UNCTAD 14
Nairobi, 17–22 July 2016
FOURTEENTH SESSION OF THE UNITED NATIONS CONFERENCE ON TRADE AND DEVELOPMENT

World Leaders Summit, High-Level Events, Round Tables, Side Events
World Investment Forum
Civil Society Forum
Global Commodities Forum
Youth Forum

What Role for Competition Policies in Regional Integration? The Cases of Africa and Latin America

Loita Room

Intercontinental Hotel, Nairobi

Wednesday, 20 July 2016

A methodological framework to capture the various economic impacts of competition enforcement

(Draft report)

CONTRIBUTION

from

Steve Davies and Peter Ormosi
University of East Anglia, United Kingdom

13 July 2016

Table of Contents

Executive summary

1. This paper provides a methodological approach to capture the various economic impacts of the work of competition authorities. As requested, we provide a literature review of best practices in relation to ex-ante and ex-post evaluations of the impact of enforcement and advocacy activities of competition authorities (CA) around the world.
2. For the ex-ante and ex-post evaluation methodologies, we have adapted to the developing country context our previous contributions to the evaluation of competition policy. For evaluation of advocacy, we also suggest a methodology whilst noting that this topic involves many uncertainties, such that no CA in the developed world has ever attempted to formally quantitatively evaluate the impact of its advocacy activities.
3. UNCTAD asked us refer to Kenya as an example of where these methodology might be applied. In the course of the project we were informed that it would not be possible to use Kenya, and we were asked to refer to Tunisia instead.
4. UNCTAD also asked us to suggest how to evaluate the distributional effects of CA interventions. Again, this has never been previously attempted for any country in the world

Ex-ante assessment (section 2)

5. In principle, there should be an ex-ante assessment undertaken at the end of every year. This should cover all decisions taken by the CAK on mergers, cartels and abuses. Initially this would be along the lines of Davies' proposed guidelines to OECD (Davies/OECD 2013), i.e. based on rules of thumb consistent with those used (when more precise methods are infeasible) by OFT and other authorities. This would not be time consuming and makes only relatively minor demands on data. It should be undertaken in-house by the CA's own staff.
6. At the year end, a designated 'evaluation team' within the CA should collate the set of all mergers cases, cartels & agreements, and abuses of dominance cases in which the CAK has intervened. It should estimate the consumer benefits associated with each intervention, by multiplying the turnovers of the parties involved by an assumed price change (3% price rise avoided for mergers, or 10% overcharge averted for cartels/abuses), and then by an assumed duration (between 1 and 6 years, usually 3, discounted to a present value using whatever is the appropriate social discount rate for the country concerned.) See section 2.4.

More detailed ex-post evaluation (section 3)

7. Each year, a sample case should be carefully selected by the CA management for a detailed retrospective evaluation.This would be a specific market for which the intervention decision was taken at least 2 (preferably 3 or 4) years previously in order that sufficient time has elapsed to observe impact.
8. There should be 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.
9. 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.
10. Identifying the causes of the analysed events and controlling for all factors that might have triggered the analysed event can help identifying cause and effect.

11. In the analysis of mergers, if estimates show that post-merger prices increase, it does not necessarily mean that the authority made a decision error.
12. Any ex post-study should carefully check the robustness of estimates to possible variations of the model/method.
13. Requesting a peer-review (typically by an academic) can improve the credibility of findings.
14. The reason that we believe that the CA should do both types of evaluation is that the ex-ante approach is the most viable way of generating an aggregate annual impact estimate in $, while the ex-post approach is the better way of drilling down to examine the detail of decision-making within the CA.

Advocacy (section 2.5)

15. We suggest that the CA's advocacy activities might be included in the methodology. While it is rare for CAs worldwide to include this in their formal evaluations of impact, advocacy is arguably the most important arm of a CA's activity, especially in its early years. If it is to be included, it will require the CA to identify all of its advocacy activities in a given year and then quantify them in a similar way to that used for cartels/abuse, but discounted, where appropriate, to acknowledge the fact that other parts of government might also have played an instrumental role.

The effect on distribution (section 4)

16. Making simple, but very reasonable assumptions, it is easily shown that competition policy is likely to have a positive impact in reducing inequality
17. This impact is likely to be greatest in markets for necessity products. This could be borne in mind in any prioritisation within the CA.
18. Given the uncertainties in measurement, we would caution against quantifying these impacts. However, if this is considered to be desirable, then a rough quantification would be possible if one is able to estimate the impact of an intervention on consumer and producer surplus, and to make a rough estimate of the share of consumption of the POOR. Assuming POOR is defined as the 40% poorest households in the population, their share of consumption would be at least 40%, often much more, except for luxury goods (with income elasticity >1).

