

**DIRECTORATE FOR FINANCIAL AND ENTERPRISE AFFAIRS
COMPETITION COMMITTEE**

Working Party No. 2 on Competition and Regulation

**WORKSHOP ON EX-POST EVALUATION OF COMPETITION AGENCIES' ENFORCEMENT
DECISIONS**

-- Note by Tomaso Duso and Peter Ormosi --

26 October 2015

This paper by Prof. Tomaso Duso (DIW Berlin and DICE) and Dr. Peter Ormosi (UEA, Norwich) is circulated FOR DISCUSSION under Item IV at the 60th meeting of Working Party No. 2 on 26 October 2015.

The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

*More documents related to this discussion can be found at:
www.oecd.org/daf/competition/evaluationofcompetitioninterventions.htm*

Please contact Ms. Cristiana Vitale if you have any questions regarding this document [E-mail: cristiana.vitale@oecd.org].

JT03382670

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

**CAPACITY BUILDING WORKSHOP ON THE EX-POST EVALUATION OF COMPETITION
AUTHORITIES' ENFORCEMENT DECISIONS:
A CRITICAL DISCUSSION**

By Tomaso Duso and Peter Ormosi***

1. Introduction

1. In this paper we reflect on the debate that took place during the OECD's "Capacity building workshop on ex-post evaluation of competition authorities' enforcement decisions."¹ The purpose is to elaborate on the general discussion that emerged during the workshop, to highlight our opinions on the various topics touched upon, as well as to provide a synthesis of the different views expressed by the delegates. Naturally, the discussion proposed in this paper is not comprehensive and it is restricted to the content of the workshop.² Where possible, references are given for a wider and more in-depth discussion of the relevant topics.³

2. Following the structure of the workshop, the discussion can be grouped into two main parts: i) a general debate on the pros and cons of ex-post evaluation exercises with an emphasis on the lessons learnt from previous experiences and ii) a more micro-assessment of the specific tools for ex-post evaluation with their strengths and limitations, which was achieved by looking at three different case studies.⁴ The two central sections of this paper will cover these two parts separately. We will then conclude with a summary of the overall findings of the workshop and highlight the way ahead by identifying key questions that should be addressed by academics and policy makers.

* Tomaso Duso, Duesseldorf Institute for Competition Economics (DICE), Heinrich-Heine-University, Universitaetsstr. 1, 40225 Duesseldorf – Germany. Email: duso@dice.uni-duesseldorf.de.

** Peter Ormosi, Norwich Business School, University of East Anglia, Norwich – UK. Email: p.ormosi@uea.ac.uk.

¹ More information about the workshop can be found at:
<http://www.oecd.org/daf/competition/workshop-expost-evaluation-competition-enforcement-decisions.htm>.

² This paper assumes intermediate econometric knowledge. While we mention and discuss specific methodologies and tools we do not explain the most basic concepts.

³ An excellent overview of the existing literature on ex-post evaluation studies in competition policy can be found in the study by Ilzkovitz and Dierx (2015). The paper builds on several of the issues discussed during the OECD workshop which the two authors actively attended as delegates from DG Competition.

⁴ The text of the three decisions discussed during the workshop can be downloaded at:
<http://www.oecd.org/daf/competition/workshop-expost-evaluation-competition-enforcement-decisions.htm>.

2. The Pros and cons of Ex-Post evaluations: lessons learnt from ex-post evaluations

3. The use of ex-post or retrospective studies of competition policy enforcement has substantially increased over the last two decades. However, quite few of the over 110 competition authorities around the world are yet making use of this instrument. One of the specific purposes of the workshop was to understand how the experience of those authorities that do make use of ex-post evaluations can motivate and help other, less experienced institutions to follow this path. Moreover, the existing studies can already help us identify and define some best practices and institutional arrangements, which can be useful to conduct serious and successful retrospective studies.

4. All invited speakers at the OECD workshop were asked to highlight their experience with ex-post studies. Hence, their views represent those of the authorities that have already at least some experience with retrospective analyses — specifically, the US, the UK, the Netherlands and Mexico. During the debate, the delegates for the European Commission, DG Competition, and the external experts also presented their views. The delegates were asked to provide concrete examples of specific lessons learnt (e.g. lessons on how to assess effects on competition in a certain sector, lessons on the effectiveness of particular types of remedies, lessons on some factors that should have been taken into account during the ex-ante evaluation, etc.). We summarise and categorise the issues discussed in the next two sections. We first focus on the reasons and usefulness of ex-post evaluations and then turn to discuss some institutional details that might play a key role in how these studies might be best organised.

2.1 Why do we need ex-post evaluations?

5. A starting point that all delegates highlighted was that ex-post evaluations are quite positively perceived by those authorities that perform them, as well as by policy makers in general. Moreover existing experience so far— for instance from the US, UK, and even Japan— suggests that retrospective studies are well accepted by case handlers who are particularly keen to understand how the market evolved after the decision they have contributed to reach. Finally, there is also a widely-held consensus among academics that we need to evaluate policy to better understand whether the intervention was necessary and effective, so to make competition policy enforcement more effective. The discussion on this issue was quite elaborate and several reasons were mentioned why retrospective studies might be useful. We focus on the three main issues in the discussion.

2.1.1 Improve policy making: learning and correcting mistakes

6. First and foremost, ex-post evaluation studies are understood as a tool for better learning whether and why some particular policy actions have been successful and, thus, improve future decision making. The use of case studies that look at the nitty-gritty of the specific policy decisions and are tailored to analyse the specificities of the market under consideration, might be mostly important to better understand why that particular market reacted in that particular way. This might provide important insights and useful additional information that can be used in future cases.

7. For instance, Dan Hosken from the US Federal Trade Commission (FTC) highlighted the findings from the two largest series of ex-post evaluations conducted by the US FTC, one on hospital mergers⁵ and the second on in the petroleum industry.⁶ He stressed how the outcomes of the evaluation

⁵ A special issue of International Journal of the Economics of Business (2011) includes these studies and other hospital merger related papers.

⁶ The US FTC's Bureau of Economics conducted six studies of eight consummated mergers in the petroleum industry. These studies are reported in a series of FTC's Working Papers: 270, 278, 291, 300, 319 available

exercises had important policy consequences. In the late 1990s, the Department of Justice (DOJ) and the US FTC lost 7 consecutive hospital mergers. Afterwards, the US antitrust authorities essentially stopped challenging hospital mergers. This triggered the US FTC Chairman Muris to conduct retrospective analyses of a number of consummated hospital mergers. These studies revealed the shortcomings of using the Elzinga-Hogarty test in the hospital market, thus providing new robust empirical evidence that market definition had previously been too broad. Moreover, they showed that, similarly to for-profit hospitals, not-for-profit hospitals also exercise market power, and highlighted the importance of adequately modelling the bargaining between hospitals and insurers. All of these findings contributed to an improvement of US merger control in the hospital sector. Specifically, they have explicitly helped the US FTC to successfully challenge a consummated merger (FTC vs. Evanston/Northwestern/Highland Park) and, more in general, to support a very active prospective merger enforcement program. Indeed since 2008, six hospital mergers were blocked or abandoned.

