Appraisal of:

"Feasibility Study on The Microeconomic Impact of Enforcement of Competition Policies on Innovation"

by Anna Rita Bennato, Steve Davies, Franco Mariuzzo, and Peter Ormosi

Tomaso Duso*

August 15, 2017

*Deutsches Institut für Wirtschaftsforschung (DIW Berlin), DICE Düsseldorf, Berlin Centre for Consumer Policies (BCCP), and CESifo, Mohrenstr. 58, 10117 Berlin, Germany. E-mail: tduso@diw.de.
1 Overview

The relationship between competition and innovation has been of central interest for economists for a very long time. Although a substantial literature on the matter exists, very few retrospective studies analyze the impact of competition policy enforcement, such as merger control decisions, on innovation. In its effort to expand the still limited body of knowledge on the effectiveness of competition policy tools on different market outcomes, rather than just prices, the Directorate General for Competition (DG COMP) commissioned a team of academics led by Peter Ormosi of the Centre for Competition Policy, University of East Anglia, to develop a feasibility study to assess how specific decisions by antitrust authorities – specifically DG COMP itself – impacted innovation.

The report "Feasibility Study on The Microeconomic Impact of Enforcement of Competition Policies on Innovation" first provides a systematic overview of the existing literature on the nexus between competition, competition policy, and innovation. It then proposes a methodological framework that can be used to evaluate specific cases. This framework is then applied to a pilot study aiming at assessing two specific decisions made by the European Commission in the hard drive (HDD) market, namely the unconditional clearance of the Seagate/Samsung merger at the end of 2011 and the conditional approval of Western Digital’s (WD) acquisition of Hitachi Global Storage Technologies in 2012. In the latter case, the Commission accepted the commitments of the merging parties to divest the entire 3.5” HDD production to Toshiba.

I was asked to provide an external appraisal of this report. I received a draft of the report at the end of March 2017 and provided my comments on the draft at a workshop held at DG COMP one month later. This is my appraisal of the final report and it has four further sections. In section 2, I present a general introduction of the study and specifically discuss the issue of how to measure innovation. In section three, I discuss the analysis of effect of the merger on R&D, while in section four I discuss the analysis of the mergers’ effect on patents and product characteristics. Finally, in section five, I conclude with some general remarks and lessons learned from the study.

2 General Remarks & Measurement Challenges

First off, I think it is important to stress from the start that this report is a ‘feasibility’ study. Hence, its ultimate goal is to verify whether it is viable to assess the effect of a competition policy decision on innovation, what are the main methodological challenges, and which kind of data is necessary. The application to two specific merger decisions is surely an additional important piece of information but should not be seen as the core of the report.

Keeping this in mind, I think that this is a very important study that helps expand the knowledge frontier on the ex post assessment of competition policy interventions. Compared to
more ‘traditional’ ex post evaluations, which generally focus on studying the effect of (mostly) mergers and merger control decisions on prices (e.g. Ashenfelter et al., 2014), when looking at the impact of merger on innovation, researchers face two major challenges. First, the definition of the outcome variable is even more difficult, as there is no perfect indicator for innovation. Second given that in most markets innovation activities are quite global, it is even more difficult than for other studies to define an appropriate counterfactual. Indeed, many retrospective merger studies use geographic variation in the outcomes—either at the local/regional or (at least) at the country level—as a source of identification. This is almost impossible when analyzing innovation markets, which are generally more global than product markets.

To overcome some of these issues, the study team proposes a rich and pragmatic approach. With the aim of producing robust evidence, they propose using multiple outcomes, each capturing different notions of innovation: from innovation as an input (R&D spending), to innovation as an output (patents), and to innovation as an outcome (product characteristics/positioning). Moreover, for (almost) each outcome they propose different counterfactuals as well as different methodologies (difference-in-differences and synthetic control method) to identify the causal effect of the mergers and merger control decisions. Finally, they also propose using placebo tests to check the consistency of their results. In all of these steps, the authors try to build on different strands of literature and adapt them to fit the specific issues faced in the particular case study.

The richness of this approach has to be evaluated very positively, as the convergence (but even the divergence) of results might help better understanding the effect of the merger and the Commission’s decision on the evolution of the market analyzed in the specific case. However, its application to a specific case study has also clear consequences. First, it essentially leads the authors to conduct three different case studies based on three different data-sets and identification strategies. Second, because of the budget and time constraints, it leads to incompletely developed case studies, thus leaving the reader a bit puzzled and wanting for more. Both of these points have further consequences on the feasibility of following studies and the design of such projects, something that I will return to in the concluding remarks.