Tunisia as a case study (section 5)

19. UNCTAD was unable to provide us with the quantitative data needed to conduct an illustrative real world example, although we do include an hypothetical numerical example.

1 Introduction

Terms of Reference (May 2016)

We have been asked “to provide a methodological approach to capture the various economic impacts of the work of competition authorities. Micro-economic impact as well as macroeconomics effects of competition authorities' enforcement and advocacy activities should be captured by the methodological approach”… Our report should also “include a literature review of best practices in relation to ex-ante and ex-post evaluations of the impact of enforcement and advocacy activities of competition authorities around the world, including their advantages and limitations.”

The methodological approach should capture the various economic impact of the work of competition authorities. Micro-economic impact as well as macroeconomics effects of competition authorities' enforcement and advocacy activities should be captured by the methodological approach. The methodological approach should evaluate the possibility for applying the following three methodologies in the context of developing countries:

- Evaluation for accountability;
- Evaluation of specific cases (anticompetitive cases such cartel, and abuse of dominance) or mergers; and
- Evaluation of distributional effects of the intervention of a competition authority.

Contents

2 Evaluation for accountability

Increasingly, CAs around the world are quantifying the aggregate benefits to consumers of their activities in an 'Impact Assessment'[1]; sometimes, the CA publishes a ratio of impact to its costs, and we believe that this is good practice. Even if not obligatory, such assessments should typically be in the CAs' own interests in demonstrating publicly that they deliver 'value for money' – assuming, as is likely, that the gains from preventing/detecting just a few anticompetitive mergers or cartels more than outweigh the relatively small budgets of most CAs. Moreover, any government agency has a responsibility to the taxpayer to be seen to be socially valuable. This is not to deny that there may be counter arguments as to why impact assessments should not be conducted; but, if so, these should be articulated, and this in itself requires the CA to consciously consider the pros and cons of this form of accountability.

This section describes a methodological approach for evaluating, for accountability, the impact of competition policy and law in the context of developing countries. Section 2 surveys the previous literature. In particular, it identifies those CAs in the developed world who have been most active in developing methodologies for impact assessment. It identifies what might be described as the essential features of Impact Assessments as they have been conducted to date. It goes on to provide more micro details on the assumptions and methodologies used to evaluate impact in the core constituent parts (merger control, cartels and abuses of dominance) by the five CAs who regularly report the results of their Assessments. Section 3 briefly refers to the CMA as a best practice example of an Impact Assessment at work. Section 4 then proposes the methodology that we believe is both practical and relatively un-time consuming: we believe this would be practicable in the developing world context.

2.1 Literature Survey

By drawing on the existing practices of those CAs which already publish regular appraisals – Competition and Markets Authority (CMA)/Office of Fair Trading (OFT) in the UK, the Department of Justice (DoJ)/Federal Trade Commission (FTC) in the US, European Commission (EC) in the EU, and the Dutch Competition Authority (ACM or its previous acronym NMa) – it is possible to identify ten defining features of an impact assessment.

Defining features

With some minor variations, they reflect common practice, and are, in our opinion, appropriate for the present purpose.

1. Impacts are assessed on a **regular, usually annual,** basis, during the following year (but see point (7) below).
2. They are **relatively undemanding in cost and time**, usually utilising information collected at the time of interventions and/or simple default assumptions[2].
3. Estimates are generally performed using **ex-ante information**. Impact assessments are conducted once interventions have been undertaken, but using only the information

[1] An OECD questionnaire survey (OECD, 2012) reveals that in 4 countries this is an obligation, and in 12 others it is conducted in some form on a voluntary basis.

[2] Having observed, advised and reviewed the CMA/OFT's impact assessments over a number of years, Professor Davies make this observation not just as a disinterested outside academic.

available ex-ante, i.e. that available at the time of the intervention. The practitioner projects forward comparing what would happen with and without the intervention. In general, ex-ante evaluation is simpler to conduct and makes fewer demands on data than does ex-post evaluation which can only be conducted some years after, when accurate data becomes available on what actually did happen (see, for example, Davies and Lyons (2007 pp.106-7)). Given that impact assessments are typically conducted in the year following intervention, they must, almost inevitably, use only the information available ex-ante. This helps to achieve (3) above since little data collection or analysis should be required. It also leads inevitably to some of the following.

