EUROPEAN COMMISSION

Brussels, 17.10.2011
SEC(2011) 1216 final

COMMISSION STAFF WORKING PAPER

BEST PRACTICES FOR THE SUBMISSION OF ECONOMIC EVIDENCE AND DATA COLLECTION IN CASES CONCERNING THE APPLICATION OF ARTICLES 101 AND 102 TFEU AND IN MERGER CASES

Accompanying the document

Commission Notice

Best Practices for the conduct of proceedings concerning Articles 101 and 102 TFEU

{C(2011) 7329 final}

DG COMPETITION

BEST PRACTICES FOR THE SUBMISSION OF ECONOMIC EVIDENCE AND DATA COLLECTION IN CASES CONCERNING THE APPLICATION OF ARTICLES 101 AND 102 TFEU AND IN MERGER CASES

STAFF WORKING PAPER

1 SCOPE AND PURPOSE 4

2 BEST PRACTICES REGARDING THE CONTENT AND PRESENTATION OF ECONOMIC AND ECONOMETRIC SUBMISSIONS 6

2.1 Formulating the relevant question 7

2.2 Data relevance and reliability 8

2.3 Choice of empirical methodology 9

2.4 Reporting and interpreting the results 11

2.5 Robustness (non implemented proposal: place robustness before reporting) 13

2.6 Further recommendations 13

3 BEST PRACTICES ON RESPONDING TO REQUESTS FOR QUANTITATIVE DATA 14

3.1 General motivation for Data Requests 15

3.2 Common elements of a Data Request 16

3.3 Main criteria to consider when responding to a Data Request 17

3.3.1 Completeness 17

3.3.2 Correctness 18

3.3.3 Timely submission 18

3.4 Other Recommendations 19

3.4.1 Cooperation in good-faith 19

3.4.2 Early consultation with the Commission to inform about what type of data is available 19

3.4.3 Consultation on a Draft Data Requests and data samples 20

3.4.4 Transparency regarding data collection, formatting and submission 20

3.4.5 Direct access 21

1 SCOPE AND PURPOSE

1. Economic analysis plays a central role in competition enforcement. Economics as a discipline provides a framework to think about the way in which each particular market operates and how competitive interactions take place. This framework further allows formulating the possible consequences of the practices under review, whether a merger, an agreement between firms, or single firm conduct. In certain cases it may also provide tools to identify the direction and magnitude of these effects empirically, if appropriate and relevant. In a number of cases, economic analysis may involve the production, handling and assessment of voluminous sets of quantitative data, including, when appropriate, the development of econometric models[1].

2. Economic analysis needs to be framed in such a way that the Commission and the EU Courts can understand and evaluate its relevance and significance. As an administrative authority the Commission is required to take a decision within an appropriate or sometimes a statutory time limit. It is therefore necessary to: (i) ensure that economic analysis meets certain minimum technical standards at the outset, (ii) facilitate the effective gathering and exchange of facts and evidence, in particular any underlying quantitative data, and (iii) use in an effective way reliable and relevant evidence obtained during the administrative procedure, whether quantitative or qualitative.

3. In order to determine the relevance and significance of an economic analysis for a particular case, it is first necessary to assess its intrinsic quality from a technical perspective, i.e. whether it has been generated and presented in a way that meets adequate technical requirements prevalent in the profession. This involves, in particular, an evaluation of whether the hypothesis to be tested is formulated without ambiguity and clearly related to facts, whether the assumptions of the economic model are consistent with the institutional features and other relevant facts of the industry, whether economic models are well established in the relevant literature, whether the empirical methods and the data are appropriate, whether the results are properly interpreted and robust and whether counterarguments have been given adequate consideration.

4. Second, one must assess the congruence and consistency of the economic analysis with other pieces of quantitative and qualitative evidence (such as customer responses, or documentary evidence)[2].

---

[1] The assessment of mergers and potential infringements "by effect" often requires a complex economic assessment by the Commission, as well as the use of statistical or econometric analysis.

[2] Economic models or econometric analysis, as is the case with other types of evidence will rarely, if ever, prove conclusive by themselves. The Commission can always take into account different items of evidence. The General Court has held that "*It is the Commission's task to make an overall assessment of what is shown by the set of indicative factors used to evaluate the competitive situation.*

5. The present document formulates best practices concerning the generation as well as the presentation of relevant economic and empirical evidence that may be taken into account in the assessment of a case concerning the application of Articles 101 and 102 of the Treaty on the Functioning of the European Union (TFEU)[3] or merger case[4]. These Best Practices are organised along two themes.

   i) First of all, it provides recommendations regarding the content and presentation of economic or econometric analysis. This is meant to facilitate its assessment and the replication of any empirical results by the Commission and/or other parties.

   ii) Second, the document provides guidance to respond to Commission requests for quantitative data[5] to ensure that timely and relevant input for the investigation can be provided.

6. The desire to ensure transparency and accountability, these Best Practices apply to all parties involved in proceedings concerning the application of Articles 101 and 102 TFEU and mergers, that is the parties to the case and interested third parties (including complainants), as well as the Commission.

7. These Best Practices do not create any new rights or obligations, nor alter the rights and obligations which arise from the TFEU, secondary EU law and the case-law of the Court of Justice of the European Union. The Best Practices also do not alter the Commission's interpretative notices and established decisional practice.

8. The principles contained here may be further developed and refined by the Commission in individual cases when appropriate in light of future developments. The specificity of an individual case or particular circumstances may require an adaptation of, or deviation from, these Best Practices. The recommendations contained in this document should be interpreted in light of procedural and resource constraints.