8. US merger retrospectives in the petroleum industry were also quite useful, though in a very different way. The major concern in this industry during the 2000s was that increased concentration resulting from petroleum mergers that occurred in the late 1990s was responsible for increased gasoline prices. The series of studies conducted by the US FTC's staff did not find evidence of a significant increase in retail price following the studied mergers.⁷ Hence, the overall lesson that the US FTC learnt from these studies was that factors other than mergers were most likely responsible for large increases in gasoline prices. This had a significant impact on how the sector has been treated afterwards.

9. Similar comments were highlighted by the other delegates, although the experience with ex-post evaluation exercises in other jurisdictions is more limited than in the US. Mathieu Pearson, from the UK Competition and Market Authority (CMA), also mentioned how retrospective studies had important consequences for competition policy enforcement in the UK. Specifically, the review of eight merger decisions between 2004-2006 by Deloitte (Deloitte, 2009) helped shaping institutional details such as the new merger assessment guidelines. Similarly, part of the body of evidence collected during the evaluation of the UK Office of Fair Tradings' (OFT) 2008 market study into Personal Current Accounts fed into the 2014 OFT market study and subsequent referral for a market investigation.⁸

10. The Dutch Authority for Consumers and Markets (ACM) has started more recently to assess competition policy decisions. The Delegate Ron Kemp highlighted the very open attitude of the agency with respect to impact assessment and its willingness to improve and enlarge their ex-post evaluation efforts in the future because of the positive feedback it has had until now. Indeed, the Dutch ACM has been very active in the past few years both in running in-house and external studies, as well as by shaping the policy debate through the organisation and participation in international workshops.

2.1.2 *Evaluate the predictive power of methodological tools*

11. A second important role of ex-post studies is to inform competition agencies on the validity of the methodological tools potentially adopted in the ex-ante analysis. This is particularly relevant for the evaluation of merger policy where authorities are asked to make decision based on the speculative assessment of what the potential effects of a proposed merger in the future might be. The ex-post

at <http://www.ftc.gov/be/econwork.shtm> and are summarized in Hosken, Silvia, and Taylor (2011). Recently, The US FTC also reviewed several mergers in grocery retailing. See Hosken et al., 2013.

⁷ Other studies using different modelling approaches found only modest price increases following some of these mergers, e.g., Hastings (2004), Hastings and Gilbert (2005).

⁸ The OFT was one of the two UK competition agency together with the UK Competition Commission (CC). In 2014 these were merged to form the CMA.

evaluation of merger decision is, therefore, particularly important to understand not only whether this speculative exercise was successful, but also whether the instruments used to make this assessment were appropriate, precise, and effective.

12. For instance, the US delegate mentioned that several researchers have focused recently on verifying how accurately ex-ante merger simulations can predict observed ex-post outcomes. This kind of exercise can also be seen as an ex-post assessment.⁹ Commonly, a differences-in-differences analysis based on ex-post data is used to estimate the ‘true’ effect of the merger on prices and this is compared to the ex-ante estimates obtained by means of merger simulations using pre-merger data. While these exercises are useful, they are less often performed as a policy instrument but rather driven by academic interest, as they are quite costly to conduct both in terms of time and required skills. We will come back on this point in the next section dedicated to the methodologies.

13. Some other methodological lessons that can be learnt from retrospective studies are related to the kind of information the authority should be focusing on during the case investigation. For instance, one conclusion of the LEAR evaluation of two UK merger decisions, commissioned by the UK Competition Commission (CC) (Aguzzoni et al., 2011) was that the UK CC should give further thought to how it can obtain data and information from the parties after the merger took place, particularly with regard to non-price-information.¹⁰ This is a very important point, especially in markets where non-price considerations might play a crucial role such as in the creative goods’ industry.

14. Finally, retrospective studies can also help authorities better understand what specific economic market models are more adequate to represent specific industries. For instance, in the aforementioned case of the retrospective analysis in the US hospital industry, the series of ex-post evaluation studies led the US FTC— and indeed many academic scholars— to start explicitly modelling strategic interactions in the hospital industry as a bargaining process between hospitals and insurers. This was not done before and the previous, less appropriate, models of strategic interactions used to analyse hospital markets had led the US FTC to potentially erroneous conclusions on the extent of the insurers’ and hospitals’ market power. For example, before the ex-post evaluation they used to assume that not-for-profit hospitals do not exercise market power.

2.1.3 *Serve advocacy purposes and improve transparency and accountability*

15. Another central aim of ex-post evaluations appears to be its advocacy purpose. The UK delegate stressed the need for the UK CMA to externally and internally demonstrate the impact of their enforcement work and build reputation. The delegates of the European Commission equally emphasised how ex-post evaluations might be particularly useful in this respect. Moreover, retrospective studies are expected to improve the transparency of decision making, as well as the accountability of the agency, by informing politicians and other external observers about the rationale for the existence of competition policy and their effectiveness.

16. Also younger authorities underlined the advocacy role of ex-post evaluations. José Nery Pérez Trujillo of the Mexican Federal Economic Competition Commission (COFECE) explained that, although COFECE’s Competition Policy Evaluation Committee is very young and has been operating since 2014, the main aim of its evaluations is competition advocacy. The figures on consumers’ savings obtained

⁹ See for instance Peters (2006), Weinberg and Hosken (2008), and Weinberg (2011) with US data as well as Björnerstedt and Verboven (2012), and Friberg and Romahn (2012) with European data. See also Duso, 2012 for a discussion.

¹⁰ Two papers have been subsequently published from this report in peer-review academic journals: Aguzzoni et al., (2014), and Aguzzoni et al. (forthcoming).

through competition policy enforcement provided by ex-post evaluations can be easily communicated to the public to clarify the important role played by COFECE's activities. Furthermore, the communication of the results from retrospective studies provides elements to the public to report anticompetitive conducts and to deter firms from involving in this type of behavior.

2.2 *Institutional details matter*

17. While the implementation of ex-post evaluation studies might face quite some challenges ranging from methodological issues, to the lack of in-house expertise, limited resources, etc., the most positive message that emerged from the workshop is that a significant stock of knowledge now exists. Even less experienced institutions can build on this newly established common knowledge to explore the potential for conducting their own retrospective studies. Clearly, there are several institutional arrangements that affect the effectiveness, not only of competition policy enforcement, but also of ex-post evaluations. The discussions on what set of rules or which kinds of approaches are better suited to produce high quality retrospective studies was lively, as institutional differences among countries are still large notwithstanding a certain convergence.