While the case study picked by DG COMP and the study team has many advantages that I will thoroughly discuss throughout this appraisal, the existence of some concurring events with the mergers and the EU decision makes the causal assessment of the EU-specific event more difficult to identify. Specifically, both analyzed mergers were also scrutinized by MOFCOM, the Chinese competition authority. Although MOFCOM conditionally approved both mergers, it imposed some additional remedies in the second merger, which delayed the full integration of Western Digital and Hitachi until the lifting of these additional obligations in 2015. This implies that due to a lack of data and time (R&D expenditures were available only until the first quarter of 2016), some effects could simply not be identified as the second merger was not fully
consummated until relatively late. This specific challenge highlights the importance of carefully choosing the case study, as it can have serious implications for identification.

To conclude these general remarks, I think one should appreciate the effort of the authors to be upfront on what these kind of studies can and cannot achieve. While in some cases smart methodological designs or the ingenious use of specific characteristics of the analyzed market can help to improve, if not fully achieve, a clean identification of the causal nexus between the firms’ conduct (e.g. the merger), the antitrust authority’s decision, and the market outcome, in some other cases this is simply not possible. Nonetheless, retrospective studies can be very helpful to better understand or monitor the evolution of the industry after the merger, which is a valuable exercise.

3 The Effect of the Mergers on R&D spending

3.1 Outcome variable

The first outcome variable used in the study to proxy for innovation is based on R&D expenditures. This is a standard measure of innovation as an input that is frequently used in the literature. Therefore, and notwithstanding the many criticisms that this measure has attracted, it seems to be a reasonable starting point for analyzing the effect of mergers on innovation. The main results for this part of the analysis are that while Seagate’s R&D intensity significantly increased after the merger, WD’s R&D intensity was not significantly affected. Moreover, the findings show that Toshiba’s R&D intensity significantly decreased after the decision of the EU Commission in 2012.

The authors are very upfront in their discussion of the drawbacks of this measure. First, the main specific problem for using this measure in ex post evaluations it that it is too aggregated. Most times, in fact, measures of R&D expenditures are not product-specific and gathered at the overall firm level. If the merger – or any other antitrust infringement under analysis– only affects specific product markets and the involved firms are highly diversified, it might be difficult to measure the effect of the merger simply because R&D expenditures are recorded for the whole firm and cannot be attributed to the relevant markets. This point is not new and not solely specific to innovation\(^1\), and has important consequences for the feasibility of such studies. A similar criticism relates to the geographic dimension of innovation. If firms are operating globally, then it is also likely that innovation activities are carried out globally. Therefore, it might be very difficult to use geographic variation to identify a relevant control group based on geographical differences. The authors carefully discuss these points and mention how they can threaten the ability to cleanly identify the causal effect of the merger.

---

\(^1\)Similar criticism is raised for using stock market event studies to assess merger’s and merger control’s effects (see Duso et al., 2015)
A second issue related to measuring the effect of a merger on R&D expenses is that, after the merger, R&D data measure the overall expenditures for the new entity, which, in most cases, represent the combination of the two merging parties. The authors propose two steps for solving this problem and suggest (1) to remove the period of the merger from the analysis; and (2) to look at growth in R&D intensity rather than absolute figures. While these seem to be adequate steps, my suggestion of summing up R&D expenditures for merging parties also pre-merger (e.g. Gugler and Siebert, 2007) was not followed or discussed.

To account for some of the issues discussed above, the precise measure used in the regressions is \textit{R&D intensity growth}. While I think this is a reasonable choice that is well motivated by the authors, I still think that it might have been useful to not only discuss but also use other measures such as the level of R&D expenditures as well as R&D intensity itself to better understand what drives the empirical results. Clearly, this would make this sub-analysis even longer and more demanding.

3.2 Choice of the Counterfactual

The key identification issue in any evaluation study using a difference-in-differences framework is the definition of the counterfactual. I think that this is possibly the most critical issue for the analysis of the impact of competition policy enforcement on innovation that is stressed in several parts of the report. The authors carefully review this point and are upfront about what one can effectively do. The pragmatic approach in the R&D analysis is to try several different counterfactuals, each with pros and cons, trying to identify a consistent story that explains the different results.

The authors propose three different counterfactuals. They start by making clear that using rival firms as control is not viable in this setting, as there are essentially no rivals that have not been actively involved in one of the mergers or remedies. Hence, following other \textit{ex post} evaluation studies (e.g. Ashenfelter et al. 2013), they first propose using other products (SSD and flash device) that are similar to those affected by the merger (HDD) to be considered a good counterfactual. While I do like this idea in general, the main issue in the context of this pilot case study is that the different product markets are very much related. Although the Commission and the parties agreed that HDD is a relevant product market, per se, there might still be demand and supply side spillover effects across these markets and technologies that makes this identification challenging. On the other hand, in the spirit of the discussion above, although this issue might impact the identification strategy, the results may facilitate the understanding of the evolution of the industry post-merger.