It is possible, in that regard, for certain items of evidence to be prioritised and other evidence to be discounted. That examination and the associated reasoning are subject to a review of legality which the Court carries out in relation to Commission decisions on concentrations". See Case T-342/07, Ryanair v Commission, [2010] paragraph 136

[3] Proceedings before the European Commission concerning Articles101 and 102 TFEU, in accordance with Council Regulation (EC) No 1/2003 of 16 December 2002 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty (OJ L 1, 4.1.2003, p.1, as amended).

[4] Proceedings under the Council Regulation (EC) No 139/2004 of 20 January 2004 on the control of concentrations between undertakings (OJ L 24, 29.1.2004, p. 1).

[5] Quantitative data means, generally, observations or measurements, expressed as numbers. For the purposes of these Best Practices, this concept is used to refer to large sets of quantitative data submitted and/or obtained for the purposes of the conduct of an assessment of an economic (and often econometric) nature.

2 BEST PRACTICES REGARDING THE CONTENT AND PRESENTATION OF ECONOMIC AND ECONOMETRIC SUBMISSIONS

9. Economic reasoning is employed in competition cases notably in order to develop in a consistent manner or, conversely, to rebut because of its inconsistency, the economic evidence and arguments in a given case.

10. Any economic model which explicitly or implicitly supports a theoretical claim must rely on assumptions that are consistent with the facts of the industry under consideration. These assumptions should be carefully laid out and the sensitivity of its predictions to changes to the assumptions should be made explicit. While it is not necessary for economic submissions to actually formalize verbal arguments in a model, this will sometimes be helpful to clearly spell out the assumptions underlying an argument, to check its logic consistency, to assess effects of a high degree of complexity, or to use the model as the theoretical basis for an empirical estimation[6].

11. An economic analysis may support an assessment of the anticompetitive or pro-competitive effects of a merger. Such analysis usually involves a comparison of the actual or likely future situation in the relevant market with the absence of the proposed merger.

12. By their very nature, economic models and arguments are based on simplifications of reality. It is therefore normally not sufficient to disprove a particular argument or model, to point out that it is "based on seemingly unrealistic assumptions". It is also necessary to explicitly identify which aspects of reality should be better reflected in the model or argumentation, and to indicate why this would alter the conclusions.

13. In many cases, economic theory is used to develop a testable hypothesis that is later checked against the data. In that case, the economic analysis makes predictions about reality that can be tested by observations and potentially rejected or verified. Thus, whenever feasible, an economic model should be accompanied by an appropriate empirical model - i.e. a model which is capable of testing the relevant hypotheses given the data available.

14. Very often simple but well focused measurement of economic variables (prices, cost, margins, capacity constraints, R&D intensity) will provide important insights into the significance of particular factors. Occasionally, more advanced statistical and econometric techniques may provide more useful evidence[7]. In any case, otherwise

---

[6] If an economic submission is well-reasoned, then the fact that a particular argument is "theoretical" or "general" is often a strength rather than a weakness of the submission. This is the case when one has deduced a general conclusion (which holds irrespective of the precise magnitudes of the parameters of the analysis) from a set of assumptions that are considered consistent with the facts of the case. For instance, an economic submission may try to substantiate that irrespective of the size or existence of efficiencies, a particular conduct cannot possibly harm consumers.

[7] For instance, an econometric analysis of the extent to which prices of an undertaking have been affected by the observed entry of a competitor may provide evidence of the competitive constraint exercised by that entrant. In turn this could provide insights with respect to the likely degree of harm, that would result if an incumbent dominant undertaking were to engage in practices resulting in anticompetitive foreclosure in that or related markets.

valid economic analysis may not always produce unambiguous results when applied to the facts of a competition or merger case. Contradictions may result from differences in the data, differences in the approach to economic modelling or in the assumptions used to interpret the data or differences in the empirical techniques and methodologies.

15. The following sections provide practical advice on the generation and communication of economic and econometric analyses. The goal of these recommendations is to ensure that every economic or econometric analysis developed by any party involved submitted for consideration in a case states to the largest possible extent the economic reasoning and the observations on which it relies and explains the relevance of its findings and the robustness of the results. This should allow the Commission and all interested parties to scrutinise the economic evidence submitted during the proceedings so as to avoid that empirical results that are not robust be disguised as such and key assumptions in theoretical reasoning be presented as innocuous. Economic or econometric analysis that does not strictly meet the standards set out in these Best Practices will normally be attached less probative value than otherwise and may not be taken into consideration.

2.1 Formulating the relevant question

16. The first step in any economic analysis, theoretical or empirical, is the formulation of a question that is relevant to the case at hand.

17. The question of interest should be:

    (a) precisely formulated so that its answer can be interpreted without ambiguity,

    (b) properly motivated taking into account the nature of the competition or merger case, the institutional features of the markets under consideration and the relevant economic theory[8].

18. An economic or econometric report should explicitly formulate not only the hypothesis to be tested (the "null hypothesis"[9]) but also the alternative hypothesis (or hypotheses) under consideration, so that rejection of the null hypothesis can be properly interpreted[10].

---

[8] Occasionally the parties might submit a literature survey or review regarding an economic question of particular relevance for the case. A literature review may be useful when it is accompanied by an explanation on the merits and shortcomings, of the existing studies and explains how the party's own reasoning or analysis relates to past research, academic or otherwise.

