affects a very small group (even one person at the extreme).

In this article we concentrate on the proper antitrust decisions of FAS.

3.2 Specific features of the Russian dataset on competition enforcement

We use a dataset based on data on the judicial review of FAS antitrust infringement decisions during the period of 2008-2015. The first year of observation is the year following the new law ‘On protection of competition’ (better harmonized with European rules than the previous version was). Thus, the first observations in our dataset concern the first claims to annul decisions made according to the new law. We use the content on judicial decisions instead of a FAS dataset. There are two reasons for this. The first is that in Russia, commercial courts present the data on judicial review for all instances²⁷ in a very systematic and user-friendly manner. More importantly, the dataset of courts’ decisions covers all of the decisions made by commercial courts. For the dataset of FAS decisions, this is not a case. A second reason is that the outcomes of judicial review form a crucial part of our analysis. An important question is whether more extensive and/or deeper economic analysis makes decisions more convincing or whether the degree of the disputability of the decisions and hence the probability of annulment increases with the amount of economic analysis applied.

The total number of observations on proper antitrust decisions in our dataset is 1133 (including decisions on exploitative practices). The observations are claims to appeal infringement decisions made under art. 10 and 11 (including 11.1)²⁸ on the law ‘On protection of competition’. In the analysis below, we classify conducts by conduct groups. As mentioned above we concentrate on non-exploitative practices. Thus the conduct groups that we consider are:

- conduct group G1 consists of violations of art. 10 and art. 11, which have strong market power-enhancing effects. They include price fixing, bid rigging, boycotts, market sharing and exclusive territories;

²⁷ In Russia, there are four instances in commercial courts: first instance, appellate instance, cassation instance and highest (Supreme) court. Supreme Court has the discretion to consider or not consider cases. Supreme and cassation courts can send cases back to the first instance court. The same decision can result in a cassation instance. As a result, the number of instances regarding the case can be large (the maximum number in our dataset is 10).

²⁸ That concern agreements and abuse of dominance practices.

- conduct group G2 includes concerted practices and coordination;
- conduct group G3 are vertical restraints that, according to conventional wisdom, can have both competition restriction and welfare-enhancing effects.
- conduct group G4 consists of tying and non-price discrimination but by dominant companies and due to agreements between groups of market participants.

The total number of observations for the above conducts (i.e. excluding exploitative abuses) in our dataset is 987. However, for constructing our aggregate EB indicators (SEB, or LSI) following the methodology described in previous section, a number of decisions (in which “gaps” in economic analysis are present), cannot be taken into account. As a result, the number of observations in our dataset drops to 539.

In the analysis of legal standards in Russian antitrust enforcement, we also use groups of variables that reflect the following:

- Characteristics of violators (size) - this is captured by a dummy variable that depends on whether the company is among 400 largest undertakings of the Russian Federation;
- Whether or not monetary penalties are imposed on violators (MONPEN);
- Period-related characteristics that divide the overall period into two equal sub-periods: from 2008 to 2011; and from 2012 to 2015. The reason for creating two sub-periods is the so-called ‘third antimonopoly package’ – a list of changes to competition legislation that was adopted at the end of 2011 and which among other things also change the penalty regime and clarified the distinction between exploitative and exclusionary types of violations.
- Variable *REBPS* reflects whether or not the violation was examined under article 13. Russian competition law’s art. 13 is similar to the EC’s article 101(3). This allows agreements under art. 101, that would be otherwise characterized as Per Se violations (under art. 101(1)), to be rebutted by the defendants. Rebuttal requires that a number (of four) cumulative conditions hold, the showing of which requires significant additional economic analysis²⁹. While under art. 101(3) the burden of proof for rebuttal falls on the defendants, the CA is obliged to examine and assess their arguments in its decisions.
- Outcomes of judicial review, as an important type of independent external assessment of the analysis undertaken by FAS. We consider both annulments of infringement decisions in the courts of first instance and final annulments (which might be made by higher courts). We note here that even though the means of the two variables coincide (Table 3), and the correlation is rather high, in our sample, every 7th decision supported by the court of first

²⁹ See: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=LEGISSUM:126114&from=EN>.

instance is annulled by one of the higher courts, and at the same time, every 6th annulment decision is reversed by one of the higher courts.

- Number of instances that particular claims are considered and duration of the litigation (in days) as the indicators of the resources spent by courts, FAS and litigants considering the claim.

Table 3 presents descriptive statistics of all the variables that we use.

Table 3. Characteristics of infringement decisions in the dataset

Variables	Description	Number of observations	Mean ³⁰	St. dev.	Min	Max
	<i>Conduct group variables</i>				0	1
G1	price fixing, bid rigging, boycotts, market sharing and exclusive territories	539	0.226		0	1
G2	concerted practices and coordination	539	0.245		0	1
G3	RPM min, RPMmax, vertical exclusionary agreements, exclusive territories	539	0.267		0	1
G4	tying and non-price discrimination	539	0.262		0	1
	<i>Legal standards applied</i>					
LSI = 1	S1: {A}	539	0.379		0	1
LSI = 2	S2: {A, B}	539	0.241		0	1
LSI = 3	S3: {A, B, C1}	539	0.317		0	1
LSI = 4	S4: {A, B, C1, C2}	539	0,063		0	1
	<i>Outcomes of judicial review</i>					
ANNULF (annulment on the first instance)	=1 if the infringement decision is annulled by the court of first instance, =0 otherwise	539	0.466		0	1
FANNUL (final annulment)	=1 if the infringement decision is finally annulled after all instances, =0 otherwise	539	0.484		0	1
	<i>Costs of litigation</i>					
INST	=number of court instances that considered the claim to annul infringement decision and further appeals on the decisions of lower courts	539	2.924	1.429	1	10
DUR	=duration of court proceeding in days	539	376	236	31	1616
	<i>Legal environment</i>					
LP	=1 for cases initiated after the changes of the Russian antitrust legislation in 2011, = 0 otherwise	539	0.488		0	1

³⁰ The mean can be calculated using the numbers in Table 5 below.