4. It is assumed that no intervention can have a negative impact. This is a direct consequence of (3) - given that the information used is confined only to that available at the time of the original intervention, and by definition the CA must have projected a positive outcome from intervening. Of course, it is important that CAs should also evaluate from a more self-critical perspective, but that is the role of the ex-post evaluation. As mentioned already, these are only possible some years after the intervention. Since ex-post assessments are invariably much more costly, they can only be conducted on a more ad hoc basis, usually focusing on only a small sub-sample of cases.
5. Estimates are deliberately **'conservative'**. This is essential, given the ex-ante basis and point (4) above. Obviously, the term 'conservative' is relative, and in the concluding section below, it is distinguished from 'lower bound' which is meant to indicate the least positive possible outcome.
6. Estimates are in terms of **static consumer benefits.** The emphasis on consumer benefits is appropriate, reflecting the fact that competition enforcement is guided by the consumer welfare standard. Typically, only static benefits are calculated, i.e. reduced price/increased consumer surplus. Dynamic dimensions, e.g. innovation and future products are rarely quantified (see (10) below).
7. **Annual moving averages** are usually employed[3]. As mentioned in (1), estimates are usually made annually. However, there can sometimes be considerable variability between years due to the erratic frequency over time in cases from abnormally large markets. Therefore it is the practice of some CAs to report estimates in moving average form. For example, the estimate for 2015 would be an average of the estimates for 2013-2015. This smooths, rather than entirely removes, the sensitivity of estimates to say very large but infrequent mergers or cartels.
8. **Estimates are 'point' estimates, rather than as a range of plausible values.** In principle, the alternative would be to present estimates in the form of a range, based on a statistical confidence interval. But this is impossible given the ad-hoc nature of many of the estimates. Sometimes however a CA might present alternative estimates for different assumptions – typically 'high' and 'low'.
9. Assessments typically cover **mergers and cartels,** and usually **abuses of dominance.** Ideally, all areas of competition policy should be included. In the UK, so too are market studies, and in the Netherlands so is some regulation.
10. The **deterrence** impact of competition policy is sometimes mentioned in impact assessments, but it has **never been estimated** in detail. Similarly the possible beneficial

[3] Of the 5 CAs reviewed below, the FTC uses a 5 year moving average; the OFT and NMa use 3 years.

effects of competition policy on productivity, innovation and growth are not estimated. As stated in (6) impact assessments are essentially static.

2.2 The Detailed Assumptions of Impact Evaluation

This sub-section describes the detailed assumptions used by the CMA[4], DoJ, FTC, EC and ACM to estimate the impact of individual cases in merger control, cartels and monopoly abuse, which go to make up the aggregate estimates. These are the five CAs who regularly publish the results of their assessments. At the end of the section, other CAs are discussed more briefly.

To estimate the impact of any individual case intervention, information is required on:

(a) the size of the affected turnover,
(b) the price increase removed or avoided and
(c) the length of time the increased price would have prevailed absent the intervention.

The product of (a) and (b) provides an estimate of the magnitude of consumer overpayment avoided, per annum. For mergers, an adjustment is also often made for the deadweight loss (i.e. surplus lost by consumers who are deterred from making any consumption, see below).[5] The full impact is then derived by multiplying this product by (c), the number of years the intervention is assumed to be effective (i.e. for how many years the cartel or abuse would have continued, or before an anti-competitive merger would provoke entry or offsetting expansion by rivals). The gains in future years are discounted by the CMA, EC and NMa.

Tables 1-3 summarise the assumptions of the five authorities, for the OFT and NMa these are in the form of 3 year moving averages, the FTC reports its estimates as 5 year moving averages. A previous paper (Ormosi 2012, Table 1) provides some analysis of time series data for four of these authorities.

2.2.1 Size of affected turnover

In principle, this 'volume' measure is the easiest to estimate – the information is normally available for the CA and can be easily recalled from the original intervention. However, there is a judgement to be made concerning precisely what is defined as the affected turnover. The narrow option is to define it merely as the turnover of the directly intervened firms, i.e. the merging parties in those markets in which they overlap, or the turnover of cartel members, or the firm(s) abusing their dominance. However, theory suggests that there should be knock-on effects to the prices of the parties' rivals. So the wider option is to define the affected turnover as that of all firms in the relevant market.