[9] The null hypothesis is generally that which is presumed to be true initially. A null hypothesis is a hypothesis set up to be nullified or refuted in order to support an alternative hypothesis.

[10] For example, consider an empirical project aimed at testing whether certain conduct would lead to higher prices. One could define as the null hypothesis that prices did not increase in which case a rejection of the null hypothesis would imply that the agreement had a positive price impact. Alternatively, one could have defined as the null hypothesis that prices did not change as a result of

19. Sometimes, an empirical exercise which is being carried out may provide only partial verification of an accompanying economic model or theory of competitive effects. This evidence may be nonetheless useful but should be properly qualified[11].

2.2 Data relevance and reliability

20. The intrinsic quality of an economic theory depends on the extent to which the underlying assumptions match the corresponding economic facts. Likewise, empirical analysis depends on the relevance and the reliability of the underlying data.

21. First, it is necessary to identify the relevant facts to validate the theoretical assumptions and employ data which is appropriate to respond to the empirical question under investigation[12].

22. Second, not all facts can be observed or measured with high accuracy and most datasets are incomplete or otherwise imperfect. Hence, parties and/or the Commission should become familiar with the facts and data and acknowledge its limitations explicitly. As regards quantitative data, for example, this requires (i) a thorough inspection of the data, including summary statistics and graphs, and (ii) a sufficient understanding of how the data were gathered, the sample selection process, the measurement of the variables and whether they bear a close relationship with their theoretical counterparts. Quantitative data may contain anomalies because of miscoding or other errors, which should be discussed with the data providers to decide how to best adjust the data to address these problems.

23. Failure to observe and validate all key assumptions or deficiencies in the data should not prevent an economic analysis to be given weight, though caution must be exercised before relying on its conclusions[13]. Furthermore, statistical techniques have been developed to deal with measurement errors, missing observations and sample selection problems. While these techniques may not be able to improve the data, they may help to deal with some of its imperfections.

---

the agreement. A rejection of the null hypothesis in that case would be harder to interpret: did prices rise or fall as a result of the specific relationship between buyer and seller?

[11] For example, the analysis of scanner data (retail prices and quantities) may provide valuable evidence in the context of a merger between producers of fast moving consumption goods, even when the direct impact of the transaction would be felt at the wholesale level and not at the consumer level.

[12] For example when discounts are important, the analysis of the price impact of a merger, agreement or practice must focus on prices paid by consumers rather than on list prices.

[13] For example, assumptions regarding firms' expectations regarding the identity of the market leader may be inferred indirectly through observation of which firm first announces its future prices.

2.3 Choice of empirical methodology

24. The choice of methodology to empirically test a hypothesis or to validate the predictions of an economic model should be properly motivated, and its pros and cons should be made explicit, including potential identification problems[14].

25. Identification can be understood as clarifying the basis upon which one theory can be preferred to another. Similarly, the term can be used to refer to any situation where an econometric model will invariably have more than one set of parameters which generate the same distribution of observations.

26. One should explain how the chosen methodology exploits the variation in the data, to at least partially discriminate between the tested (or null) hypothesis and the alternative hypotheses. At the very least, an economic model or argument should generate predictions that are consistent with a significant number of relevant observed facts.

27. The choice of methodology must take due account of (a) the dataset and its potential limitations, (b) the features of the market under investigation, and (c) the economic issues under consideration — i.e., it should be designed to test the hypothesis of interest (see also section 2.1 above).

28. If statistical and/or econometric methods are used, it is strongly recommended that important methodological choices are explicitly justified, in particular:

    i) specification (what is the range of sensible general forms for the relationship under evaluation, including the relevant variables, the way they could interact, and the nature of errors or uncertainty?).

    ii) observation (how well do the measurements approximate the variables they are intended to represent?).

    iii) estimation (what do the data in the sample suggest as to the range of plausible relationships among variables?).

29. Moreover, a reasoned justification should be given when applying statistical techniques that deviate from generally accepted methods commonly used to assess the question of interest. In particular, one should motivate the changes, describe the modified technique or model, and document the likely biases, if any, that the new or adapted method is likely to introduce.

30. In general, it is recommended to follow a “bottom-up” approach. In the context of multiple regression analysis, this would mean estimating simple models first and then

[14] Problems of inference can be separated into statistical and identification problems. Studies of identification seek to characterize the conclusions that could be drawn if one could use the sampling process to obtain an unlimited number of observations. Studies of statistical inference seek to characterize the generally weaker conclusions that can be drawn from a finite number of observations.

engage in more refined estimation exercises if necessary in order to avoid bias[15]. Large-scale surveys of final consumers may usefully supplement qualitative or other documentary evidence obtained from targeted requests of information to market participants. Whilst the evidential value of replies to information requests from market participants lies in the substance of the information provided by players with intrinsic industry or market knowledge, the specific purpose of large-scale surveys of final consumers is to obtain statistically relevant data in order to estimate the characteristics, behaviour and views of a larger group of final consumers from the responses received from a smaller sample. The objectives of a high quality sample survey should be specific, clear-cut and unambiguous. Further, the definition of the relevant population of consumers (and the associated sampling frame) is crucial because there may be systematic differences in the responses of various differentiated consumer segments. Identification of a survey population must be followed by selection of a sample that accurately represents that population. The researcher can apply probability sampling in large-scale surveys of final consumers to some aspects of respondent selection to reduce the likelihood of biased selection[16].