REBPS	=1 if specific conduct is considered according to art.13 (that allows Per Se violations to be potentially rebutted) =0 otherwise	539	0.449		0	1
	<i>Violator-specific characteristics</i>					
SIZE	=1 if the alleged violator is one of the 400 largest companies in Russia, =0 otherwise	539	0.212		0	1
MONPEN	=1 if monetary penalties are imposed on the alleged violator, =0 if otherwise	539	0.191		0	1

Source: calculated by authors using dataset

4. Legal standards in Russian competition enforcement: empirical analysis

4.1 EBS scores and LSI scores over time

This section provides statistical analysis for the *EBS* and the *SEB* (or *LSI*) measured using the approach described above. The objective of this description is threefold. We want first, to assess the time trends in the application of economic analysis in Russian competition enforcement; second, to analyze the degree of standardization (consistency) of the approach that Russian competition authorities adopt regarding different allegedly illegal types of conduct; and third, to compare the empirical estimate of LSs (the *LSI*) to the LSs suggested by modern economic theory.

Table 4 summarizes intertemporal descriptive statistics of legal standards in the investigation of particular types conduct groups, as well as the outcomes of judicial review.

Table 4. EBS and LSI for different types of conducts and conduct groups (2008 – 2015)

		2008	2009	2010	2011	2012	2013	2014	2015	Av. for period
All non-exploitative conducts										
EBS	mean	2,31	2,25	1,86	2,06	2,13	2,34	2,17	2,01	2,11
	N	68	144	222	113	86	89	184	81	987
LSI (SEB)	mean	2,41	2,27	1,65	1,94	2,05	2,52	2,30	2,03	2,06
	N	46	78	152	67	39	48	81	28	539
Share of finally annulled decisions		0,44	0,40	0,52	0,48	0,48	0,52	0,38	0,28	0,44
By conduct groups										
Group G1										
EBS	mean	2,14	2,62	1,93	2,13	2,38	2,05	2,14	1,96	2,11
	N	7	29	80	24	29	37	90	56	352

LSI (Set)	mean	2,17	2,56	1,52	1,85	2,50	2,17	2,89	1,80	1,98
	N	6	16	48	13	8	12	9	10	122
Share of finally annulled decisions		0,86	0,48	0,53	0,46	0,38	0,51	0,41	0,25	0,44
Group G2										
EBS	mean	2,46	2,52	2,15	2,34	2,15	3,00	2,50	1,50	2,36
	N	24	48	33	32	13	3	2	2	157
LSI (Set)	mean	2,52	2,71	2,15	2,32	2,22	-	3,00	1,50	2,40
	N	21	35	33	31	9	-	1	2	132
Share of finally annulled decisions		0,21	0,33	0,67	0,41	0,69	0,33	0,50	0,50	0,43
Group G3										
EBS	mean	2,17	1,85	1,45	1,73	1,83	2,43	1,95	1,83	1,82
	N	12	54	58	40	24	23	38	12	261
LSI (Set)	mean	2,40	1,36	1,21	1,00	1,62	2,67	1,88	1,67	1,57
	N	5	22	43	16	13	15	24	6	144
Share of finally annulled decisions		0,42	0,35	0,41	0,53	0,38	0,43	0,55	0,33	0,43
Group G4										
EBS	mean	2,28	2,08	2,04	2,24	2,10	2,58	2,35	2,55	2,26
	N	25	13	51	17	20	26	54	11	217
LSI (Set)	mean	2,36	2,20	1,96	2,57	2,11	2,62	2,38	2,60	2,33
	N	14	5	28	7	9	21	47	10	141
Share of finally annulled decisions		0,56	0,62	0,55	0,53	0,60	0,62	0,20	0,36	0,47

Source: calculated by authors using dataset; N = observations

We can make a number of important observations. First, the extent of economic analysis, measured either by the *EBS* or *LSI* is very low (close to 2 when the maximum value is 7) – indicating that low (close to *Per Se*, specifically *MPS*) legal standards are adopted. Also, it does not increase over time. This conclusion of no definite trend over time is also true for decision annulment. The data differ for different time periods and different conduct groups. The share of finally annulled decisions increased up to 2013 and then started to decrease. In group B, the share of finally annulled decisions has tended to decrease. For other groups, the trends are not obvious.

4.2 Weighted Average Legal Standards (WALS) and indicators of legal uncertainty and of the quality of enforcement

Table 5 provides information about the share of the value (between 1 and 7) of the *EBS* and *LSI* indicators for our four main conduct groups. We observe that variations in both the *EBS* and *LSI* are very large, demonstrating absence of consistency in the use of legal standards in assessing any given group of conduct. For example, in Group G1, in 48% of the cases *LSI* is 1 (i.e.

the SPS LS is applied), while in another 29% of the cases LSI is 3 (i.e. a TEB LS is applied). This implies that, from the point of view of firms, there is a large degree of legal uncertainty in the enforcement of competition law that we also measure empirically and show in Table 6 together with a number of other indices.

Table 6 deepens the statistical analysis by presenting measures of the weighted average legal standard (WALS) adopted and of both the quality of enforcement and legal uncertainty. We consider several indicators of the deviation of legal standards from what is considered best practice in international antitrust enforcement: deviation from the WALS, from the legal standard with the highest share and from the two neighboring legal standards with the highest cumulative share in the sample.

WALS is calculated using Table 5: it is the sum of the values of LSI, each value weighted by its respective share³¹. The higher the WALS the closer is the LS to effects-based (with full effects-based requiring a value of 7). The first observation is that, for the conduct group (G1) that is traditionally illegal *per se* (price fixing and market sharing), the weighted average legal standard is, as expected, close (though not as close as it should) to 1 (the optimal value of the LS in this case). The second observation is that in all other cases (conduct groups G2, G3 and G4) the WALS is far from its theoretical optimum (which we assume to be full effects-based): in all cases, the CA (i.e. FAS) is choosing LSs that are much closer to Per Se than to effects-based.

Two indicators in Table 6 measure legal (un)certainty. The first is the index of the concentration of legal standards (the HHI concentration index calculated as the sum of the squared shares multiplied by 100), and the second is the standard deviation of LSI. The concentration of LSI is very low (with highest attained for conduct group G3) indicating that FAS' approach to assessment is not predictable. And a high concentration coexists with a large variation. This fact indicates that, although most decisions are made using standards that are on average closer to per se rules than effects-based, FAS oscillates between SPS and TEB.