From the wording of the public documents of the 5 CAs, it appears that, for cartels, the narrow definition tends to be adopted, but for mergers and abuse, it is less clear which definition is used.

The more challenging parts of the evaluation are how to estimate the magnitude of the price effect and duration. For all five authorities, the preferred option is to draw, if possible, on information collected at the time of the original investigation (for example, on the magnitude of a cartel

[4] In the UK, the separate Phase I and Phase II agencies, the Competition Commission (CC) and the Office of Fair Trading (OFT) were merged to form the Competition and Markets Authority (CMA). Until 2014, the Impact Assessments were conducted by the OFT, but thereafter this became the responsibility of the CMA.

[5] In that case, the estimate of consumer savings is the pre-merger turnover multiplied by the estimated price rise the merger would have caused minus the deadweight loss that this would have caused.

overcharge). However, in many cases, such suitable data are unavailable, and the CA then employs simple simulation or assumes some default value for each of these two factors in their evaluations. These are now briefly described as follows.

2.2.2 The Price Effect and Duration

Cartels

Table 1 summarises the assumptions used in the evaluation of the impact of cartel investigations. Where the case details are insufficient to use a case-specific estimate of overcharge, the norm is to assume 10 per cent, although the CMA and EC (sometimes) employ a 15 per cent default.

Assumptions on the expected future life-span of cartels show wider dispersion. At one end of the spectrum, the DoJ and NMa assume just a notional 1 year[6], while at the other end, the CMA assumes 6. The EC lies in between, using 1, 3 or 6 years depending on its judgement as to the future sustainability of the cartel at the date of detection.

Table 1 : Assumptions used in cartel cases

	EU	USFTC	USDOJ	CMA	NMa
Affected consumers	Affected market	n/a	Volume of commerce	Turnover affected goods	Affected market
Price effect	10-15%	n/a	10%	10-15%	10%
Duration (years)[1]	1/3/6	n/a	1 or number of months for shorter lived	6	1

Sources: European Commission DGCOMP (2012), USDOJ (2012), USFTC (2012), CMA (2015), NMa (2010), NMa (2012).

[1] CMA and EC discount future gains at 3.5% p.a.

[2] £ for CMA, $ for FTC and DoJ, euros for EC and NL

Mergers

In merger cases, if the price impact of a merger was estimated during the original investigation, this can be used in impact evaluation. If not, simulation models appear to be the most preferred option, projecting how prices, demand, and market shares might have changed had the merger been cleared without the CA's intervention. In other cases default assumptions are made on the price impact, as summarised in Table 2.

Table 2: Assumptions used in merger cases

	EU	USFTC	USDOJ	CMA	NMa
Affected consumers	Size of relevant market	Volume of commerce	Volume of commerce	Turnover of affected goods	Size of relevant market

[6] Indeed, in the case of cartels which are less than 1 year old at the time of detection, DoJ projects future duration to be the same number of months as its age at detection.

Price effect	Simulated[1]	simulated, if not 1%[2]	simulated, if not 1%[2]	simulated, if not, average of simulated [2]	Case specific, sometimes simulated. If not, 1% [2] [1]
Duration (yrs)[1]	2-7	2	1	2	1

[1] plus assumed 1% due to enhanced efficiency

[2] plus deadweight loss estimate

The simulation models are chosen, as appropriate, from three candidates – Cournot for homogeneous products and Bertrand (e.g. PCAIDS or ALM) for differentiated product industries. Increasingly in recent years, for the CMA/OFT, 'simulation' is based on estimates undertaken at the time of the investigation of the likely Upward Pressure on Prices (UPP) or related techniques (OFT 2012). Previously the EC assumed that future customer savings resulting from corrective merger decisions corresponds to 10% of the size of the relevant market(s). It has now been changed to a practice whereby price effects are typically simulated on a case-by-case basis.

Where simulation is not appropriate – either because none of the above models adequately describes the nature or competition, or because data for calibration are unavailable, a very low default assumption of just a 1% price raising effect is made (FTC, DoJ and NMa), although the NMa adds 1% to this, representing an 'efficiency effect'.

The duration assumptions almost universally being either only one or two years, although the EC employs more case-specific discretion.