18. Our view as academic experts who have been heavily involved in the ex-post evaluation of competition policies over the past decade is that these exercises should nowadays be seen as an integral part of the enforcement of competition policy. As we discussed above, their role is so indispensable in the modern, evidence-based enforcement of public policies that every competition authority should—according to their possibilities—engage in retrospective studies. They should be run on a regular basis and according to minimum quality standards that have been set by the international community in several recent publications. Even more specifically, the OECD is going to publish at the end of 2015 a Reference Guide on the ex-post evaluation of competition policies agencies' enforcement decisions that builds on the existing stock of knowledge developed in the past decades and has been intensively debated with the major competition authorities.¹¹

19. During the workshop, the discussion focused on quantitative methods as they are perceived to be a sort of 'first-best' approach to ex-post evaluation, if rigorously applied. Nevertheless, often competition agencies have very limited time, resources, and skills. In these cases— but also to complement more rigorous quantitative exercises— ex-post evaluations can also be based on qualitative techniques, such surveys or other methods.¹² Clearly qualitative studies cannot measure the effects of competition policy enforcement as precisely as quantitative ones do. Yet, there are important lessons that one can learn from them.

2.2.1 *Compulsory or not?*

20. Although all the delegates at the workshop generally agreed that ex-post evaluations are an important and useful tool, several different institutional settings are conceivable. The UK CMA and the Mexican COFECE stressed that ex-post evaluations are compulsory by law in their jurisdictions. They highlighted that this arrangement is particularly useful as it helps integrate ex-post evaluations in the process of policy making and enforcement. As is well-known, resources available to competition authorities are quite limited and there is competition between their use for regular enforcement activities and for retrospective studies. Therefore, institutionalizing ex-post assessments by making them compulsory might help establish a fairer division of resources.

¹¹ Starting in November 2015, this guide will be available on the OECD website

¹² See for instance the study conducted by the New Zealand Commerce Commission (Csorgo and Chitale, 2015) where it is qualitatively analysed whether those anticipated developments that were important to particular decisions did take place. If not, it is also attempted to understand the reasons why such developments did not take place as expected.

21. Some other delegates did not quite agree that this institutional arrangement is best, or even easily politically enforceable, if resources are very limited. They suggested that it might be more useful to use ex-post evaluations as a flexible instrument and internally motivate staff to conduct them by providing them with an interesting research opportunity that may also become an academic publication. The US experience is quite revealing. Retrospective studies are considered to be an important part of the staff's research activities, which have a dedicated budget so as to avoid competition with the resources dedicated to the enforcement activities. Hence, those staff members working on these projects might be particularly motivated by this additional reward, which could help produce high quality studies. Indeed, several of the ex-post evaluations carried out by the US FTC staff have been published in academic journals, some of which very highly ranked.¹³

2.2.2 *Resources: budget, data, people*

22. Ex- post evaluation studies that involve a quantitative assessment of the effects of a decision are time and resource intensive. All delegates strongly agreed on this issue. Generally, several people are involved in each evaluation exercise and this takes several months up to more than one year to be conducted. In some of the major authorities there exists a dedicated budget for evaluations or, more in general, for research activities, although this is often quite limited. For many authorities, such dedicated budget does not exist. This is still one major issue, as reported by several delegates and supported by our own experience. To ensure a minimum quality standard, both the quantity and quality of the involved staff, as well as the quality of the data has to be assured. If the necessary resources are not available, this might lead to suboptimal outcomes in terms of quality that can even be counter-productive. We will discuss this in more detail below, under quality checks. However, it is very important to stress that an adequate budget is a necessary condition for a successful ex-post evaluation exercise that provides precise and reliable quantitative results.

23. Ex-post evaluations are almost inevitably data intensive. Often data has to be purchased; hence it is crucial to allocate a budget for data acquisition already when planning the evaluation. Data quality has increasingly improved over the past years and there are several public or commercial data sources that can be used to perform econometric exercises. In some cases, the researchers can also collect or construct their own data.¹⁴ Optimally, competition authorities should think about the possibility of requiring the parties involved in the analysed case to provide data for the ex-post evaluation. While this would make the evaluation exercise much easier, it became clear during the discussion that it is unfortunately unlikely to happen. Indeed, it might prove to be legally challenging to impose this burden on the firms, as remedies in general need to be linked to the nature of the competition problem to be addressed and such a reporting obligation cannot easily be related to it. Nevertheless, especially in cases that involved specific interventions such as divestitures or behavioural remedies, it might be possible to think about such possibilities once ex-post evaluations have been more closely and institutionally integrated within the policy making process.

¹³ For instance Hosken et al. (2011), Ashenfelter and D. Hosken (2011), and Ashenfelter et al. (2009, 2013).

¹⁴ For instance in the study for the UK Competition Commission conducted by LEAR (Aguzzoni et al. 2011), the researchers collected product-specific data to match the data gathered by the CC during the investigation phase so to be able to perform a difference-in-difference analysis. Similarly in the study conducted for DG Economic and Financial Affairs, Buccirossi et al. (2009) collected data on the institutional and enforcement features of several competition authorities over a time span of 10 years to construct indexes of the quality of competition policy's regimes (see also Buccirossi et al., 2011 and Buccirossi et al., 2013).

24. The choice of the data to use in the analysis is crucial. The exact design of the evaluation framework, as well as the specifically adopted methodological tools, is heavily influenced by the structure and quality of the data. Low quality data often only allows low quality analysis or, at best, an analysis where the causal interpretation of the results might be quite challenging. Moreover, the data gathering as well as the data preparation are generally very resource and time intensive. Hence, once the (generally costly) data have been acquired and the (generally costly) preparation has been concluded, there is little room to go back and modify them. The general lesson is, therefore, that the first phase of an evaluation study, that includes the setting of the general framework and identification of available data should be particularly careful.

2.2.3 *Quality check: in house studies or outsourcing?*

25. According to the delegates' and our own experience, a well-organized research team is a crucial condition for a successful retrospective study. Generally, the team will be constituted of one or more senior economists with expertise in competition policy evaluation and the use of relevant empirical methods, as well as junior economists with strong empirical skills. The role of the former is to set up the evaluation framework, adapt the existing methodologies to the peculiarity of the analysed decisions, guide the junior staff along the empirical steps of the project (data generation and preparation as well as descriptive and econometric analysis) discuss, summarize and triangulate the results. The junior team members are generally more heavily involved with the data preparation and analysis, which is often quite time consuming.

26. Quite often economists inside competition agencies do not have the time to keep up to date with the most recent and advanced developments in econometrics and academics can help to fill this knowledge gap. Academics should be able to transfer the knowledge frontier, as well as the most advanced methodologies, from the academic work to the policy arena. They might be heavily involved as the main consultants who actively run the evaluation themselves or they might play the role of 'quality-checkers' or 'peer reviewers'. They might also be helpful as 'teachers' who help building up and updating the stock of knowledge inside the agency. It is very important that they get involved from the beginning as many — often irreversible — decisions about several methodological details need to be made during the evaluation and often (especially on data issue) at the beginning.