Two indicators measure the quality (Q) of enforcement. The first expresses the distance of WALS from the theoretically "optimal" legal standard. The second, measures the distance of most typical legal standards used by FAS from the optimal. As the value of Q increases the quality of enforcement increases. We notice that the quality of enforcement is extremely low for all conduct groups except G1 (price fixing and market sharing) for which a Per Se approach is the appropriate one.

³¹ For example the value of WALS for G1 in Table 6 is 1,98. This is obtained, using Table 5 as: $0,48*1+0,16*2+0,29*3+0,08*4 = 1,98$.

Table 5. Values of EBS and LSI for the different conduct groups

		Value of the EBS and LSI indicators							N
		1	2	3	4	5	6	7	
Group G1									
EBS	Share	0,16	0,59	0,22	0,03	0,00	0,00	0,00	352
	N	58	206	78	10	0	0	0	
LSI (Set)	Share	0,48	0,16	0,29	0,08	0,00	0,00	0,00	122
	N	58	19	35	10	0	0	0	
share of missed observation= 0.65									
Group G2									
EBS	Share	0,13	0,37	0,50	0,00	0,00	0,00	0,00	157
	N	21	58	78	0	0	0	0	
LSI (Set)	Share	0,16	0,28	0,56	0,00	0,00	0,00	0,00	132
	N	21	37	74	0	0	0	0	
share of missed observations = 0.16									
Group G3									
EBS	Share	0,39	0,44	0,12	0,05	0,00	0,00	0,00	261
	N	103	114	32	12	0	0	0	
LSI (Set)	Share	0,72	0,08	0,12	0,08	0,00	0,00	0,00	144
	N	103	12	17	12	0	0	0	
share of missed observations = 0.45									
Group G4									
EBS	Share	0,10	0,59	0,25	0,06	0,00	0,00	0,00	217
	N	22	129	54	12	0	0	0	
LSI (Set)	Share	0,16	0,44	0,32	0,09	0,00	0,00	0,00	141
	N	22	62	45	12	0	0	0	
share of missed observations = 0.35									
Total number of LSI:		204	130	171	34				539

Source: calculated by authors using dataset

Table 6. Estimates of WALs, degree of legal uncertainty and indicators of quality of enforcement for particular conducts and conduct groups

	Price fixing and market sharing (Group G1) (Assumed optimal = 1)	Concerted practice (Group G2) (Assumed optimal = 7)	Vertical agreements (Group G3) (Assumed optimal = 7)	Exclusionary conduct (Group G4) (Assumed optimal = 7)
Number of observations	122	132	144	141
'Best practice' legal standards	1	7	7	7
Weighted Average Legal Standard, WALs (1 to 7)	1,98	2,40	1,57	2,33
LS with highest share, share in brackets	1 (0,48)	3 (0,56)	1 (0,72)	2 (0,44)
Two LSs with highest sum of two neighboring shares (sum of shares in brackets)	1,2 (0,64)	2,3 (0,84)	1,2(0,80)	2,3 (0,76)
Index of Concentration of LSs, I_{CON} (index of legal certainty)	34	42	53	33
Index of Uncertainty, I_U	0,18	0,21	0,26	0,18
Quality of Enforcement: $I_{Q,1}$, $0 \leq I_{Q,1} \leq 6$	5,02	1,40	0,57	1,33
Quality of Enforcement: $I_{Q,2}$, $0 \leq I_{Q,2} \leq 6$	6	2	0	1

Source: calculated by authors using dataset

$$I_{CON} = 100 * \sum_{i=1}^7 s_i^2; 14,3 \leq I_{CON} \leq 100$$

$$I_U = \text{standard dev. of shares}, 0 \leq I_U \leq 1.$$

$$I_{Q,1} = 6 - D1, D1 = \text{deviation of WALs from optimal}, 0 \leq D1 \leq 6$$

$$I_{Q,2} = 6 - D2, D2 = \text{deviation of LS with maximum share from optimal}, 0 \leq D2 \leq 6$$

A final comment concerning legal uncertainty is in order here. The LSs reported in Tables 5 and 6, reflect the economic analysis used by FAS. But these may not be the levels of economic analysis and LSs that are considered appropriate by the Courts. Thus, we can interpret findings in two different ways. One is to assume that FAS always predicts correctly the LS that Courts will apply, so then what is recorded in Tables 5 and 6, is also the level of economic analysis and the LSs considered appropriate by Courts. This in turn implies that the high degree of legal uncertainty faced by firms, in predicting the LS that will be used in assessing their conduct, is the outcome of the significant variation (between SPS and TEB) in the LSs adopted by the Courts.

An alternative, is that FAS can predict imperfectly the LSs considered appropriate by Courts, a more reasonable interpretation especially given that this is a young jurisdiction. So in

the, for example, 48% of the cases that in Group G1, the LSI = 1 (so an SPS LS is used) this is the legal standard adopted by FAS *but it is not* what would be for Group G1 the appropriate LS of Courts. We will rely on this interpretation when we discuss the effects of the LS adopted by FAS on the probability of annulment by appeal courts.

5. Econometric analysis of the effect of legal standards on the probability of annulment and on the costs of enforcement

5.1 Legal standard as a predictor of the outcome of judicial review

This sub-section aims to assess the impact of legal standards on the probability of annulment of FAS decisions (that is, on the outcome of the judicial review). As dependent variables, we use both the binary variable of FAS decision annulment by the court of first instance and the binary variable of final annulment. It makes sense to look both at the decision of the court of first instance and at the final decision, given that, because of the easy and inexpensive access to litigation, in Russia decisions of first instance courts are often appealed and reversed by higher courts.

What factors are expected to influence and hence are responsible for the annulment rate of appealed infringement decisions that we observe in practice? One important factor is that the more extensive is the economic analysis and so the higher the LSs adopted by competition authorities, the more opportunities exist for the alleged violators to criticize this analysis, attempting to show ambiguities, omissions or errors and, hence, to show the opposite effect to that shown by the CA. In other words, *disputability of economic analysis* increases with the legal standards applied³² thus increasing the probability of annulment. This is why we expect that if CAs are well informed and always choose the same LSs as the LSs that are the courts' optimal choices, there will be an increasing relationship between probability of annulment and LSs adopted.