Abuse of dominance

As argued by Werden (2008), Davies (2012) and Ormosi (2012, section 1.5)[7] abuse of dominance cases arguably pose the greatest challenges for impact assessment. Similarly to other case types, if case-specific information is not available, default assumptions are made about the price effect and likely future duration of the infringement (see Table 3.) Here, there appears to be a dichotomy between the OFT and NMa on the one hand, which treat abuse similarly to cartels, and the FTC and DoJ on the other hand, which treat them similarly to mergers. The EC now no longer publishes estimates of its impact in this area, explaining that it fears that the small number of cases involved might prejudice confidentiality.

Table 3: Assumptions used in abuse of dominance cases

	EU	USFTC	USDOJ	CMA	NMa
Affected sales	N/A	Volume of commerce	Volume of commerce	Turnover affected goods	Turnover affected goods of abusing firm

[7] See also Davies' presentation (2012) at the previous meeting of this Working Party.

Price effect	N/A	simulated, if not 1%[1]	simulated, if not 1%[1]	10%	10%
Duration (yrs)[2]	N/A	2	1	6	1

[1] plus deadweight loss estimate

[2] CMA and EC discount future gains at 3.5% p.a.

Others areas of policy

Beyond these three cornerstones of competition policy, the CMA also includes market studies in its evaluations. In addition, some of the CAs also conduct some assessment of other advocacy activities and of their consumer protection initiatives. For example, the CMA compares the pre- and post-intervention number of consumer complaints, and a reduction in the number of complaints is converted into a financial estimate of avoided consumer detriment by valuing each complaint at a proportion of the purchase value. As a performance measure, the FTC reports the number of consumer complaints and the percentage of the FTC's consumer protection law enforcement actions that target the subject of consumer complaints to the FTC. Other methods such as consumer satisfaction surveys are also applied in other countries.

2.2.3 Other Competition Authorities

As far as is known, the five above authorities are the only ones who conduct regular annual impact assessments with results placed regularly in the public domain. However, it is clear from the responses to the OECD (2012) questionnaire that a number of other authorities undertake similar sorts of assessment, albeit sometimes only on a particular part of policy, and/or non-routinely.

The most comprehensive of these appear to be Hungary and Mexico, who employ similar defaults for assumed price rises to those of the five above authorities (although Mexico employs a 20% default price effect for mergers.) Germany and Japan have also both evaluated the impact of cartel interventions using default price increases.

On the other hand, one authority (New Zealand) reports that it has trialled a methodology along the lines of CMA but has decided not to adopt it – at least for the present – because so many impacts are inherently un-measurable. In particular, it is concerned with the deterrence issue. Not only is this difficult to measure, but, more fundamentally, the authority is concerned that a CA's success in deterrence will have the perverse effect of lowering measurable impact – at least if impact is measured along the lines described in this paper.

2.3 Best Practice: CMA's Impact Reports

This section describes the impact evaluations undertaken by the UK's Competition and Markets Authority (CMA). The CMA is highlighted because it (and the OFT before it) have been the most transparent and prolific, of all authorities around the world, in publishing details of its work in this area. The CMA is therefore used here as an illustration of what can be called world best practice.[8]

[8] Davies has worked with and advised the CMA/OFT's impact evaluations over a number of years, and is very familiar with them at first hand;

Each year the CMA publishes an Impact Assessment, which quantifies the value of the welfare benefits of its interventions compared to how much it costs the UK taxpayer. The most recent published example is for year 2014/15 (CMA, 2015).[9] Its methodology corresponds very closely to the 10 defining features listed in the previous section. Ideally, the methodology involves evaluating the anticipated benefits for consumers of all the cases CMA handles in the given year.

Mergers

The CMA's merger control methodology entails assessing the impact of all its interventions in merger cases. Hereafter, we use the term merger interventions to refer to remedies (assurances and commitments agreed with the parties before clearance is agreed) and outright prohibitions (which occur only very rarely). It relies on what might be called simple simulation techniques. In the early years of PI, the evaluation team would choose and then simulate a model for each merger, as appropriate, from three alternative candidates – Cournot for homogeneous products and PC Almost Indeal Demand System (PCAIDS) or Anti-Trust Logit Model (ALM) for differentiated product industries. Increasingly in recent years however, the simulation is based on estimates undertaken at the time of the investigation by case teams of the likely Upward Pressure on Prices (UPP) or related techniques (OFT 2012). Invariably, these are derived from estimates of diversion ratios (often estimated from consumer surveys). The model is calibrated using low, medium and high assumed values of the industry elasticity (which is often a key input in the model), and uses the medium as the chosen point estimate. The conservatism relates to the range of the elasticity assumptions. Off model adjustments may also be necessary to accommodate any properties of the merger or market not picked up by the models.