31. The use of probability sampling techniques in large-scale surveys of final consumers enhances both the reliability and representativeness of the survey results and the ability to assess the accuracy of quantitative estimates obtained from the survey as regards the relevant population of consumers. Probability sampling in large-scale surveys of final consumers offers two important advantages over other types of sampling. First, the sample can provide an unbiased quantitative estimate of the responses of the relevant consumers from which the sample was drawn; that is, the expected value of the sample estimate is the population value being estimated. Second, the researcher can calculate a confidence interval that describes explicitly how reliable the sample estimate of the population is likely to be.

32. If possible, given time and data constraints, conducting multiple empirical analyses relying on different methodologies would help determine whether the conclusions of the empirical investigation are robust to different tests or models (see also section 2.5 below).

---

[15] For example, it is sound practice to estimate an Ordinary Least Squares (OLS) regression first and then, to the extent endogeneity is suspected to be a problem in the case at hand, move on to an instrumental variable (IV) estimation.

[16] Probability samples range from simple random samples to complex multistage sampling designs that use stratification, clustering of population elements into various groupings, or both. In simple random sampling, the most basic type of probability sampling, every element in the population has a known, equal probability of being included in the sample, and all possible samples of a given size are equally likely to be selected. In all forms of probability sampling, each element in the relevant population has a known, nonzero probability of being included in the sample.

2.4 Reporting and interpreting the results

33. The results of economic and econometric analysis must be presented clearly, taking the reader through each step of the reasoning[17]. All empirical analysis, even descriptive statistics of relevant variables (e.g. price series) should be accompanied by all the documentation needed to allow timely replication, as well as a deep understanding of the methodology of any prior data management efforts. Reports which do not allow for replication and in particular econometric analysis not including the code and data in electronic form will receive less consideration and are consequently unlikely to be given much weight.

34. An empirical submission should not only discuss the statistical significance of the results but also their practical relevance. In general, with very large samples coefficients may be statistically significant even if they are of trivial magnitude[18]. This creates the potentially misleading impression that certain variables are important. Therefore, the magnitude of the coefficients must always be examined and discussed. This requires interpreting the results in connection with the hypothesis that is being tested, so as to draw implications for the case under investigation.

35. Commonly, results from economic analysis and statistical information are presented in tables. Although it is not necessary to comment on or restate every piece of information that a table contains an interpretation of the data in it must be provided.

36. The results of the empirical analyses should be reported in the standard format found in academic papers. For example, when reporting multiple regression results, one should report on the statistical significance[19] of the parameter estimates by following the convention of reporting coefficients, p-values, standard errors and the size of the sample. Where the coefficient of interest is economically significant, the emphasis should be on statistically significant findings, for example to the 5% or 10% level (i.e. p-value<0.05 or 0.10). However, just because some hypothesis cannot be rejected in a statistical sense does not necessarily mean that the empirical analysis has no evidentiary value.

---

[17] Any mathematical notation should either (a) follow the standard notation in the literature or (b) be very self-explanatory.

[18] Statistical significance is determined, in part, by the number of observations in the data set. The more observations used to calculate the regression coefficients, the smaller the standard error of each coefficient. A smaller standard error reflects less random variability in the estimated coefficient (or estimate). Other things being equal, the statistical significance of a regression coefficient increases as the sample size increases. If the data set is sufficiently large, results that are economically significant are often also statistically significant. However, when the sample size is small it is not uncommon to obtain results that are economically significant but statistically insignificant.

[19] A statistically significant result is one that is unlikely to have occurred by chance. In hypothesis testing, the significance level is the criterion used for rejecting the null hypothesis. The p-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true. If the obtained p-value is smaller than or equal to the significance level, then the null hypothesis is rejected and the outcome is said to be statistically significant.

37. It may be that a particular analysis can be criticized in terms of its accuracy. However, it is often possible to evaluate that inaccuracy, for example by providing confidence intervals around an estimate. Also, depending on the question of interest, an approximate economic or econometric result can be informative if, for example, it is the direction of the effects rather than its magnitude that are most relevant. Similarly a particular estimate may be criticized because some facet of the methodology introduces bias. However, it is often the case that an estimate is biased in a particular direction; if this is the case it may be known that the estimate is too large, or too small. This may not matter in the context of a particular case. If it is known that the estimate is too large, and yet it is insufficient in size to reach some critical value, then the bias does not invalidate the conclusion that the critical value will not be reached. Detailed information should also be provided on all other specification tests and statistical diagnoses (see also section 2.5 on robustness).

38. The results of any statistical or econometric analysis should also be assessed with respect to the relevant economic theory[20]. When discussing the results of a multiple regression analysis, this requirement includes assessing not only the coefficient(s) of direct interest, but also the coefficients of all other explanatory variables, as they often provide a signal on the reliability of the analysis. For example, a finding that the sign of a particular coefficient is counter to what would be expected by economic theory[21] may be an indication of an omitted-variable problem[22], a selection bias[23], or some other identification problem[24].

39. In the case of large-scale surveys of final consumers the report should disclose essential information about how the research was conducted to allow judging the reliability and validity of the results. All data must be fully documented and made available (subject to appropriate safeguards to maintain privacy and confidentiality). Non-sampling error, in particular the non-response rate and response bias[25] should also be taken into account in the analysis. Conclusions from large-scale surveys of final consumers should be carefully distinguished from the factual findings.

---

[20] For example, econometric estimates of the elasticity of demand for a given product implying an upward sloping demand curve should be discarded in almost all cases, unless the product in question can be shown to be a Giffen good—i.e., a product for which a rise in price of this product makes people buy even more of the product.