However, there is another factor at work, that may lead to the observed probability of annulment to *decline* with LSs. This is the following. While CAs will not have incentives to choose

³² Neven (2006), looks at all the appeals against EC decisions in the period 1994 – 2006, and computes the proportion of cases in which the Commission prevailed (so decisions were not annulled). He finds a success rate of art.82 (abuse of dominance) decisions of 98% which, as he comments, “is striking” (for mergers and art.81 cases the fraction is much lower – 75%). To explain this, he notes that “Article 82 has remained focused on form, whereas the merger regulation and increasingly Article 81 (at least with respect to vertical agreements) are focusing on effects, which involves the development of economic theories and evidence. *Such differences in success rates are consistent with the view that the scope for disagreement (and decision annulment) is greater when economic theory and evidence are important.* This is probably the most important insight from (our findings).....” (my emphasis). The evidence presented by Neven is confirmed with a larger dataset (that considers EC decisions until 2016) by Katsoulacos, Makri and Metsiou (2018). See also Katsoulacos (2018a) for an extensive discussion.

LSs that are higher than those anticipated to be chosen by courts³³, they might find it optimal to apply legal standards lower than they expect that the courts will adopt. In doing so, they know that the probability of annulment will be higher than that if the “correct” LS was chosen³⁴ – the probability of annulment at the level of evidence associated with the lower LS will be higher than if the level of evidence associated with the “correct” LS was chosen – but this increase in the probability of annulment of using the “wrong” (lower) standard is outweighed by the reduction in cost as a result of using the lower standard. This implies that we can observe that a lower legal standard (applied by the CA) is associated with a higher probability of annulment, when the LS is lower than that considered appropriate by the Courts.

A U-shaped relation between LSs and probability of annulment can then be observed if, when the LSs considered optimal and adopted by courts are relatively low (say, MPS or TEB) the CA adopts in these cases even lower LSs (so the second effect above is present), while when the LSs adopted by courts are higher (higher than MPS or TEB), the CA follows the courts and also adopts the higher LSs (so then the first effect above applies)³⁵.

This discussion suggests that it is not possible to predict theoretically what the empirical relationship between the LSs adopted and the probability of annulment will be, as this depends on whether the CA anticipates correctly the LSs adopted by Courts and on whether it adopts lower LSs than those adopted by the Courts. Our empirical results, described below, are consistent with a U-shaped relationship in which the increase in LSs reduces the probability of annulment up to some threshold beyond which the probability of annulment increases.

In deriving our empirical results we also include the legal period and conduct-specific binary variables for the conduct group, decision-specific binary variables for monetary penalties and violator-specific (size of the company) control variables. Also we include the binary variable concerning whether or not the case was treated under art. 13 of the Russian law – as explained above, treatment under art. 13 implies enhanced use of economic analysis, which is expected to increase, *ceteris paribus*, the annulment rate (by increasing the disputability of decisions). Our expectations are that monetary penalties induce companies to exert greater efforts to provide more evidence for achieving annulments. The legal period is expected to exert the same effect, as the amendment in competition law after 2011 modernize antitrust assessment by making clear that restrictions of competition must be the basis for an infringement decision. Finally, we expect that larger companies can devote more resources to invest in the annulment of infringement decisions.

³³ Given that both litigation costs increase as LSs increase (move closer to effects-based from Per Se) and the probability of annulment also increases when CAs and courts move to higher LSs – by the argument just given above.

³⁴ I.e. the LS considered correct by the courts.

³⁵ The model in Katsoulacos (2018a) of a reputation maximising CA whose utility declines with the probability of annulment derives these results.

The baseline regression model for testing how annulment at the first instance court is influenced by the legal standard adopted has the following probit regression specification:

$$P(\text{annulf} = 1 | \text{appeal}, LSI_i, x) = F_1(\alpha_0 + \sum_{i=2}^4 \alpha_i LSI_i + \sum_M \alpha_M' X_M')$$

where $\text{annulf} = 1 | \text{appeal}, LSI_i, x$ indicates the probability that an infringement decision is annulled under judicial review in the first instance. The regression model for final annulment is similar.

Table 7 presents the results. One first thing to note is that evidence that increasing the extent of economic analysis increases the probability of annulment (by increasing the degree of disputability of the decisions) is provided by the statistically significant positive influence of the REBPS variable on the probability of annulment, given that REBPS captures cases examined under art. 13 which, as just noted, requires the application of more economic analysis.

The hypothesis of a U-shaped relationship between LSs and probability of annulment seems also to be confirmed, in Table 7, both for the decisions on annulment by the court of first instance and for the final decisions on annulment. The increase in the legal standard up to TEB ($LSI=3$) decreases the probability of annulment by the court of first instance court on average by 12%. For final decisions, the increase in the legal standard up to MPS and then up to TEB ($LSI = 2$ and $LSI = 3$, respectively) decreases the probability of annulment on average by 15% and 20% correspondingly. A further increase in legal standard to $LSI = 4$, seems to increase the probability of annulments relative to its value when $LSI = 3$ (by 14%) though this effect is not statistically significant until we distinguish between different types of conduct.

Table 7. Determinants of the annulment of infringement decision by the first instance court, and final annulment respectively (marginal effects are reported)