The second proposed counterfactual consist of all other IT firms as "\textit{the demand for high-tech products is likely to be driven by the same underlying economic conditions}". I am more skeptical about the validity of this counterfactual, as the demand for HDD was declining while
demand for flash memories (SSD) was increasing, so what does this tell us about common demand shocks between memory storages (which were following different paths) and other IT sectors? Nevertheless, the authors are very careful and try different definitions for this alternative counterfactual in order to determine how well each works based on a battery of standard tests, which I will discuss below.

The third counterfactual, which essentially is a refinement of the previous one is to look at a weighted sample of IT firms. This seems to work best. The authors use a matching procedure to generate the weights. The underlying idea is that the propensity score—a statistic measuring the similarity between treated and non-treated firms that avoids the usual dimensionality problem of looking at similarity along many characteristics—can be used as a good measure for the weights. This is surely a good idea that is already used in the literature and makes sense in this context.

There is a typical criticism to this approach that also applies in this case. By using a propensity score approach—or, as I will discuss below a synthetic control method—the selection of the counterfactual is only based on observable characteristics, indeed each treated unit is compared to a non-treated unit that is as similar as possible along those observables. However, unobservable characteristics, such as innovation capabilities, intangible assets, and the ability to capture spillovers across technologies, might play an important role in this context. While one could potentially identify sources of ‘exogenous’ variation to account for the endogeneity of the treatment due to selection on unobservables—as Gugler and Siebert (2007) do for the DRAM industry—this is questionable and has its own issues. Some easier additional robustness checks that the authors could try include examining if there was a structural break along other dimensions (i.e. control variables) around the merger. Moreover, they could try applying the approach proposed by Altonji et al. (2005), who suggest that the amount of selection on the observed explanatory variables gives an indication to assess the amount of selection on the unobservables.

### 3.3 Inference & Robustness

The study team was quite careful in checking the quality of their identification strategy as well as the robustness of their inference and results. First, they carefully analyze the so-called common trend assumption. Indeed, key for identification is that the control and treated groups behave in a similar (parallel) fashion before the merger took place, thus indicating that the control group can be considered to be a good counterfactual for the unobserved behavior of the treated unit after the merger took place. The authors use both graphical and regression analyses to perform this check. While they are quite careful in this step, future work could be even more formal on this issue. For instance, both Ashenfelter et al. (2014) and Dobkin et al. (2014) propose some refined tests to check the common trend assumption. The empirical results of these tests show clearly that some counterfactuals do not work particularly well. I find it very important that the study team is very upfront on this issue and never tries to oversell their results.
For inference, the issue of serial correlation also plays an important role. As before, the authors are quite careful and actually much more transparent than many published papers. They conduct several tests and show when they work and when not. Specifically, the authors show that they find autocorrelation when using the R&D intensity data, but then claim that when using R&D intensity growth, i.e. by taking differences, they only find serial correlation when using IT firms as a counterfactuals. While this is reassuring, a possible additional approach to tackle this issue could be to use bootstrapped standard errors.

A final nice set of tests presented by the study team is provided through placebo analyses. First, the authors show that in a specification where other IT firms (and not the merging parties) are assumed to be the treated (i.e. to have merged), positive treatment affects are not found. In this case, however, the reader would have profited from a deeper and more transparent discussion. Second, they also tested for placebo treatment times by assuming that the time of the merger was different from when it actually happened. They then run a very large set of regressions and show that only for Toshiba they estimate significant placebo effects. Again, if compared to existing (even published) work, the approach of the case team appears to be very transparent and thorough.

4 Effect of the Mergers on Patents and Product Characteristics

The next chapters of the report focus on alternative proxies for innovation. The study team stressed in several parts of the report that, due to time and budget limitations, these additional analyses are less well developed than the previous one. Hence, the reader should bear this in mind when assessing the empirical evidence reported in this part of the study. Accordingly, also my comments will be more limited as it does not seem fair to discuss things that could not have be done due to limited resources.

4.1 Patents

4.1.1 Outcomes

Patents are also a commonly used as a proxy for innovation as an output. There is an immense literature on how (not) to use patent measures to assess innovation. The study team is well aware of this and discusses, at least partially, this enormous literature when describing which specific measures to adopt. For the (preliminary) analyses presented in this final report, they discuss patent counts, citation weighted patents, a ’factor variable’ identified through factor analysis, patent count stock, patent intensity, and moving average of patents.