27. Peer-review 'quality-checks' are central to a successful ex-post evaluation. Often, competition authorities do not have enough in-house expertise to successfully engage in such activities.¹⁵ Furthermore, most authorities have limited staffs, who are necessarily allocated to the active ex-ante policy enforcement. In these cases, ex-post evaluation studies can (and should) be outsourced to academic experts. However, even in these cases, it is important to strike the right balance between in-house and outsourced activities. All authorities should build an in-house stock of knowledge to be able, if not to run retrospective analyses by themselves, but at least to understand and check the quality of the studies outsourced to external consultants. Therefore, one or more agency's officials should always be directly involved in order to internalise the lessons learned during the evaluation. These staffs should be able to follow, check, and probe the work of the external consultants and constantly ask detailed questions during different phases of the study to make sure that no corners were cut. Finally, it is prudent practice to let studies carried out by external experts be reviewed by other experts or, more in general, be discussed by the relevant community. This peer-review process is fundamental to assure the robustness and reliability of the presented results.

¹⁵ The US FTC is a major exception with more than 80 PhD economists in their staff. Most of the retrospective studies are done in-house and only sometimes academics have been involved. Other large authorities such as the UK CMA and DG Competition of the EU Commission follow a dual track with some studies outsourced to external experts and some done in-house (yet, often with a peer-review from an academic expert). Also smaller, but very active, authorities such as the Dutch ACM have followed this path.

28. Additionally, it is worthwhile to consider publishing the results of ex-post evaluations in academic journals subject to a serious peer-review process. This might help reinforce the belief that policy making is based on serious, scientific, evidence-based research and, hence help reaching advocacy goals. Yet, such peer-review process should also happen before the study is made public. As we already mentioned above, misleading results from poorly run studies can to some extent be almost detrimental.

3. The implementation of ex-post evaluations faces several methodological challenges

29. The ex-post evaluation of competition policy poses several methodological challenges. The Workshop discussed some cardinal methodological issues, and hereby we address them one-by-one, whilst acknowledging that an exhaustive methodological discussion is beyond the remits of the paper.

30. Reliable and statistically robust estimates come at the cost of increased methodological complexity and data demand. Although simple evaluations (such as comparisons of before and after prices or comparisons to yardstick competition) exist, their susceptibility to bias makes them less attractive to use. More sophisticated methods of course come with more assumptions, and the violation of these assumptions is likely to lead to biases in the estimates.

31. The exact design of the ex-post study depends on the specific circumstances of the case. Nevertheless, there is an ever growing body of literature applying these methods to evaluate competition policy, which can be used for guidance to develop some general rules of thumb.¹⁶ There is also a growing body of works that summarise the most important substantive points from the relevant literature. The forthcoming OECD Reference Guide provides a good synergy of the discussion of these methods. Another summary document, a forthcoming report by DG COMP of the European Commission (EU, 2015) on merger retrospectives in the EU gives a detailed overview of the methods used for the estimation of the impact of mergers. On cartels, the European Commission gives a detailed account of methods used for estimating the damages caused by cartels¹⁷.

32. Probably the most unambiguous message of the Workshop was that the specificities and details of the analysed case are important to determine what to exactly evaluate, how to do it in the specific case, as well as several details of the applied methodology. The Workshop primarily focused on reduced form causal analysis. Relatively less was said about simulations, because these are rarely used by the agencies themselves for the ex-post analysis of mergers.

33. Out of the three cases discussed, two were using difference-in-differences methods to estimate the price effects of mergers, and the third one used a mixture of statistical tests to establish the existence of a structural break and estimate the magnitude of the price and welfare impact of a cartel. Although not discussed in much detail, simulations based on a fully-fledged structural model were also briefly mentioned on several occasions; therefore, we provide a short reflection on this as well.

34. Although most of the discussion focused on estimating the price effect of the analysed behaviour, it is important to emphasise that understanding the non-price effect of mergers is an equally important task. For example developments in market structure (entry, exit) are relatively simple to observe and may be analysed using qualitative methods. Similarly, one of the most speculative elements of merger control is merging parties' efficiency claims, yet we have very few studies that examine whether anticipated cost savings took place post-merger. These may not require sophisticated econometric methods or highly granular data. The UK¹⁸ and New Zealand¹⁹ are good examples where the authority conducts ex-post

¹⁶ The forthcoming OECD Reference Guide references most relevant work.

¹⁷ European Commission (2013).

¹⁸ For example see Deloitte (2009).

studies of this sort. As the Workshop only superficially mentioned these methods we leave a more detailed discussion on them to future work.

3.1 *Difference-in-differences*

35. The most widely used method has been difference-in-differences (diff-in-diff), where the idea is to compare the change in prices in a market affected by the analysed practice (cartel, merger or abuse of dominance) with the change in prices for the same period in a different market unaffected by the analysed practice.

36. The reliability of the diff-in-diff method hinges on the correct design of the study, especially the choice of the counterfactual.²⁰ The success of this is of course dependent on data availability; the kind and quality of the available data strongly affects the research design. The discussion highlighted a few important points about how data can drive the research design.

37. First, there is the choice of the **counterfactual**. If there is sufficient local and/or regional variation in the data then a carefully selected set of local markets can be used as counterfactual. If, on the other hand, such data is not available and there is no local variation (e.g. if firms compete nationally rather than locally) then a different counterfactual, such as a competitor or a group of competitors, may have to be used.

38. The Workshop discussion rightly pointed out that depending on the counterfactual, some of the assumptions required for unbiased estimates might be more likely violated. For example, a key assumption for diff-in-diff studies is that there are *no spill-over effects* (i.e. the analysed event only affected the treatment group and not the counterfactual). This is unlikely to be the case if the counterfactual is the rivals' prices, which are likely to have been affected by the merger, or cartel, or abuse of dominance. The specific problem in this case is that the counterfactual price also absorbs some of the price effects, therefore any estimate from a diff-in-diff will be biased downwards.

39. When data is available for local markets and the data shows sufficient variation across these markets then choosing a better performing counterfactual is easier. Nevertheless, it is important to emphasise that this choice should not be arbitrary, but rely on formal methods. Propensity Score Matching (PSM) is a frequently used device for finding a counterfactual that is sufficiently similar to the treatment market (i.e. it converges to the ideal situation, where the counterfactual is the same as the treatment but for the analysed event).²¹

40. Other, less frequently used techniques exist for choosing a counterfactual that is not affected by spill-over effects. Probably the most elaborate of these was proposed by Choné and Linnemer (2013), who design an iterative technique using radii around spatial points to select a local market that was not affected by a merger. The motivation behind this technique is that even markets where only one merging firm is present might have absorbed some of the merger effects.

41. We cannot over-emphasise the importance of carefully designing the counterfactual. Even where formal methods for counterfactual selection are not possible, the merger study should always examine the implications of the given counterfactual choice. For example, if local markets cannot be used, and competitors are used as control then the study should examine in detail whether any assumptions of an unbiased diff-in-diff estimator have been violated and what this may imply about the estimate. Arbitrarily

¹⁹ Csorgo and Chitale (2015).

²⁰ In this paper we use counterfactual and control interchangeably.

²¹ For an example of the use of the PSM technique in the context of ex-post merger analysis, see: Aguzzoni et al (forthcoming) and OFT (2014).

selecting a group of competitors to use as control, without discussing its implications on the estimates, is not prudent practice.