[21] For example, a study showing that an increase in the marginal costs of production of a given good is associated with lower prices for that product should, ceteris paribus, be discarded automatically.

[22] That is, when a relevant explanatory variable, which is correlated with the dependent variable has been omitted from the analysis, so that the coefficients of some or all other explanatory variables suffer from a bias of a priori unknown sign or magnitude.

[23] The bias that arises when the selection process influences the availability of data in a way that is related to the dependent variable.

[24] See note 13 *supra.*

[25] Response bias refers to situations were, for a host of reasons, respondents fail to answer questions truthfully, fully and/or were influenced by the interviewer.

2.5 Robustness (non implemented proposal: place robustness before reporting)

40. Economic and econometric analysis should to the greatest possible extent be accompanied by a thorough robustness analysis, except where its absence is appropriately justified. In any event, any formal economic model or econometric analysis needs to be generally consistent and reasonably predict observed past outcomes and behaviour.

41. Other common robustness checks that may be appropriate include assessing whether empirical results are sensitive to changes in (a) the data, (b) the choice of empirical method, and (c) the precise modelling assumptions[26]. Similarly, the relevance and credibility of an economic model can be significantly enhanced if accompanied by a sensitivity analysis with respect to the key variables.

42. It is strongly recommended to address explicitly (i) to what extent, the results of the analysis are in line with past results using similar methods, and whether the results can be generalised[27]. Congruent and convergent results based on methods supported by academic and practitioners' are likely to be given greater significance than widely divergent results.

2.6 Further recommendations

43. The credibility of an economic submission may be enhanced when the limitations with regards to accuracy or explanatory power of the underlying data and methodology are explicitly acknowledged. In this regard it is often advisable to address rather than minimize uncertainty.

44. The parties rely sometimes on data that they do not have the means to audit and verify. Hence, they should be careful not to misleadingly present economic opinions as statements of fact. The sources of information should be carefully acknowledged, and the facts properly documented and described without ambiguity. This applies whether the economic or econometric analysis is a stand alone report or part of a broader submission.

45. It is advisable that the parties consult DG Competition regarding the types of empirical analyses that they consider useful in testing the anticompetitive and/or efficiencies theories. In particular, the parties can suggest potential analyses which may be easier for DG Competition to conduct, given its access to data from third

---

[26] For example, in a multiple regression analysis, one should indicate whether the results are severely affected by how the variables were defined, by the set of explanatory variables incorporated to the analysis, or the functional form.

[27] For example, if the elasticity of demand for a given product has been estimated for a given country, where data is available, but the case at hand would require estimates of the elasticity of demand for various countries, one should consider whether or not, and under which assumptions, her results for one country apply to the others. Similarly, if an economic model assumes that firms make take-it-or-leave-it offers when interacting with intermediate buyers with certain characteristics, it may be necessary to assess whether such assumption extends to all types of intermediate buyers.

parties. DG Competition, in turn, may propose analyses it believes might be useful for the parties to conduct. Similarly, it is recommended that the parties consult the DG Competition regarding the most suitable robustness checks for a given methodology. Experience suggests that such consultation can be most effective if the parties are prepared to share any relevant preliminary results in advance of a formal submission.

46. Where economic submissions rely on quantitative data the parties should provide the data and codes timely, in an appropriate format and in accordance with the criteria laid down in section 3 of this document. In particular, the absence of all the necessary elements needed for replication and assessment of an economic submission can constitute grounds for not taking it further into consideration.

47. When granting access to the file, the Commission may provide upon request the data and codes underlying its final economic analysis or, to the extent that they have been made available to the Commission, that of third parties on which it intends to rely or take into account. Where necessary to protect the confidentiality of other parties' data, access to the data and codes will be granted only at Commission premises in a so-called data room procedure[28], subject to strict confidentiality obligations and secure procedures[29]. Third parties or complainants are equally expected to submit all the underlying data used in the analysis. They are also expected to authorise the Commission, where appropriate, to offer data room access to the parties upon request.

48. When conducting large-scale surveys of final consumers to address a case-specific issue the parties might want to involve the Commission in the questionnaire development and design[30]. Subject to time and resource constraints it is often desirable to conduct a pre-test or pilot[31].

3 BEST PRACTICES ON RESPONDING TO REQUESTS FOR QUANTITATIVE DATA

49. Pursuant to Article 18 of Regulation 1/2003 and Article 11 of the Merger Regulation, the Commission is empowered, in order to carry out its duties, to require

---

[28] See Commission Notice on Best Practices for the conduct of proceedings concerning Articles 101 and102, paragraphs 97 and 98.

[29] Similarly, the Commission will endeavour to organise access to a data room, normally to the parties' economic advisors and external counsel, if necessary to ensure their rights of defence are fully respected.

[30] Occasionally, the Commission may take the initiative to commission its own large scale consumer survey. In that case, it will normally consult the parties and interested third parties on the questionnaire design and instruments of data collection, subject to confidentiality safeguards and to the extent such consultation does not delay or otherwise jeopardize the investigation.

[31] All questions should be pretested to ensure that (i) questions are understood by respondents, (ii) can be properly administered by interviewers, and (iii) do not adversely affect survey cooperation

undertakings and associations of undertakings to provide it with all necessary information. It is the Commission that defines the scope and the format of requests for information.