Where simulation is not appropriate – either because none of the above models adequately describes the nature or competition, or because data for calibration are unavailable – the CMA uses a default assumption. It is assumed that the ratio of consumer savings to turnover for the case is equal to the average of the lower bound ratios across all the mergers which could be simulated in that year.

Cartels[10]

Where possible, estimates are based on information from case teams, but where this is not available, "rules of thumb that are consistent with international best practice and recent academic research" – typically, an avoided price rise of 10% or 15% is assumed.

Abuse of dominance[11]

There has been typically very few Article 102 cases evaluated. The methodology is similar to Article 101, but with the difference that the assumed default price rise is 10%.

Market studies

In the UK system, a market study/investigation is conducted by the CMA where there are reasons for believing that a market is not functioning well – not necessarily because of any anti-competitive behaviour by the firms. These sometimes lead to remedies, and the impact of these remedies also

[9] The 2015/16 report is imminent at the time of writing.

[10] Often referred to as Article 101 in EU parlance.

[11] Often referred to as Article 102 in EU parlance.

forms part of the CMA's impact evaluation. These are again conducted using conservative assumptions is stressed, but this is the one area where the evaluations are not always ex-ante. Monitoring of subsequent market developments may lead to revised assumptions and updated impact estimates.[12]

Consumer protection

CMA also attempts a partial evaluation of its consumer protection work. The methodology compares the pre- and post-intervention number of consumer complaints, and a reduction in the number of complaints is converted into a financial estimate of avoided consumer detriment by valuing each complaint at a proportion of the purchase value of the product or service concerned.

Table 4 summarises the CMA's estimates of consumer savings for 2012-15. As can be seen, in aggregate, the ratio of benefits to costs just exceeds 10:1. Table 2 shows that similar ratios have been achieved in all previous years since the PI began.

Table 4: Estimated annual CMA/OFT/CC consumer savings and costs for 2012-15 (£mn)

Competition enforcement	65	
Consumer protection enforcement	79.2	
Merger control	23.7	
Market studies	576.6	
Total benefits		744.5
Total OFT/CMA costs (averaged over 2012-15)		66.5
Benefit/Total CMA costs		11.2:1
Source: CMA Impact 5ssessment 2014/15, CMA47 (July 2015, p.3.)		

Bearing in mind the CMA's activities not included in the estimates, and the deliberate strategy of using conservative assumptions to estimate the benefits in individual cases, CMA remarks (2015, p. 3) that "we regard our estimates of direct financial benefit as being on the conservative side."

2.4 A Proposed Methodology

Against this backcloth, we now propose the following methodology for ex-ante impact evaluation for developing countries.

A good starting point are the 10 essential features outlined in section 2. Broadly speaking, 1-8 describe what is already more or less common practice (albeit with some exceptions on some of the points) amongst the incumbents. These should underpin any impact assessment.

On the coverage of the assessment (point 9), it is common practice amongst CAs already undertaking impact assessments to evaluate the core (mergers, cartels and abuse) even if separate figures for each are not always published. Again, it would also be sensible for CMAs from developing countries to do the same. Beyond this, some CAs also evaluate other areas (e.g. market studies in

[12] See section 4.5 below. Specific examples of OFT's are published on its website at: http://www.oft.gov.uk/OFTwork/research/evaluation/Evaluation-completed. One resent example is Bank Personal Current Accounts.

the UK, and sometimes consumer protection in some authorities), but because these may be idiosyncratic, and anyway often very difficult, these are best thought of as optional extras.

On point (10) – widening the scope of evaluation to include deterrence, productivity, growth etc. – these are matters all requiring future research (see Davies (2012) and Ormosi (2012)). Without doubt, the absence of an estimate of the benefits from deterrence is the most uncomfortable gap in impact evaluations. But, while some CAs have made tentative first steps towards quantifying deterrence[13], there is still insufficient understanding of the likely magnitudes involved to support the routine inclusion of, say, a deterrence multiplier in impact evaluations.

Assume then that an impact assessment is defined by 1-8 of the essential features and is confined to mergers, cartels and abuse, which methodologies and assumptions should be used for cases in each of these three areas?

Again following the general existing practice, the first best solution is to draw wherever possible on the original investigations for case-specific information on the likely price rise of a merger or overcharge by a cartel, or dominant firm abusing its dominance. Similarly, case details may sometimes support informed guesses on likely duration. To operationalize such case-specific information, it may be appropriate to employ back of the envelope simulations, for example in the case of mergers using diversion ratios and UPP methods to estimate the price effect.