42. Whilst – at least in the abstract – it may seem simple to see what the ideal counterfactual in a merger case is (i.e. the no-merger world), the same is not true for cartels. The motivation for the assessment of cartels is fundamentally different from merger retrospectives. A merger-decision is by definition ex-ante and, thus, the main goal of the assessment is to decide whether the authority got it right. Cartels on the other hand are ex-post decisions, their assessment assumes that the authority made the right decision and focuses purely on the magnitude of the negative effects that were avoided through successful intervention.

43. The Workshop briefly touched upon two main types of counterfactuals that are used in cartel overcharge estimates. One is based on a benchmark market (products not included in the cartel, or jurisdictions without the cartel, or pre-cartel or post-cartel prices), and the other is based on a more involved cost-plus approach, where the product costs and a reasonable merging are assumed to be the non-collusive price (the rigour of the estimation is probably very high, as cartel overcharge estimates are frequently used in front of Courts in private damage claims). In practice of course, the choice of counterfactual in cartel cases is also driven by data availability. The problem is that most of the counterfactuals used in cartel cases are likely to suffer from the same problem: they might not closely reflect the ‘but for’ world. To find a well-performing counterfactual, one would need to know what would have happened in the absence of the cartel. Would the market (and prices) be competitive or would there still be some level of (tacit) collusion? In practice however either pre-, or post-cartel prices are used without answering these questions. There is a growing size of literature on the dynamics of markets after cartel breakdown that find that the markets in the post-cartel world are often highly collusive or are characterised by intensive merger activity to make it more collusive.²² In this case any overcharge estimate that is based on the assumption that without the cartel the price would be competitive, would be upward biased.

44. Another key methodological point that the Workshop discussed in detail relates to the **dimensions of the panel data**. The data used for difference-in-differences analysis has two key dimensions: the number of individuals observed (e.g. markets, firms, or products), and the number of time periods. Datasets vary in their disaggregation, which determines the exact dimensions of the data, which in turn can be used to inform the researcher on the adequate methods (for example how to model the error terms). Panel data with a large number of time periods (time-series dominant) can be prone to all the issues that one would encounter in time-series analysis (e.g. non-stationarity, autocorrelation). Such time-series issues are less likely to be a problem in data with large number of individuals and a small number of time periods (cross-section dominated).

45. One of the points raised at the workshop was the importance of dealing with autocorrelation (where the error terms at different time periods are correlated). If there is autocorrelation in the data and the research does not deal with it, then the estimated model is likely to return spurious results. This is due to the problem that if there is positive correlation between the dependent variable (usually price) at different time periods, then the standard errors estimated using a simple OLS regression will be biased downwards. This could lead to statistically significant findings even where they do not exist.²³

46. Ideally, if the time-series aspect of the data is correctly handled, it is good to have large frequency data that spans sufficiently long before and after the event. However, in some circumstances the data only allows a distinction between before and after prices for a large number of individuals (cross-sections). Although this situation would eliminate concerns about non-stationarity, or serial correlation, it would of

²² See for example Davies, Ormosi and Graffenberger (2015).

²³ Bertrand et al (2004) give a detailed account of how autocorrelation can affect difference-in-differences estimates.

course limit the richness of the analysis. Moreover, a simple before and after breakdown of data (i.e. respectively aggregating all data before and after the analysed event) may be problematic, because it does not allow testing of the parallel trend assumption between the treatment market/product and the counterfactual. The assumption of parallel trends is one of the conditions of unbiased estimates. If the treatment and the counterfactual time-series of prices follow non-parallel patterns then the estimated effect of an intervention will also contain some other underlying differences between the treatment and counterfactual. With many time periods we can test if the treatment and the counterfactual groups follow a parallel pattern before the analysed event, to prove that any deviation from parallel is due to the analysed event.

47. Another central assumption to the diff-in-diff method is that the analysed event is **exogenous**, i.e. the incidence of the merger, cartel, or abuse of dominance is not correlated with the error term in the model. Endogeneity can be due to unobserved factors that were driving the decision to merge/collude/abuse. For example, if the price effects of a merger is estimated and the researcher has no information on costs, but the merger is driven by the parties' willingness to reduce their costs, then the estimated price effect will suffer from endogeneity bias. Endogeneity could also be due to reversed causality, i.e. where a change in the price is what is driving the analysed event not the other way around.

48. In practice it is not trivial how to deal with such endogeneity, and many of the previous studies simply assume that the analysed event is exogenous. This is a possibility in many cases, for example, when analysing the local (e.g. country-specific), and product-specific effects of a largely diversified global merger, the decision to merge at the local level is likely to be exogenous (based on a global decision at the headquarters of the firm). Nevertheless, when the situation is not so clear cut, the most prudent practice is to try and find out what triggered the analysed event. Information on this can be collected through a qualitative analysis (e.g. a questionnaire administered to market participants or market experts). In merger cases there might be signs suggesting that the merger was proposed by the parties to reduce their costs. With this information in hand it becomes even more important to try to acquire costs data and include it in the model to reduce the chances of biased results.

49. Having as many controls as possible is very important for **identification**. It is important however that the right interpretation is given to controls. For example many studies control for costs in order to (at least partially) deal with endogeneity. But we can often see that by costs many of these papers include the price of some raw material. This is good practice, but it is important to bear in mind that the cost-minimising considerations (e.g. rationalising production) that may drive many mergers might be unrelated to the price of these input factors, therefore these controls may achieve very little in eliminating endogeneity. Another issue that we have seen is that many mergers were analysed together, which limits the appropriate identification of effects.

50. The Workshop also highlighted the use of fixed-effect controls, for example dummy variables for each local market, or each firms, or each diversified product. This helps control for some of the unobserved heterogeneity (at least the time-invariant part) across local markets, or firms, or products. For example local markets where only one of the merging firms is present is often used as counterfactual. However it is possible that there are unobserved characteristics of local markets (population demographics, local markets structure, regional tastes) that determine where each firm is present. Fixed effect dummies for these local areas can control for some of these unobserved characteristics.

51. Although the general consensus is that diff-in-diff methods are a simple, versatile, and reliable method for estimating the impact of competition policy related events, the workshop highlighted two of its **limitations**. Firstly, diff-in-diff is well designed for estimating changes in a simple outcome variable (most frequently price) in a reduced form context. However it is inadequate for welfare analysis. Assume that in two cases the diff-in-diff estimates show a price increase of 1% in each case. Based simply on a diff-in-diff study, one would conclude that the two mergers had a similar effect. However, it is possible that the two

cases have completely different welfare implications. For example, in a market where goods are homogenous, demand is inelastic and customers buy frequently (e.g. petrol) a 1% price increase may have much more serious welfare effects than in a market with differentiated products, more elastic demand, or infrequent purchases (say television sets). One would need to go beyond these diff-in-diff estimates – potentially using a structural model – to explore these differences in the two cases.