50. Most competition or merger investigations involve (1) collecting data, (2) analyzing data, and (3) drawing inferences from data. In most antitrust and merger cases, the Commission will gather evidence by sending targeted requests for information pursuant to Article 11 of the Merger Regulation and Article 18 of Regulation 1/2003 to the main players in the market (e.g. competitors, direct customers and other parties with specific knowledge of the market). This document, however, provides specific guidance to respond to a request for quantitative data[32]. However, many of the principles here identified apply, more generally, to responses to any request for economic information, quantitative or qualitative.

51. Quantitative data may help the Commission to conduct statistical analysis to define markets, establish a counterfactual, assess the potential anti-competitive effects of a notified merger, validate efficiency claims or predict the impact of remedies. In order to do that the Commission needs to get accurate data, with sufficient time to analyze it.

52. The Commission is aware of the costs that its procedures may impose on undertakings. An important objective of this section is, therefore, to provide recommendations to reduce the burden on the involved parties and on the Commission posed by the production and processing of quantitative data, while at the same time ensuring and enhancing the effectiveness of the Commission's substantive review.

53. These best practices are intended as general guidance and do not supersede any specific instructions in any Data Request issued by the Commission in specific cases.

3.1 General motivation for Data Requests

54. The primary objective of a Data Request is to obtain accurate information concerning quantitative variables such as prices, turnover, capacity and entry or exit decisions within the possible relevant markets over a reasonable period. Quantitative data may be necessary to understand current market conditions and competitive dynamics. In some cases, reliable quantitative data may allow to conduct statistical or econometric analysis to be submitted as evidence in an antitrust or merger investigation.

55. The Commission will endeavour to ask for the appropriate amount of data to carry out the required analyses. The Commission is mindful of time constraints and must balance the usefulness of each request against the time left before any legal or procedural deadline. In appropriate cases, DG Competition may discuss in advance with the addressees or other affected parties the scope and the format of the Data

[32] For statistical purposes, “quantitative data” means a series of observations or measurements, expressed as numbers. A statistic may refer to a particular numerical value, derived from the data. For example, an HHI measure and a correlation coefficient are statistics.

Request. DG Competition may also explain the analysis that it intends to perform with the requested data in order to improve the efficiency of the data collecting process and to ensure the data is of adequate quality. This is particularly the case in the later stages of an investigation as early requests could be of a more general nature and aimed primarily at better understanding the functioning of the market in question.

56. The Commission will carefully consider what the proper sample to characterize a population is. Inferences from the part to the whole are justified only when the sample is representative[33].

57. A further issue that may influence the scope of the Data Request is whether third party data will be necessary and available to conduct any meaningful analysis.

3.2 Common elements of a Data Request

58. Examples of data necessary for a competition investigation include data on costs, output, sales, prices, capacity, product characteristics, delivery flows, customer characteristics, tender details, entry barriers, business strategies, and market shares of the parties involved and of the other participants in the relevant market.

59. The source of the information can be the parties involved in the procedure, third parties, trade associations, trade press, independent consultants, survey information or government sources.

60. Data may be costly to collect or hardly accessible in the relevant time frame. Often, however, requests for quantitative data in merger proceedings seek data that is readily available to the involved parties. Readily available data refers to data that is routinely collected and maintained for a reasonable period as part of the firm's normal business operations, for example to inform business strategy or for internal reporting. Readily available data also includes data that is regularly purchased from third parties, such as scanner data or survey data[34]. In any event, in its investigations, the Commission is not limited to request only data that is readily available to the parties (see point 77 below). Deadlines for submitting data which is difficult or costly to retrieve will be decided by the Commission on a case-by-case basis.

61. A Data Request often includes the following sections, but each request will be tailored to the specific information needs and circumstances of the case:

    (i) a glossary of terms, in particular key variables;
    (ii) a list of the variables;

---

[33] For example, in certain circumstances it may be appropriate to limit the data request to a certain representative subset of the involved firms' customers, or to a particular geographic market which stands out for a valid given reason.

[34] Where econometric analyses are to be conducted, the sample needs to be of sufficient size for meaningful inference. For instance, in the absence of cross-section variability, requests would generally cover at least a three year period of monthly observations.

(iii) for each variable: the units of measurement; the level of aggregation over time (e.g. monthly); the time range (e.g. the last three fiscal years) and the geographic scope (e.g. countries, regions or cities);
(iv) the preferred electronic format (stata file, excel file, etc);
(v) suggestions or specific requests on data formatting, variable classification and tests to detect data inconsistencies;
(vi) deadline for compliance with the request.

62. In some instances, particularly where data is requested from different parties, DG Competition may provide a template to ensure all submissions are compatible and can be efficiently combined with minimal risk of error.

3.3 Main criteria to consider when responding to a Data Request

63. Responses to a Data Request must be: (i) complete, (ii) correct, and (iii) timely.

64. The Commission may impose on undertakings and associations of undertakings fines where, intentionally or negligently, they supply incorrect or misleading information or when, in response to a request made by decision, they supply incomplete information or do not supply information within the required time-limit[35]. Furthermore, in merger cases, the relevant time limits for initiating proceedings and for the adoption of decisions may exceptionally be suspended where, owing to circumstances for which one of the undertakings involved in the concentration is responsible, the Commission has had to request information by decision or to order an inspection[36].

3.3.1 Completeness

65. The parties should provide all data requested, in any of the stated formats and follow indications regarding presentation and consistency checks. Subsidiary data that is necessary to construct or to understand any variable requested should also be provided, except when adequately justified and with prior approval by the Commission.