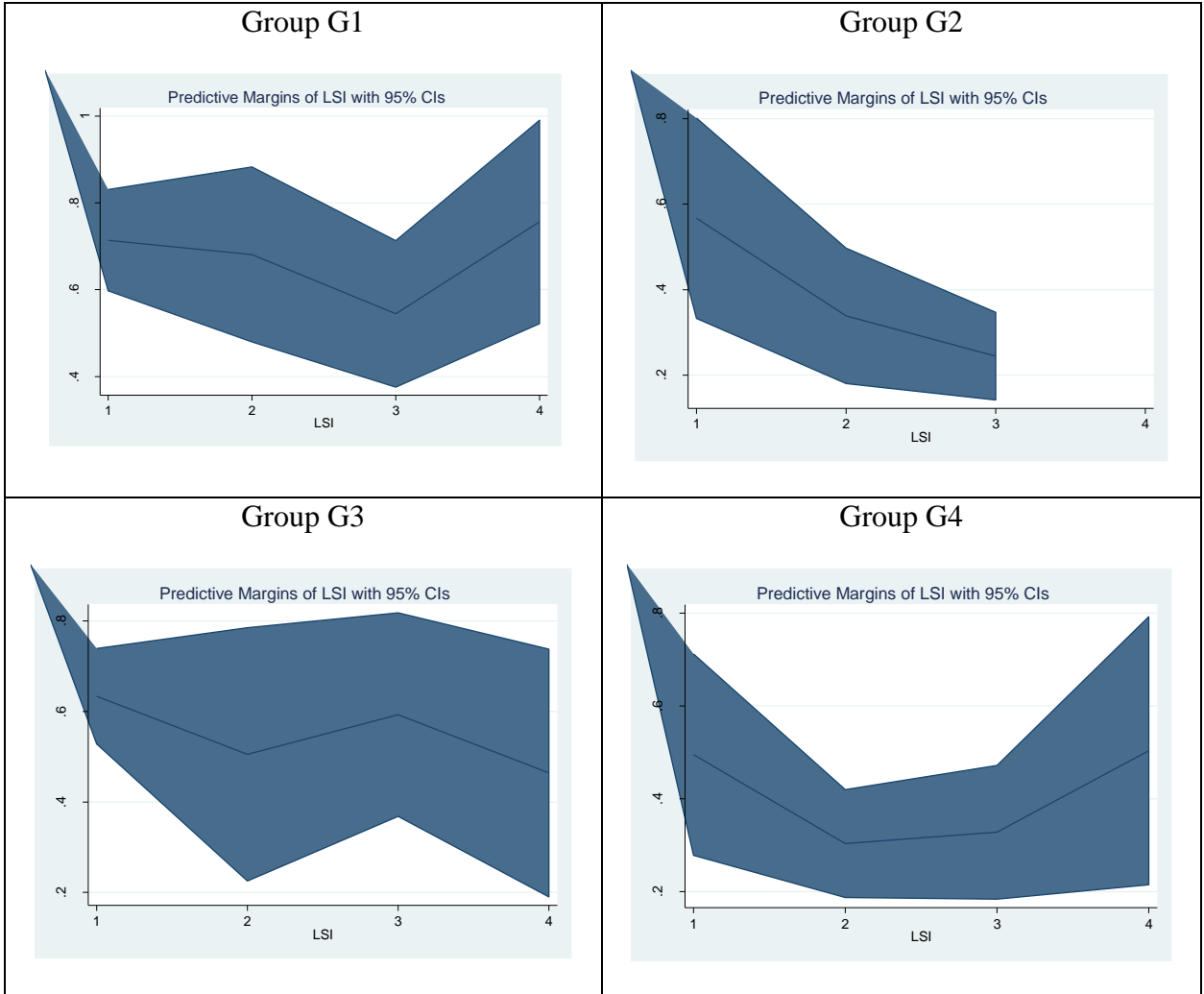
	First instance court	Final decision
REBPS	0.25*** (0.06)	0.29***(0.06)
LP	-0.06 (0.05)	0.00 (0.05)
MONPEN	0.11** (0.05)	0.08 (0.05)
SIZE	0.02 (0.05)	0.05 (0.05)
Set (LSI = 1 is base level)		
LSI = 1	base level	base level
LSI = 2	-0.07 (0.06)	-0.15* (0.07)
LSI = 3	-0.12** (0.06)	-0.20*** (0.06)
LSI = 4	-0.02 (0.09)	-0.06 (0.09)
Group control variables		
G1	base level	base level
G2	-0.42*** (0.08)	-0.36*** (0.08)
G3	-0.18** (0.06)	-0.07 (0.06)
G4	-0.40*** (0.07)	-0.30*** (0.08)
Number of observations	539	539
Prob F-stat	0.0000	0.0000

Pseudo R2	0.0633	0.0514
-----------	--------	--------

st. err. in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Graphs presented in Table 8 below visualize the estimated marginal effects on final annulment by groups of conducts.

Table 8. Marginal effects of LSI on probabilities of annulment by groups of conducts



The findings, at least for some conduct groups, are consistent with a U-shaped relationship in which the CA chooses a lower LS than what it expects courts will choose³⁶, when it expects that courts will choose MPS or TEB (LSI = 2 or LSI = 3), but chooses a higher LS, the same as courts, when the latter adopt higher LS (above MPS). Thus, at least for some groups of conduct (Group 1 and 4), a decrease in probability of annulment with LS turns to the opposite trend after some higher level of LS is reached.

³⁶ As explained above, this could be an optimal choice, balancing the increased probability of annulment when a lower (but wrong) LS is chosen with the reduced cost of the lower LS. The optimal choices of a reputation maximization, resource constrained CA are examined in detail in Katsoulacos (2018a).

Let us consider, for example, conduct group G4. Table 5 shows the share of decisions on this conduct group that are assessed by different LSs. Assume that courts consider optimal to assess this conduct group adopting LSI = 2, and sometimes a higher LS. According to the estimates we got the probability of final annulment Φ when FAS (correctly) uses LSI = 2 is 30% and when it uses LSI = 3 the probability is slightly higher and equals to 32%. More specifically, under Final Decision, the estimates show the following, when read in conjunction with Table 5 for conduct group G4:

- In 16% of the cases (from Table 5) that FAS uses the SPS LS (or LSI = 1) but it should use the MPS LS (LSI = 2) or the TEB LS (LSI = 3). For these cases, we have according to the estimates, $\Phi = 50%$ (20% higher than at LSI=2 and 18% higher than at LSI=3).
- In 44% of the cases (from Table 5) FAS uses the MPS LS (or LSI = 2) and the estimated probability of final annulment for these cases is $\Phi = 30%$.
- In 32% of the cases (from Table 5) FAS uses the TEB LS (or LSI = 3) and the estimated probability of final annulment for these cases is $\Phi = 32%$.
- In 9% of the cases (from Table 5) FAS uses the a LS higher than TEB (specifically, LSI = 4) when the Courts also use this LS (LSI = 4). For these cases, due to first effect above (disputability increases significantly), $\Phi = 50%$ (20% higher than at LSI=2 and 18% higher than at LSI=3).