52. Second, if the purpose of the ex post study is to evaluate the quality of the competition authority's decision then diff-in-diff may not be the best for analysing all possible scenarios. For example in merger retrospectives, in many jurisdictions the ability of diff-in-diff to identify Type I errors (too much intervention) is limited. The intuition is simple: diff-in-diff estimates in merger evaluations show the difference in prices between the factual world (world with the remedied or unconditionally approved merger) and the counterfactual world (no merger). If the estimates imply that prices increased post-merger, then there was possibly a Type II error (insufficient intervention) in the decision. However if prices do not increase, and a remedy was required by the authority, then we have two possibilities: (1) the remedy was the right decision, and (2) the intervention was unnecessary, and prices would have remained the same even without the remedy (hence there was possibly a Type I error). Because in a diff-in-diff model the counterfactual is always the no-merger world, it often does not allow the researcher to compare the factual price (price with a remedy) with a counterfactual world of merger with no remedy.²⁴

53. There are of course important exceptions, where diff-in-diff studies can still inform us about Type I errors. For example in the US it is possible that the DOJ or FTC decides to block a merger, and the Court overrules this decision, and the merger goes ahead. In this case the diff-in-diff study compares prices between the merger and the no-merger world. If prices do not increase then it can be concluded that blocking the merger would have meant an unnecessarily strong intervention (Type I error). Another example is the UK system, where mergers do not need pre-notification. If a merger is consummated and there is sufficient time before a remedy is imposed, a diff-in-diff could identify if the remedy was overly strong by comparing prices without the merger, prices under the factual world of the remedied merger and prices during the period when no remedy was implemented. If the no-remedy world does not show a price increase, then it would suggest that a Type I error (unnecessary remedy) was made. Of course any such conclusion is strongly contingent on the exact details of the case. The bottom-line is: the number and types of counterfactuals that can be created from available data determines how far our interpretations of the potential agency errors can go.

3.2 *Simulations*

54. Although the Workshop focused primarily on difference-in-difference methods, other methodologies were also mentioned and, among these, simulations were discussed on various points. Although the main application of simulations is mergers, it must be noted that there is no reason why these methods cannot be used in other policy areas.²⁵ Here we do not wish to give a detailed account of simulations. Instead we focus on the few issues brought up at the Workshop, leaving the detailed introduction of structural modelling and simulations to specialist articles and textbooks.

55. There is no general consensus about the full merits of using simulations in ex-post studies. One of the main problems is that depending on the assumptions, completely different results can be derived, and whereas in diff-in-diff it is possible to empirically test the viability of the assumptions, the same is not always true in simulations, where for example one may get very different results depending on assumptions about demand, or the type of competition in the market, or the model used.

²⁴ More details on this discussion are given in European Commission (2015).

²⁵ For example structural model based simulations are increasingly being used in cartel damage estimations.

56. The academic community is also divided on the relative merits of diff-in-diff and simulations.²⁶ Without wanting to provide judgement in this debate, we take a more dovish stance and argue that both diff-in-diff and simulation methods can be useful in some cases and have weaknesses in others. Their respective usefulness depends on the research question and the available data.

57. Simulations are frequently used in merger analysis during the investigation to help predict the outcome of the merger. Here we only focus on simulations that are done after the analysed event. Based on the type of data used, we can distinguish between three ways of ex-post simulations.

- *Simulations using ex-ante information:* Whilst, strictly speaking, these do not provide ex-post estimates, they still have very important value in assessing/evaluating the decision of the authority. For example in mergers, during the investigation a fully-fledged simulation provides probably the best prediction of merger outcomes. Nevertheless, it is rare that the authority has time for these fairly complex exercises. Doing an ex-post simulation using ex-ante data can tell us what the most-informed decision of the authority would have been, from which we can identify how far away the actual decision fell from this most-informed prediction.
- *Simulations using ex-post data:* naturally, to derive truly ex-post estimates using simulations, we have to rely on ex-post data. The purpose of this exercise is the same as in ex-ante simulations, the effect of an event is simulated but in this case the structural parameters are estimated using ex-post data, and the pre-event situation is simulated using these parameters but accounting for any change that happened post-event (for example, in the case of mergers, this would account for any post-merger cost savings, or how entry and exit affected the market).
- *Simulations using ex ante and ex post data:* Although this is the most laborious of our three categories – as it requires estimating the pre and post equilibria separately –, it has the advantage that it allows testing a key assumption used in simulations: whether the pre and post structural parameters remain unchanged.

58. Although simulations relying on a full structural modelling and estimation of demand parameters can be quite involved, agencies on the ‘knowledge-frontier’ together with academics can play a crucial role to help develop and refine these methodologies.

59. In an ideal world of course one would be able to estimate effects using both diff-in-diff and simulation methods. Typically, there is a large overlap in the data required, and especially where a simulation is being conducted, it should be relatively simple to use the same data to also run simple diff-in-diff models.²⁷ Using the diff-in-diff estimate can help guide the research through the simulations and eliminate the results that are due to ‘wrong’ assumptions.²⁸ However, such an exercise would be an onerous task for the authorities.

60. Another possibility is to combine diff-in-diff studies with a more qualitative approach. Diff-in-diff could identify the cases where post-merger prices increased, and the authority could look at these cases in more detail to establish what – if anything – went wrong with the decision. For example it can assess the

²⁶ For probably the most prominently exposed debate on this see: Angrist and Pischke (2010), and Nevo and Whinston (2010).

²⁷ It is true in general that due to the improvements of diff-in-diff studies, the quality and the detail of the data used in either methods are becoming more and more similar.

²⁸ For an elaborate study of this sort see Björnerstedt, and Verboven (2012).

precision of its predictions on post-merger entry and exit and examine whether these factors (or a lack of them) contributed to the unforeseen price increase.²⁹

3.3 *Acquiring data for the estimation*

61. Data availability inevitably determines which industries are analysed. A comprehensive study on US merger retrospective by John Kwoka (2014) shows that most US merger retrospectives are concentrated around a few industries (e.g. hospitals, transportation, petrol retail), where detailed data is available either publically or as proprietary data.

62. When selecting cases to evaluate, one important issue is the complexity of the given market and its interaction with data availability. The focus on end-consumer and retail markets makes the analysis cleaner and easier. However there are several complex but very interesting markets (e.g., health sector, energy sector, intermediate markets) for which it might be interesting to make evaluations but which involve much more assumptions and require additional thinking on how to apply known methodologies in a sensible way.

63. It might be advisable to think about the evaluation of a case (and about data collection) already during the investigation of the case itself. It might help to start collecting some data during the investigation and (possibly) impose it on the involved firms to provide data in the future (behavioural remedy/monitoring).