66. It is strongly encouraged that problems of missing data are flagged to the Commission well in advance of the deadline for compliance with the Data Request to allow, if appropriate, for either a modification of the request or an extension of the deadline. Any data missing from the original Data Request must be adequately justified. In any event, a response to a Data Request may not be considered complete unless accompanied by a memo:

---

[35] Article 23(1)(a) and (b) of Regulation 1/2003 and Article 14(1)(a), (b) and (c) of the Merger Regulation.

[36] Article 10(4) of the Merger Regulation, but see also Article 8(6) thereof.

(i) describing the data compilation process: from raw data through aggregation and merging operations to the final database submitted. How was the sample selected and was it necessary to eliminate certain kinds of observations;
(ii) identifying all relevant sources;
(iii) labelling and thoroughly describing all variables;
(iv) reporting on the reasons for potential measurement error such as missing information or any changes in the collection process;
(v) describing any assumptions and estimations used to fill incomplete data; and
(vi) reporting on consistency checking and all data cleaning operations.

3.3.2 Correctness

67. It is up to interested parties to ensure the correctness of the data submitted. Tests for accuracy of all variables should always be undertaken and reported[37].

68. In order to detect incorrectness in data it will be expected that consistency checks are performed and documented prior to submission. In particular:

i) Responses to the Data Request should be consistent with responses provided to other requests for information (e.g. turnover, market shares, etc);

ii) Individual values within a variable must be consistent with the economic reality[38];

iii) When aggregation of raw data is necessary, one needs to ensure the aggregation algorithm is sensible and applied consistently;

iv) Coherence between different variables is necessary[39];

v) Over time consistency across and within variables must also be ensured.

3.3.3 Timely submission

69. Deadlines for responses to Data Requests must be strictly respected. Where parties plan to submit data in connection with an empirical analysis conducted at their own initiative, it is useful to warn in advance DG Competition of the planned timing and scope of such a submission. Results that the parties intend to rely upon or discuss in a

---

[37] For example, negative sales volumes or zero transaction prices are normally inaccurate and are often indicative of data extraction errors, systematic measurement errors or inadequate accounting of rebates or taxes.

[38] For example, transaction prices (net of discounts) should generally be positive, missing or unexpected values (i.e. sales not in line with historical levels) should be checked.

[39] For example, shipments of one product must be related to shipments of any by-products. Also, charged prices should generally remain above transportation costs (i.e. ex-works negative prices cast doubts on either the correctness of the charged price and/or the transportation cost).

meeting with DG Competition should be submitted, including data and code to facilitate replication, at least 2 working days before the said meeting.

3.4 Other Recommendations

70. This section sets down further recommended best practices concerning responses to a Data Request.

3.4.1 Cooperation in good-faith

71. Data production is an area where cooperation between the parties and the Commission is especially important. The parties will need to explain clearly the complexities that can be associated with requests that the Commission may regard as simple[40]. The Commission endeavours to define its requests as specifically and quickly as possible so the parties can understand what is being sought. This dialogue may help both sides deal more efficiently with data issues. In any event, it is for the Commission to decide the scope, format and timing of the Data Request.

72. It is important to emphasise in that regard that the integrity and efficiency of the process are undermined if, inter alia, the parties make representations about what data exist without reasonably diligent efforts to confirm their accuracy, if they ignore a carefully drafted and limited Data Request and produce large amounts of data points disregarding the submission format, scope, or data processing requirements, if they use non-obvious "definitions" of common terms in construing requests, or if they make unilateral and undisclosed inferences about what the Commission is effectively seeking.

3.4.2 Early consultation with the Commission to inform about what type of data is available

73. In some cases, the burden of compliance with Data Requests may be significantly reduced if the parties inform the Commission at the earliest opportunity on the availability of quantitative data. Early consultation allows to determine not only what data is available and its suitability, but also in what form it can be provided, thereby making it easier and faster for the parties to provide the data, in the event the Commission makes a Data Request. However, the Commission is not limited to request only data that is readily available to the parties.

74. To make these early discussions fruitful, parties must be prepared to thoroughly explain their information management systems and should be prepared to discuss certain issues such as: every field of information captured, how the underlying data is collected and formatted, the frequency of collection, what software is used, the size of the data set, what reports are routinely generated from that database, etc. It is

---

[40] Why, for example, it may be difficult, impossible or useless to simply "turn over" a "database," or the burdens and costs associated with providing data in the manner the Commission seeks.

recommended that the involved firms provide any written documentation and/or training materials to the Commission in advance of any discussion. It is also generally useful that parties create a diagram to show how the relevant data is distributed throughout the organization. In any event, as a general rule, parties should provide relevant documents to support their contentions concerning the availability, scope and production time of quantitative data.

75. Preliminary meetings or telephone conversations with those responsible for data collection or analysis in the firms are often quite useful. Parties are advised to make such personnel available as early as possible. These discussions should involve descriptions of the type of electronic (or other) data that the parties maintain (both in the ordinary course of business and what is archived, and in what form).

76. In the case of mergers, pre-notification discussions should routinely deal with data issues. Although, the Commission will endeavour to identify all issues that may require a Data Request as soon as possible, certain issues may not be identified until later in the proceedings.

3.4.3 Consultation on a Draft Data Requests and data samples

77. When appropriate and useful, DG Competition will send a "draft" Data Request for quantitative data in order to facilitate a better identification of the format, and to allow for basic consistency checks (see section 3.3.2). The purpose of the draft Data Request is to invite parties to propose any modifications that could alleviate the compliance burden while producing the necessary information. Any reduction on the scope of the Data Request can only be accepted if it does not risk harming the investigation and may trigger, particularly in merger cases, a reduction in the deadline for response initially anticipated.