So, as LSI for the LS used by FAS, increases from LSI = 1 to LSI = 4, the probability of annulment first falls from 50% to 30% and then increases a bit up to 32% but then it increases back to 50% - a U-shaped curve.

We provide some further discussion of the Table 7 results in Appendix 1.

5.2 Legal standards as a predictor of the costs of decisions

We use time, measured by days from the submission of a claim to the final decision, as an indicator of the cost of litigation. This approach makes sense since longer litigation is expected to imply higher costs in person-days, expert fees and management time, for all litigation parties, as well as for society. CA and companies bear the costs of litigation, in terms of additional resources used.

In addition to time variables, we use *the number of instances* appealed as another indicator that reflects the cost of litigation. From some perspectives, this indicator measures explicit litigation costs even better than the pure 'days' indicator because every additional instance incurs additional costs. Because of the evident dependence of duration on the number of instances that

claims for annulment are made, explicitly proven for the Russian commercial courts (Shastitko, 2016), we also use this effect to render the magnitude of influence more precise.

The main hypothesis that we test is that *litigation costs increase monotonically with legal standards*. This is expected as a result of higher LSs requiring a greater amount and more sophisticated economic analysis – something that follows from the very definition of effect-based procedures.

We apply OLS regression to test the hypothesis. In addition, we divide our sample into two sub-samples. The influence of effect-based analysis could differ depending on the value that the parties place on the decision been annulled – so we distinguish between cases with lower and higher values for the parties. For cases for which the value is high, parties exert greater efforts and resources to provide evidence. To differentiate between these two groups of cases, we use the submissions to the Higher (Supreme) court as the classification criterion. In Russia, submissions to the Supreme court is not very expensive. At the same time, the Supreme court is the only instance where discretion exists to consider or not to consider particular cases. The conditional probability of the case being considered and resolved is much lower for submissions to the Supreme court; therefore, the value of the case for the party should be greater, all other things being equal, to decide on submission. In our sample, 20% of the decisions of the lower courts were submitted to the Supreme court. The number of observations in the sub-samples is sufficient to compare these two groups.

The results presented in Table 9 confirm our hypothesis. In the sample, as well as in both sub-samples, higher legal standards require additional days to resolve the case. Two results are worth mentioning. First, there is a monotonic increase in litigation costs with the increase in legal standards. Second, in the sub-sample of high value cases, increases in legal standards result in larger increases in costs. In the sub-sample of low value cases, only a shift to TEB analysis (LSI = 3 and LSI = 4) provides a statistically significant increase in the number of days. In contrast, in the high value cases, even a small shift from *per se* to *MPS* (LSI = 2) increases the number of days for the cases to be decided. Relevant coefficients are higher for high-value cases, especially for TEB analysis (LSI = 4).

Table 9. Determinants of the duration of judicial review (OLS regression)

	All claims	<i>Low value for parties</i>	<i>High value for parties</i>
REBPS	-3.08 (36.19)	-17.78 (34.76)	-416.56* (224.23)
LP	38.41* (21.88)	61.07*** (22.54)	-1.75 (45.36)
MONPEN	-13.46 (25.51)	-33.46 (26.12)	51.81 (54.78)
SIZE	71.01*** (24.68)	64.29** (26.48)	30.72 (49.97)
Set (LSI = 1 is the base level)			
LSI = 1	<i>base level</i>	<i>base level</i>	<i>base level</i>
LSI = 2	34.40 (29.75)	12.11 (29.98)	150.55** (69.60)
LSI = 3	106.94*** (26.77)	88.04*** (28.04)	129.85** (56.43)
LSI = 4	122.84*** (42.76)	93.03*** (41.84)	363.73*** (116.31)
Conduct group variables			
G1	<i>base level</i>	<i>base level</i>	<i>base level</i>
G2	6.98 (44.59)	43.21 (44.53)	301.38 (219.38)
G3	-97.53*** (28.94)	-51.30* (29.37)	-189.53*** (68.99)
G4	-88.36*** (37.86)	-49.242* (37.00)	281.37 (222.85)
Constant	343.44*** (24.98)	283.58*** (26.00)	674.84*** (58.69)
Number of observations	539	430	109
Prob F-stat	0.0000	0.000	0.0000
Adj R-squared	0.0969	0.0697	0.1467

st. err. in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Regression of the number of instances (Table 10) shows similar results. Higher legal standards induce more re-submission of claims to higher courts which is an important predictor of litigation cost. To put it otherwise, Table 10 shows that *the probability of appealing to higher courts is increasing in the legal standards.*

Table 10. Determinants of the number of instances under judicial review (OLS regression)

	All claims	Low value	High value
REBPS	0.12 (0.22)	0.01 (0.20)	-2.02 (1.26)
LP	-0.11 (0.13)	0.06 (0.13)	-0.39 (0.25)
MONPEN	-0.30* (0.16)	-0.44*** (0.15)	0.12 (0.31)
SIZE	0.56*** (0.15)	0.52*** (0.15)	0.00 (0.28)
Set (LSI = 1 is the base level)			
LSI = 2	0.10 (0.18)	0.08 (0.17)	0.13 (0.39)
LSI = 3	0.54*** (0.16)	0.41** (0.16)	0.38 (0.32)
LSI = 4	0.66** (0.26)	0.63*** (0.24)	1.24* (0.65)
Group control variables (G1 = base level)			
G2	-0.11 (0.27)	0.02 (0.25)	1.75 (1.23)
G3	-0.41** (0.18)	-0.12 (0.17)	-0.70* (0.39)
G4	-0.23 (0.23)	-0.02 (0.21)	1.74 (1.25)
Constant	2.82*** (0.15)	2.34*** (0.15)	4.64*** (0.28)
Number of observations	539	430	109
Prob F-stat	0.0000	0.0000	0.2775
Adj R-squared	0.0813	0.0630	0.0215

st. err. in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

6. Conclusions

For the purposes of this paper we collected and analysed a unique large dataset of 1133 antitrust infringement decisions reached by the Russian CA between 2008 – 2015, which were appealed to Courts for annulment. Our analysis contributes to the development of tools for the empirical measurement of the extent of economic analysis and legal standards in competition law enforcement. Our second contribution consists in a detailed empirical description of the role of economic analysis in Russian competition enforcement. Our approach allows us to precisely assess the role of economic analysis in comparison to judgments based on anecdotal evidence. We find both an insufficient degree of economic analysis, close to Per Se legal standards, as well as that it is not applied in a consistent way. In investigations of similar conducts, the legal standards applied vary quite substantially, creating high levels of legal uncertainty for firms that we also measure empirically and compare across different conduct groups. Overall, with the exception of price fixing and market sharing conducts, the quality of enforcement (measured by the deviation of the

LSs adopted from the theoretical optimum) has been and remains throughout the period investigated very low.