3.3.1 *How to interpret the findings*

64. One reason why some authorities might be reluctant to conduct ex-post evaluations of their cases is because of their concern: what if it shows that they made an error. It was emphasised at the Workshop that the finding of an increase in, say, post-merger prices does not necessarily imply that the authority made an error in its decision. It is possible that the authority's decision was based on faulty facts, for example on surveys, where rivals wrongly anticipated increased entry post-merger. If the authority based its decision on such surveys that reflected reasonable predictions about post-merger events, then it would be too harsh to call it the authority's error if these predictions turn out to be wrong. Similarly, there is always some random variation in post-merger prices. Merger control is an ex-ante exercise, and as such it is unable to predict the post-merger world with full precision. Imagine for example that having considered all information about a merger, the competition authority predicts that with 50% probability prices will not increase post-merger, with 25% probability post-merger price will be -1, and with 25% probability post-merger prices will be 1. The rational decision by the authority would be to approve this merger as the expected price-change is non-positive (in this case zero). But even in this case (having accounted for all possible information) there is a 25% chance that prices will increase due to random variation. An observed post-merger price increase may indicate that prices went up purely due to chance. In this it would be unreasonable to conclude that the authority made an error. Finally, many authorities acknowledge that small price-increase might often follow mergers but they are outweighed by other, non-price factors (for example quality improvements).³⁰

²⁹ Similar qualitative post-merger studies are regularly conducted for example by the UK and New Zealand. Example of DiD coupled with survey/interviews are Aguzzoni et al (2014), Aguzzoni et al (forthcoming) and Buccirossi et al. (2006).

³⁰ For a more detailed account of the interpretation of the findings of merger retrospectives see Werden (2013).

3.4 Concluding on methodology

65. Probably the most important advice is that ex-post studies should be upfront about the assumptions used and discuss possibilities of violating these assumptions and the potential consequences on the credibility of results in light of these violations. The same is true for the correct identification of effects and establishing causality. More specifically, the following key points could be highlighted:

- Careful study design with a particular focus on the choice of counterfactual: the perfect counterfactual should share all attributes of the treatment group, but for the fact of the treatment itself.
- The study should acknowledge the limitations of the chosen counterfactual. For example, it might be that different local markets were used as counterfactual but the researcher recognised that prices on these markets were also affected by the merger/cartel. This would produce price-impact estimates that are downward biased (because the spill-over nips away some of the true price effects). Recognising this, the study can still conclude that the estimated price effect is a lower bound of the real price effect.
- Identifying the causes of the analysed events and controlling for all factors that might have triggered the analysed event can help avoid endogeneity issues.
- Simulations could help estimate the welfare effect of the analysed events but they are much more demanding on data and resources.
- In the analysis of mergers, if estimates show that post-merger prices increase, it does not necessarily mean that the authority made an error in the decision.
- Any ex post-study should carefully check the robustness of estimates to possible variations of the model/method.
- Requesting a peer-review (typically by an academic) can improve the credibility of findings.

4. The way ahead

66. Ex-post studies on the decisions of competition authorities serve several important goals. Probably the most important of these is to feed-back into the decision-making process and, thus, improve future decisions and their effectiveness. Retrospective studies have begun to be considered an integral part of the enforcement of competition policy. They face several methodological and analytical challenges, which have been partially addressed over the last decade by a growing body of academic and policy works. The previous discussion highlighted that a robust stock of knowledge has been built and several recent summary studies have delineated what can be considered to be best practices for ex-post evaluations. Nevertheless, there is still a great potential for improvement.

67. First, most of the existing studies focus on merger policy. Yet, other areas of competition policy enforcement are equally important and have not been systematically evaluated so far (e.g., Ormosi and Davies, 2012). While many methods and considerations that are well understood to work and exist for the evaluation of mergers can be simply adopted to assess the effectiveness of other antitrust decisions, it is important to think more carefully about the differences and peculiarities of such assessments. This is particularly true because merger control is the only area of competition policy that is enforced ex-ante, while cartels, abuses, and other infringements are ex-post interventions. This crucial difference poses some issues on how to interpret retrospective studies in these other policy areas. Specifically, it is not quite clear

whether it is really possible to disentangle the impact of the firms' misconduct from the impact of the authority's decision. More work on this issue is definitely needed.

68. Second, most of the existing studies look at the price effect of single competition policy decisions. There are many good reasons to focus on prices, as they are an important— and often (more) easily available — indicator of consumer welfare. Yet, other outcomes are also important and hence research and future evaluations should go beyond just looking at simple price effects. This includes several aspects among which:

- Non-price effects such as the impact on quality, variety, innovation, entry/exit etc.
- The identification of the broader effects of policy enforcement such as its overall welfare effects, its effect on productivity, its deterrence effects, the spill-over effects that some key decisions might have on different markets.
- A clearer identification of type I errors—too much intervention— and type II errors—insufficient intervention.

69. For all these issues, not only more appropriate empirical tools, but also more general and robust theoretical frameworks are needed. For instance, it is not clear whether an increase in products' variety or the entry and exit of new products and firms should be seen as a welfare improvement or not. Another major challenge is the identification of true causal relationships when finding an appropriate counterfactual is particularly difficult. This is particularly important when looking at the broader effects of competition policy enforcement. To get the broader picture of the effects of competition policy enforcement, researchers have to adopt more general, macro-oriented frameworks and they need to use more aggregate data, which pose very different identification issues than when using the most common, micro-founded, case-specific, difference-in-difference methodologies (e.g. Buccirossi et al., 2013). One of the main causes lies in the fact that in these cases the researchers need to more carefully think about the indirect effects of competition policy. A particular decision might directly affect some markets, which in turn are connected along the value chain to several other downstream markets. To clearly identify the effects of competition policy enforcement on final consumers in these cases might be very challenging. Finally, some more work should be invested in more cleanly identifying the extent of potential errors. While to some extent, existing methodologies allow identifying type II errors, we know much less about type I errors as we discussed before. For instance taking the example of a remedied merger, an estimated price increase following the merger would imply a type II error. Yet, if prices do not increase we do not know whether there was a type I error (i.e. if the intervention was necessary at all) because we cannot easily estimate a counterfactual scenario with the merger but no remedy.

70. Third, there appear to be still too little work on the clear identification of the effect of specific remedies, which are currently the most commonly used policy tool. Even in the more specific case studies, it might be almost impossible to separately identify the effect of the firms' conduct, the Commission decision, and the specific remedy. Institutional details might be helpful, as clearly discussed by the delegates of the UK CMA in their presentation. The different timing as well as the geographical variation of the enforcement of different remedies might potentially be exploited to more clearly assess the effect of specific remedies. However, this might be only possible in specific cases and markets and under particular institutional conditions. To answer this question, therefore, the use of structural econometric methods and simulations might be a way to go. Indeed, these tools allow generating simulated counterfactuals which otherwise would not have been observed. However as the US FTC stressed, the complexity and resource-intensity of this methodology are still such that it might take some time before they become standard tools for ex-post evaluation.

71. Fourth, while there is quite some work on the evaluation of competition agencies' decisions, there is still very little analysis of the role of judges and the evaluation of their decisions. This is unfortunate given that courts are very much involved— either directly or indirectly— in the enforcement of competition policies. A more careful analysis of both the authority's and courts' decisions and how they are related might also be helpful to better identify possible mistakes. For instance, in cases where the competition authority blocks a merger, but courts overturn this decision then the researcher might try to test whether the authority potentially made a type I error.