The complexity of patent-based measures make this part of the analysis much more difficult and, potentially, open to criticism. The authors try again to be very comprehensive, discussing the pros and cons of different measures. Nevertheless, patent data are complex, noisy, and difficult
to deal with. One first important step when working with patents is to filter out insignificant patents (e.g. Harhoff, et al. 2003). The authors are aware of this and make a distinction between patents ‘about the relevant product’ vs. patents that ‘reference the relevant product’. They only consider the former, yet this might result in missing important effects, especially because the product market definition (HDD) might not be the relevant one when we think about the merger effect on innovation. In fact, it might be that some basic research and patents that are relevant for one product market might be also relevant for the other product markets. Moreover, there might be across-technologies spillovers that make innovation in the HDD market relevant for SSD and/or flash memories and vice versa.

The idea of creating a patent indicator by factor analysis is potentially interesting. Yet because it is not the standard in such a specialized and vast literature, I think it deserves some particular care. For instance, it is not clear how the authors use the time dimension. Apparently, the indicator is calculated by applying factor analysis to a number of different measures of innovation available in the patent data for the time period 2007-2014. Is it reasonable to use a sum or average of these characteristics over a relatively long time period? Could the information entailed in the evolution of these measures over time be taken into account?

When considering future extension of this work, the authors – or whoever else will perform such ex post evaluations – should also consider other alternative measures based on patents. For instance, a recent paper by Verhoeven, et al. (2016) proposes some measures aiming to capture radical innovation as ‘technological novelty,’ which are technological discontinuities, much like inventions, that introduce a novel technological approach. Another promising avenue for future research would be to mix the patent and product characteristics approaches by looking more carefully at which patents are contained in each product. Clearly both of these suggestions are way beyond the scope of this report. However, they are surely important steps on which future work could build.

4.1.2 Definition of the Counterfactual

The definition of the control group is different in the case of the patent analysis from what is done for R&D because the data available are much richer. Yet, even in this case, the authors follow their previous approach of proposing three different control groups: 1) the HDD patents of firms most active in patenting HDD-related innovations; 2) NAND Flash patents of the firms with most Flash-related patents; and 3) top ten storage firms in terms of number of HDD-related patents.

While these are a reasonable starting point, there is more room for improvement as the possibility to define additional control groups are much larger. This also points to one of the issues of having very rich data: more flexibility leads to a larger number of potential combinations and, therefore, increases the subjectivity of some choices. In principle as advocated by the study team,
the application of novel statistical methods, such as machine learning for model selection, could be extremely useful tools and open up new viable approaches, although one should never forget the economic rationale of the choice of a specific counterfactual. Moreover, the richness and complexity (for instance the high volatility) of the data should not compromise the identification strategy. Hence, future studies should put the same amount of effort put in the previous chapter to discuss issues such as the existence of a parallel trend or issue of autocorrelation, which are unfortunately not very transparent in this chapter and mostly relegated to the appendix.

As a bottom line to this discussion, one should bear in mind that the time and resources needed for research need to increase as the complexity of the data increases.

4.2 Product Characteristics

I think that the idea of looking at product characteristics and product positioning to proxy innovation is one of the most interesting parts of the study. Even in this case I think much more work needs to be done. Yet the study team identified some very interesting feasible avenues for developing this approach, which is not completely novel in the existing literature (e.g. the CIS innovation survey), but that has, I think, enormous potential.

The study team again proposes two different outcome measures: 1) the ‘release of new products’ and 2) ‘unit cost of new products.’ These seem to be very reasonable proxies of innovation diffusion or innovation as an outcome. Moreover, these measures have the potential to make a link to a different literature in the industrial organization tradition that sees products as a bunch of characteristics positioned in the product space. This link might be very important as it could lead to adopt different methodological approaches to tackle the issue of measuring the effect of mergers on innovation. Namely, instead of using reduced-form difference-in-differences analyses, researchers could rely on structural econometric methods. One specific idea proposed by the authors that seems very appealing in this context would be to "quantify the distance between a technical artefact, which includes the innovation, and the one that does not have it". At this stage of the analysis, it is not very clear if this concept is easily measurable and generalizable to other industries.

Concerning the potential control group, the only available counterfactual for this part of the analysis is to use SSD drives. This is unsatisfactory as SSD drives follow a quite different trajectory. One possibility to deal with this issue would be to consider vintages, i.e. compare the evolution of HDD drives with the evolution of SDD drives not at a given point in time but at a given point in the technology evolution. Other potential counterfactuals could be different components of personal computers and notebook, such as DRAM or processors, which also should move similar ways as they are subject to the same demand shocks as hard drives.