Next, our results are consistent with the existence of a U-shaped relationship between LSs and the probability of annulment. This can be supported theoretically as showing that FAS adopts the wrong (lower) LSs when courts' standards are intermediate (MPS or TEB LSs) but follow the courts when the LS is higher than TEB. Finally, we show that an increase in legal standards leads to additional costs for competition authorities. Specifically, in terms of the time necessary for litigation to be completed, decisions in which FAS applies effect-based take longer than decisions based on *per se* standards. And, higher legal standards induce greater re-submission of claims (a higher appeal rate) to higher courts which is one of the main predictors of litigation cost.

References

- Avdasheva, S., Golovanova, S., & Katsoulacos, Y. (2018). The Impact of Performance Measurement on the Selection of Enforcement Targets by Competition Authorities: The Russian Experience in an International Context. *Public Performance & Management Review*, 1-28.
- Avdasheva, S., Golovanova, S., & Korneeva, D. (2016). Distorting effects of competition authority's performance measurement: the case of Russia. *International Journal of Public Sector Management*, 29(3), 288-306.
- Baker, J. B. (2003). The case for antitrust enforcement. *Journal of Economic perspectives*, 17(4), 27-50.
- Blair, R. D., & Sokol, D. D. (2012). The Rule of Reason and the Goals of Antitrust: An Economic Approach. *Antitrust LJ*, 78 (2), 471-504.
- Carlton, D. W. (2007). Does antitrust need to be modernized?. *Journal of Economic Perspectives*, 21(3), 155-176.
- Gavil, A. I. (2008). Burden of Proof in US Antitrust Law. *Issues in Competition Law and Policy*, ABA Section of Antitrust Law.1, 125-157.
- Gerber, D. J. (2015).. Adapting the role of economics in competition law: a developing country dilemma. In: *The Economic Characteristics of Developing Jurisdictions: Their Implications for Competition Law*, Ed. By M.S.Gal and others. pp.248-264.
- Hovenkamp, H. (2018). The NCAA and the Rule of Reason. *Review of Industrial Organization*, 52(2), 323-335.
- Jenny, F. (2016). The Institutional Design of Competition Authorities: Debates and Trends. In *Competition law enforcement in the BRICS and in developing countries* (pp. 1-57). Springer, Cham.
- Jones, A., & Kovacic, W. E. (2017). Identifying Anticompetitive Agreements in the United States and the European Union: Developing a Coherent Antitrust Analytical Framework. *The Antitrust Bulletin*, 62(2), 254-293.
- Katsoulacos Y. (2018a) On the Choice of Legal Standards: a Positive Theory for Comparative analysis. DP available in: <http://www.cresse.info/default.aspx?articleID=3388>

Katsoulacos, Y. (2018b). A Note on the Concepts of Legal Standards and Substantive Standards (and how the latter influences the choice of the former). *DP*, available <http://www.cresse.info/default.aspx?articleID=3388>

Katsoulacos Y., Avdasheva S., Golovanova S. 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*, forthcoming in 2018.

Katsoulacos Y, G. Makri and E. Metsiou (2018). “Antitrust Enforcement in Europe in the Last 25 Years: Developments and Challenges”, forthcoming in *Review of Industrial Organisation*.

Katsoulacos, Y., Metsiou, E., & Ulph, D. (2016). Optimal substantive standards for competition authorities. *Journal of Industry, Competition and Trade*, 16(3), 273-295.

Katsoulacos, Y., & Ulph, D. (2009). Optimal legal standard for competition policy: a general welfare based analysis. *The Journal of Industrial Economics*, 57(3), 410-437.

Katsoulacos, Y., & Ulph, D. (2011). Optimal enforcement structures for competition policy: Implications of judicial reviews and of internal error correction mechanisms. *European Competition Journal*, 7(1), 71-88.

Katsoulacos, Y., & Ulph, D. (2016). Legal uncertainty, competition law enforcement procedures and optimal penalties. *European Journal of Law and Economics*, 41(2), 255-282.

Katsoulacos, Y., & Ulph, D. (2017). Regulatory decision errors, Legal Uncertainty and welfare: A general treatment. *International Journal of Industrial Organization*, 53 (2), 326-352.

Lehmkuhl, D. (2008). On government, governance and judicial review: The case of European competition policy. *Journal of Public Policy*, 28(1), 139-159.

Lianos, I. (2012). The emergence of forensic economics in competition law: Foundations for a sociological analysis. University College London, Centre for Law, Economics and Society , Research Paper Series: 5/2012

Neven D. J (2006) “Competition economics and antitrust in Europe”, *Economic Policy*, Volume 21, Issue 48, 1 October 2006, Pages 742–791.

Padilla, J. (2010). The elusive challenge of assessing information sharing among competitors under the competition laws. *Roundtable on Information Exchanges Between Competitors Under Competition Law*, 879-924.

Shastitko, A. (2018). Empirical assessment of the role of economic analysis in the Russian antitrust: Why is economic analysis used? *European Journal of Law and Economics*, 45(2), 313-330.

Schinkel, M. P. (2008) Forensic economics in competition law enforcement. *Journal of Competition Law and Economics* 4(1): 1-30.