72. Finally, some discussion is still needed on what to do with possibly negative results from evaluations. On one hand, they provide an even more important feedback to the policy making process than positive evaluations. A pragmatic view would be just to accept that some mistakes can be made and learning from them can reduce future errors. On the other hand, it seems imperative to understand the legal consequences of such cases. Can firms sue the authority for making mistakes? Can the evidence put forward by ex-post evaluations be used in Court for example in ongoing compensation claims? Moreover, some work need to be done on whether the risk of negative evaluations might have a perverse incentive on the selection of cases that get evaluated.

REFERENCES

- Aguzzoni, L., E. Argentesi, P. Buccirossi, L. Ciari, T. Duso, M. Tognoni, and C. Vitale, 2011, "The Ex Post Evaluation of Two Merger Decisions," a report prepared the UK Competition Commission, available at http://www.competition-commission.org.uk/assets/competitioncommission/docs/2011/11_09_20_ex_post_evaluation_of_two_merger_decisions.pdf.
- Aguzzoni, L., E. Argentesi, P. Buccirossi, L. Ciari, T. Duso, M. Tognoni, and C. Vitale, 2014, "A Retrospective Merger Analysis in the UK Videogames Market," *Journal of Competition Law and Economics*, 2014, 10, 4, 933–958
- Aguzzoni, L., E. Argentesi, L. Ciari, T. Duso, and M. Tognoni, 2015, "Ex-post Merger Evaluation in the UK Retail Market for Books," *Journal of Industrial Economics*, forthcoming.
- Ashenfelter O. and D. Hosken, 2011, "The Effect of Mergers on Consumer Prices: Evidence on Five Mergers on the Enforcement Margin," *Journal of Law and Economics*, 53, 417–466.
- Ashenfelter O., D. Hosken and M. Weinberg, 2009, "Generating Evidence to Guide Merger Enforcement", *Competition Policy International*, Volume 5, No. 1.
- Ashenfelter O., D. Hosken and M. Weinberg, 2013, "The Price Effects of a Large Merger of Manufacturers: A Case Study of Maytag-Whirlpool", *American Economic Journal: Economic Policy*, 5 (1): 239-61.
- Bertrand, M., E. Duflo and S. Mullainathan, 2004, "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119 (1), 249-75.
- Björnerstedt J. and F. Verboven, 2012, "Does Merger Simulation Work? Lessons from a Case in the Swedish Analgesics Market," *CEPR Discussion Paper No. DP9027*.
- Buccirossi, P., L. Ciari, T. Duso, S.O. Fridolfsson, G. Spagnolo, and C. Vitale, 2006, "Ex-post review of merger control decisions," report commissioned by DG Competition of the European Commission.
- Buccirossi, P., L. Ciari, T. Duso, G. Spagnolo, G., and C. Vitale, C., 2009, "Development and Application of a Methodology for Evaluating the Effectiveness of Competition Policy," report commissioned by DG Economic and Financial Affairs of the European Commission.
- Buccirossi, P., L. Ciari, T. Duso, G. Spagnolo, G., and C. Vitale, C., 2011, "Measuring the Deterrence Effect of Competition Policy: The Competition Policy Indexes," *Journal of Competition Law and Economics*, 7, 165-204.
- Buccirossi, P., L. Ciari, T. Duso, G. Spagnolo, G., and C. Vitale, C., 2013, "Competition Policy and Productivity Growth: An Empirical Assessment," *The Review of Economics and Statistics*, 2013, 95, 4, 1324-1336.
- Csorgo Lilla and Harshal Chitale , 2015, "Targeted Ex Post Evaluations in a Data Poor World," report commission by the New Zealand Commerce Commission, available at <http://www.comcom.govt.nz/dmsdocument/13620>.

- Davies, S. and Ormosi, P. (2012) A comparative assessment of methodologies used to evaluate competition policy in *Journal of Competition Law and Economics* 8, 769-803.
- Davies, S., Ormosi P., Graffenberger, M. (2015) Mergers after cartels: how markets react to cartel breakdown, *Journal of Law and Economics*, 58 (Nov) Forthcoming
- Deloitte, 2009, "Review of merger decisions under the Enterprise Act 2002," a report prepared for Competition Commission.
- Duso T., 2012, "A Decade of Ex-post Merger Policy Evaluations: A Progress Report," in Dan Sjöblom Ed., *More Pros and Cons of Merger Control*, Stockholm: Swedish Competition Authority, 125-187.
- European Commission (2013) Quantifying harm in actions for damages based on breaches of Article 101 or 102 of the treaty on the functioning of the European union, Accompanying the communication from the Commission on quantifying harm in actions for damages based on breaches of Article 101 or 102 of the Treaty on the Functioning of the European Union [C(2013) 3440].
- European Commission (2015) A review of merger decisions in the EU: What can we learn for ex-post evaluations? *Forthcoming*
- Farrell J., P.A. Pautler, and M.G. Vita, 2009, "Economics at the FTC: Retrospective Merger Analysis with a Focus on Hospitals", *Review of Industrial Organization*, 35 (4): 369-385.
- Friberg R. and A. Romahn, 2012, "Ex-Post Merger Review and Divestitures," mimeo, Stockholm School of Economics.
- Hastings, J. S., 2004, "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California," *American Economic Review*, 94, 1, 317-328.
- Hastings, J. S., and R. J. Gilbert, 2005, "Market Power, Vertical Integration and the Wholesale Price of Gasoline," *The Journal of Industrial Economics* 53 (4), 469-492.
- Hosken, Daniel, Louis Silvia, and Christopher Taylor, 2011, "Does Concentration Matter? Measurement of Petroleum Merger Price Effects," *American Economic Review*, 101, 3, 45 –50.
- Hosken D., L.M. Olson, and L.K. Smith, 2012, "Do Retail Mergers Affect Competition? Evidence from Grocery Retailing", Federal Trade Commission, Bureau of Economics, Working Paper No. 31.
- Ilzkovitz F. and A. Dierx, 2015, "Ex-post economic evaluation of competition policy enforcement: A review of the literature," available at http://ec.europa.eu/competition/publications/reports/expost_evaluation_competition_policy_en.pdf.
- Kwoka J., 2014, *Mergers, Merger Control, and Remedies: A Retrospective Analysis of U.S. Policy*, MIT Press.
- Nevo, A. and M. D. Whinston (2010). "Taking the dogma out of econometrics: Structural modeling and credible inference." *The Journal of economic perspectives*: 69-81.
- OFT (2014) Shell - Rontec: An evaluation of the OFT's conditional clearance of the merger.
- Peters, C., 2006, "Evaluating the performance of merger simulation: Evidence from the U.S. airline industry," *Journal of Law and Economics*, 49, 627-649.

Weinberg M., 2008, "The Price Effects of Horizontal Mergers", *Journal of Competition Law and Economics*, 4 (2): 433-447.

Werden, G. J., 2013, "Inconvenient Truths and Constructive Suggestions on Merger Retrospective Studies." Available at SSRN 2287856.