Failing any such case-specific information, default assumptions will be required. Table 5 recalls the range of assumptions used by incumbents, and these are now discussed and compared in turn.

Table 5 Possible default assumptions (i) current practice

	Mergers	Cartels	Abuse
Affected turnover	Parties' turnover only or Market turnover?		
Deadweight loss averted	Include or exclude an allowance?		
Price rise removed/averted	1%	10-15%	1-10%
Duration (years)	1-7	1-6	1-6

Affected turnover

As mentioned earlier, economic theory suggests that the price effect of a merger/cartel/abuse is unlikely to be confined to just the parties involved. In nearly all oligopoly merger models, the rational response of rivals is to increase price in response to an increased price by the merging parties – broadly speaking, the rival price increase should match the parties' for coordinated effects mergers, or to be somewhat less in unilateral effects mergers. Similarly, it is rational for cartel

[13] The NMa and OFT have both previously commissioned surveys of competition lawyers etc, designed to ascertain typically how many cases are deterred for every case intervened, and they show how such information could be used to multiply up the observed impacts to allow for indirect deterred impacts. Apparently, the Lithuanian Authority uses such multipliers in its impact evaluations. The German Cartel Office also makes a fairly arbitrary conservative allowance in its assessment of the benefits from cartel prosecution (2010, p.16). It assumes, as a very bare minimum that at least one cartel was deterred in the period 2003-7, and this would have benefited consumers by several million euros.

outsiders to charge a higher price when competing with a cartel than when competing with the same firms not cartelised. In monopoly abuse cases, there is no unambiguous lead from theory; for example, partially foreclosed rivals may be forced to price lower if they are to maintain a foothold in the market. Generally speaking the narrow definition (parties' turnover) will lead to an underestimate of averted harm, while the wider definition (market turnover) will over-estimate it – probably particularly in abuse cases. In the spirit of conservatism, perhaps the narrow definition should be used.

Deadweight loss averted

Especially for mergers, some of the incumbents allow for the deadweight welfare loss averted by the intervention in their assessment. In other words, if a merger prohibition or remedy averts an *x* % increase in price, this benefits not only those consumers who would continue to buy post-merger, but it also deters some consumers from leaving the market and their surplus should be included in the estimate. Similarly, the striking down of a cartel benefits not only those consumers who were buying from the cartel but also those consumers now attracted to enter the market by the new lower price. In order to make this adjustment, an estimate is required of the demand elasticity.

Again in the spirit of making conservative estimates and in simplifying the exercise, there is a case for ignoring this adjustment even although it is academically uncomfortable to do so. But either way, there is also a case for treating mergers, cartels and abuse identically.

Price effect

For cartels, the common assumption is 10%, but with CMA and sometimes the EC opting for 15%. The latter seems more in keeping with the empirical evidence in the academic literature. Much of the evidence (for example, see the meta-study by Bolotova and Connor (2006)) suggests that the median cartel overcharge lies between 17 and 30 per cent[14]. However, 10% probably better qualifies for the 'conservative' label.

For mergers, the generally adopted default of incumbents is only 1%, although this is probably the area where the default is used least frequently. On the face of it, 1% appears to be an absolutely low interpretation of the term 'conservative'. It is extremely unlikely that any CA would attempt to intervene in any merger for which the expected price increase was only 1%. Unfortunately, unlike for cartels, the academic literature sheds little light on what might be a general ball-park figure for how much the 'typical' anti-competitive merger raises price, although there are some illustrative US studies on intervened US mergers. Kwoka (2013) reviews price estimates of the effect of 53 transaction (46 of which were mergers). The average price change after the 53 transactions was 6% (the average price increase for mergers only was 7.3%). Looking at those transactions only that increased prices the average effect was 9.4% (9.85% for mergers only). As a matter of personal opinion, we believe that raising the 1% default to say 3% would be justified – recognising that most CAs would be reluctant to intervene in any merger unless it felt confident that it would raise price by at least that amount. This would help to rectify what we see as an anomaly in present practice which weights each intervened cartel 10 times more heavily than each intervened merger.