78. In this connection, providing samples of the data is generally very helpful as it helps the Commission to determine what data is available and would be useful. As a result, on the basis of the sample it may be possible to draft a more focused Data Request, limiting the eventual burden on the parties.

3.4.4 Transparency regarding data collection, formatting and submission

79. A transparent process allows for all parties involved to be aware of any incidences during the data collection process and thus react more rapidly and effectively.

80. The parties are advised to submit quantitative data in a format that minimises the time and manipulation required to process the data for analysis. Parties should always be able to answer all the following questions:

    i) How applicable is the data to the analyses under consideration;

    ii) How reliable or "clean" is the data;

    iii) Is it enough to conduct a meaningful analysis;

iv) What institutional factors specific to the industry setting and/or company may impact the proper interpretation of the data?

81. The involved parties should draw the Commission's attention early on to any limitations in the data. They should make clear how raw data has been compiled and what steps have been taken to ensure its reliability[41].

82. The involved parties are also strongly encouraged to conduct their own descriptive analysis to detect data problems before submitting the data to the Commission. Also the Commission may sometimes welcome efforts by the involved parties to deal with any remaining data imperfections using statistical analysis. In some cases statistics allow in various ways to average out errors in measurement and yield statistically sound estimates. All such statistical analysis should be adequately reported. In any event, raw data should be provided wherever possible because the aggregation and cleaning of data may have a significant impact on the outcome of statistical or econometric analysis. Also parties should provide the program files that manipulate, clean and complete the raw data in preparation for the analysis.

3.4.5 Direct access

83. In some instances, the Commission will accept that as part of its response to a Data Request the involved parties provide direct electronic access to the underlying data. This alternative can provide an inexpensive and fast way to provide access to large amounts of data. Limited direct access can also provide a means to assess the value of certain corporate information.

84. The terms and conditions for direct access can be discussed in advance, addressing issues such as the availability of technical assistance, the ability to print or otherwise retrieve the data, the number of log-ins the company should provide, assurances that the activities of the services of the Commission will not be tracked, that underlying data will not be removed without agreement of the Commission and, most importantly, continued access throughout the entire course of the investigation. In limited instances, when providing direct access to corporate resources is unworkable, the Commission may submit a set of queries to the firm so that reports can be generated.

[41] For example, if the raw data is based on a sample of individual customer accounts, an explanation of how these accounts have been chosen and why they are representative of all customers should also be provided.

ANNEX 1

STRUCTURE AND BASIC ELEMENTS OF A SOUND EMPIRICAL SUBMISSION

This Annex briefly describes how to structure an empirical submission in a competition or merger case according with the principles set out in the preceding sections (esp. section 2 above). A sound economic or econometric submission should contain the following sections and elements:

A. The relevant question

- The research question must be: (i) formulated unambiguously and (ii) properly motivated, taking into account both the nature of the competition issue, the institutional features of the markets and industries under consideration, and the relevant economic theory.
- The hypothesis to be tested (or null hypothesis) must be clearly spelled out as well as the alternative hypothesis or hypotheses under consideration.

B. The data

- A clear description of data sources must be provided as well as hard copies of the databases employed in the analysis. Normally, an accompanying memo would describe how previous intermediate data sets and programs were employed to create the final dataset as well as the software code employed to generate the final dataset. All efforts made to correct for anomalies in the data should be clearly explained.
- One should also report how the data were gathered, the sample selection process, the measurement of the variables and whether they match with their theoretical counterparts, etc.
- In addition, the data should be thoroughly described. This includes reporting the sample time frame and the statistical population under consideration, the units of observation, a clear definition of each variable, any data cleaning procedures, etc. This information should be accompanied by descriptive statistics (including means, standard errors, maximums, minimums, correlations, and histograms, residual plots, etc) of all relevant variables.

C. Methodology

- The choice of empirical methodology should be properly motivated. One should discuss their methodological choices in light of: (a) their data limitations, (b) the features of the market under investigation, and (c) the economic issues under consideration (the relevant question).

– Alternative methodologies should also be discussed and if possible, given time and data constraints, employed to verify the robustness of the results to the choice of model. An economic model or argument must generate predictions that are consistent with a significant number of relevant observed facts.

D. Results and implications

– Parties should explain the details of their models, and share any documentation needed to allow timely replication (e.g. the programming code used to run the analysis).

– The results of the empirical analyses should be reported in the standard format found in academic papers. For example, when reporting multiple regression results, one should report both the estimated coefficients and their standard errors for all relevant variables. They should also provide detailed information on all other specification tests and statistical diagnoses.

– One should discuss not only the statistical significance of their results but also their practical relevance. This requires interpreting the results in connection with the hypothesis that is being tested, so as to draw implications for the case under investigation. The results of the statistical and econometric analyses should also be assessed with respect to the relevant economic theory.

E. Robustness tests

– All empirical work should be accompanied by a thorough robustness analysis that (i) checks whether the empirical results are sensitive to changes in the data, the choice of empirical method, and the precise modelling assumptions; (ii) tests whether the results of the analysis can be generalised; and (iii) compares the results of the empirical work in question with previous results in the relevant literature.

– An economic model should generally be accompanied by a sensitivity analysis with respect to the key variables, to the extent only the plausible but not the exact value of each variable can be determined. All results from the sensitivity analysis conducted should also be reported and not only those that support the argument.