The issues with the choice of the right counterfactual brings me back to the point I made
above. Perhaps, and especially for this latter analysis, a structural approach could be feasible and (potentially) superior to econometric tools based on difference-in-differences.

5 Conclusions and Methodological Lessons for Future Work

To conclude, I find this report a very interesting and important step forward in the development of tools to retrospectively assess the working of antitrust agencies and their enforcement activities. The choice of the case study – with the caveats on the existence of concurring events highlighted above – seems to be appropriate as the industry under consideration has many appealing features for these kinds of analyses. The methodological steps discussed by the study team are thought through and well motivated, although some are better developed than others. The very rich and comprehensive approach, where the authors analyze several measures of innovation, using several outcomes for each, coupled with several potential counterfactuals, produces a very rich set of results. This is admirable even though it runs the risk of making it very difficult to find a simple, clear, and coherent story. Yet, when things are complex, as in innovation markets, perhaps it is not possible to tell a simple, clear, and coherent story.

Because of the richness of the approach and the many dimensions the study team examines, even a simple case study becomes a quite long venture. Therefore, some parts of the study still appear to be a bit ‘rouglier’ than others and would need more time to be fully refined. The authors themselves are upfront on this issue by prominently stressing that this is a pilot feasibility study and not a full-fledged ex-post evaluation.

This feasibility study highlights some major issues with respect to existing, more traditional retrospective studies. First, when looking at innovation, it is much more difficult to measure the outcome(s). The report presents some possibilities and discusses, I think carefully, what are the pros and cons of them. The discussion during the workshop at DG COMP highlighted that there are other possibilities. While clearly a work on innovation cannot avoid thinking about patents, especially because there is an enormous amount of research on which one can build on, other measures seem to be quite promising. The idea of looking at the outcome of innovation, such as product characteristics or productivity, seems especially novel. Furthermore, future research could think more carefully about the effect of mergers on cumulative innovation (e.g. Galasso and Schankerman, 2014).

As usual in ex post evaluations, the quality and granularity of the data is critical, having huge implications for the identification strategy. When looking at innovation, researchers are mostly facing variation within a firm over time and, potentially, variation across product markets. Rarely is geographic variation available due to the global nature of innovative activities. However, researchers could be more innovative and not only use differentiation in the product market but
also differentiation in the technology space to improve their identification strategy (e.g. Bloom et al. 2013).

While the core of this report is to discuss the feasibility of retrospective studies of competition policy enforcement on innovation, the pilot case study offers plenty of useful results worth discussing. The main results highlight quite some heterogeneous response across the three firms involved in the mergers. According to the findings, Seagate became more innovative after the merger: It increased R&D intensity growth, it increased its patent activities, and it increased its number of new products. Western Digital, instead, only seems to have increased patents (if at all), while Toshiba even seems to have decreased R&D intensity, patents, and product introduction. Looking at this heterogeneous response can be an interesting source of identification. Even if not, it is surely useful to better understand the evolution of these markets. Moreover, it would be interesting to understand whether these sources of heterogeneity can be imputable to different determinants, such as a firm’s innovation history, the existence of spillovers and absorptive capacity, other firms’ characteristics and/or product characteristics like size, diversification, and profitability.

There are two further important lessons for the development of the ‘ex post evaluation agenda.’ First, one key trade-off competition authorities often face is between considerations about static allocative efficiency (prices) and those about dynamic efficiency (investment, innovation, product positioning). I believe that a natural extension of these kinds of studies, especially for such dynamic and innovative industries, would be to simultaneously look at both dimensions. Recent research makes this point, even if not exactly in the context of merger evaluations (e.g. Genakos et al., 2016). Industries, such as the one analyzed in this report, would represent a perfect case study, as products are very well defined, while price and quantity data are probably easy to obtain.

Second, the experience of this report stresses again one important issue related to the amount of resources needed to perform a careful ex post evaluation. It is recognized that accurate retrospective studies are resource intensive and need time (Duso and Ormosi, 2015). The study team especially needs more time when the breath of the study increases exponentially, as in the present report. This is not only because the collection and use of different datasets and econometric techniques is time consuming but, and probably foremost, because the study team should have the time to incorporate the comments of different stakeholders, including the original case team, the Commission itself, as well as external, independent reviewers and/or seminar participants. The latter group is extremely important – at least it was for the studies I have conducted – to assure a substantial quality check and peer review. This is common practice in academic work and should also become common practice in retrospective studies. It would undoubtedly increase the quality and acceptance of such evaluations in the broader community and, thus, improve their usefulness for policy advocacy.
6 References