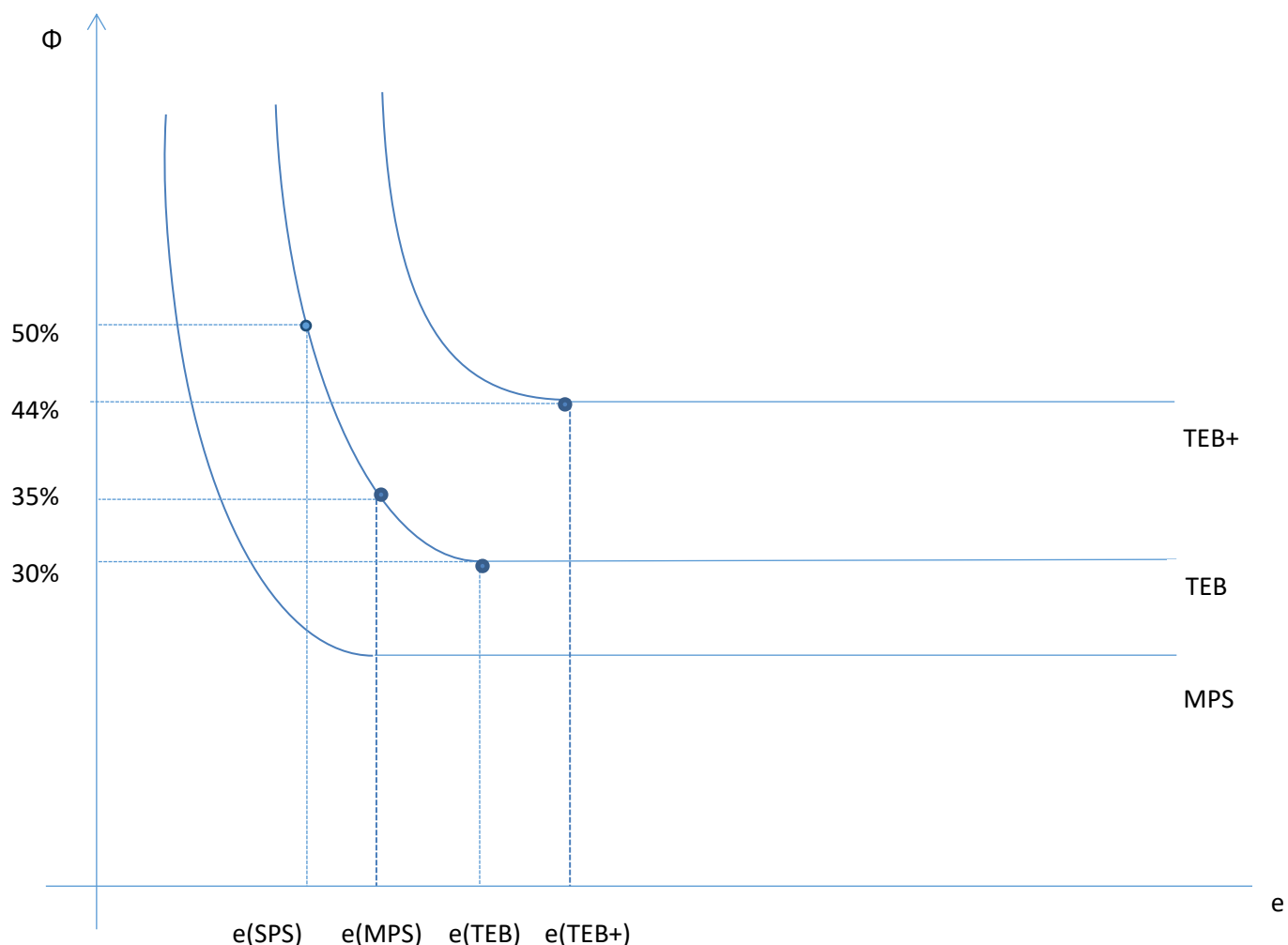
Trochev, A. (2016). Legitimacy, Accountability and Discretion of the Russian Courts. In *Politics and Legitimacy in Post-Soviet Eurasia* (pp. 121-147). Palgrave Macmillan, London.

Vickers, J. (2005). Abuse of market power. *The Economic Journal*, 115(504).

Appendix: Further discussion of results depicted in Table 7

Here we provide further discussion of the results on the relationship between the rate of annulment and the legal standard adopted in Table 7. Specifically, we can use the hypothetical example presented at Figure 1 (adapted from Katsoulacos (2018a)) to illustrate.

Figure 1



In Figure 1 the vertical axis shows the probability of annulment Φ , while the horizontal axis shows the optimal degree of economic analysis and evidence (e) associated with different LSs. Specifically in Figure 1 on the horizontal axis we show the optimal economic analysis associated with SPS, MPS, TEB and TEB+ (the last been the legal standard associated with Set 4 in text). The 3 curves in Figure 1 show how Φ varies with e , given the LS, for MPS, TEB and TEB+. Each curve shows that Φ declines as e increases up to the point where e reaches its optimal level for the LS represented by the curve. After that point, Φ remains constant.

Figure 1 shows that as the LS increases the value of Φ for any given e increases. Thus, consider some e , e.g. a value between $e(\text{MPS})$ and $e(\text{TEB})$ (so, higher than $e(\text{MPS})$) and lower than $e(\text{TEB})$). The value of Φ at this value of e , if MPS is adopted, is the same as for $e(\text{MPS})$, i.e.

it is $\Phi(e(\text{MPS}))$). The value of Φ at this value of e , if TEB is adopted, will be higher, specifically, equal to $\Phi(e(\text{TEB}))$ or higher – it is higher in Figure 1. $\Phi(e(\text{TEB})) > \Phi(e(\text{MPS}))$ because as we shift to a higher legal standard, increasing e , disputability of decisions increase.

In the hypothetical example presented at Figure 1 we assume that the optimal LS (adopted by Courts) is TEB or TEB+ and also that at $e(\text{TEB})$, $\Phi = 30\%$. The CA, though it anticipates that TEB is the standard to be used by Courts can find it optimal to use sometimes a lower standard (as explained in the text and shown in Katsoulacos, 2018a), specifically it uses sometimes $e(\text{SPS})$ making $\Phi = 50\%$ and sometimes $e(\text{MPS})$ making $\Phi = 35\%$. When it uses $e(\text{TEB})$ then $\Phi = 30\%$. If the CA uses $e(\text{TEB+})$ and TEB+ is the legal standard that should be used, then there is an increase in Φ to $\Phi = 44\%$ (due to the increase in disputability when moving from $e(\text{TEB})$ to $e(\text{TEB+})$).