The lack of any wide-ranging empirical academic survey on abuse of dominance cases makes it difficult to assess the prevailing default assumptions used by the incumbents, although the dichotomy between the US (1% in the absence of simulation) and the European authorities (10%) is

[14] More recently, Smuda finds a mean of 20% for a cleaned subset of this database.

stark. As discussed in previous meetings of this working group, more research is surely needed in this area. In the meantime, there is a case for treating cases of abuse identically to cartels, i.e. assuming a common default of 10%.

Duration

As seen earlier in Table 1, there is considerable diversity between CAs in their default duration assumptions for cartels. In assessing these alternatives, it is tempting to turn to previous empirical studies of cartel duration such as Block, Nold et al. (1981) and Levenstein and Suslow (2006). However, these studies were conducted on the duration of cartels at the time when they were detected, and these cartels are likely to differ from undetected cartels, which are more appropriate for the implicit counterfactual used here, namely, how much longer would these cartels have survived *had they not been detected?* In principle, the expected future duration at each point in its lifetime can be predicted using a survival analysis of past cartels – but only those which died a natural death. As far as is known, no such study along these lines has been attempted to date. In the meantime, we do know from the academic theoretical literature that there is a variety of determinants of cartel duration (see Ormosi 2012, p.5). These include the severity of fines and leniency programmes, as well as the type of the industry, specific market conditions, and entry conditions. Given this potential heterogeneity, the EC's case-dependent approach is quite persuasive – it selects a future duration depending on how sustainable it assesses each cartel to be at the date of duration. However, this does require a significant judgemental input, and if this was considered unattractive, then convergence on a single number, probably somewhere between 1 and 6 years, might be appropriate.

For mergers, the corresponding concept is for how long would the adverse consequences of an anti-competitive merger continue before offsetting market self-correction would occur - new entry or rival expansion. In this case, it seems unlikely that a CA would ever choose to intervene if it believed that post-merger offsetting market correction would occur rapidly within the following one or two years. Thus again one might argue that the incumbent authorities are being overly conservative. The one exception to these very small lower bounds is again the EC, which classifies mergers into one of three groups: "significant", "high" and "very high", depending on its assessment of the height of barriers to entry or rival expansion. As for cartels, there is a trade-off here between introducing discretion (and implicitly heterogeneity between CAs) and choosing a fairly arbitrary common default. Again, a future research survey might be helpful – in this case a survey of retrospective studies of anti-competitive mergers. But pending that research, the EC approach is perhaps the most appropriate.

Very similar comments on abuse cases apply as to those made above for cartels. Arguably, self-correction by markets to abusive foreclosing behaviour might typically be longer than for cartels, but there is no body of available previous research surveys to corroborate this expectation.

Table 6 pulls this discussion together by making what we consider to be defensible conservative default assumptions. Needless to say, each would merit further discussion.

Table 6 Possible default assumptions (ii) tentative proposals

	Mergers	Cartels	Abuse
Affected turnover	Parties' turnover only		
Deadweight loss averted	No allowance		
Price rise removed/averted	3%	10%	10%
Duration (years)	1-6*	1-6*	1-6*

* Assumed duration to lie in this range, depending on the judgement of the CA. Alternatively, 3 years might be an appropriate conservative norm.

2.5 Advocacy

The question of whether or not to include advocacy in impact assessment merits separate discussion. On the one hand, as far as is known, no CA has attempted to quantify the impact of advocacy, in terms of estimated benefits to consumers. On the other hand, competition authorities typically allocate a significant proportion of their budget to advocacy activities[15].

There is no doubt that the task of quantitatively estimating the impact of advocacy is challenging – much of this activity is general and intangible in nature, with specific effects that are not amenable to measurement, or disentangling from other influences on policy decisions. In the circumstances, most of the little evaluation work that has been done is predominantly qualitative. For example, the OFT (2010b) conducted a survey of officials across various government departments asking how far its competition related advice was taken into consideration and influenced policymakers. Judged on the replies from 43 respondents, it did indeed have a significant impact – leading to important changes in policy approach in half of cases, and changes in objectives in one quarter. However, this study did not attempt to quantify this impact.

The OFT also describes three case studies which illustrate quite well that, while advocacy can have important positive impacts, they are difficult to measure. The first was advice given to the Ministry of Justice, warning against licensing regulations that would have posed a serious barrier to entry into the market for will-writing; the second related to energy-efficiency in light bulbs, on which the OFT warned against voluntary arrangements which might facilitate collusion; the third was on how to improve procurement guidelines for competitive tendering in public procurement of waste management services. Impact in all three cases would be difficult to quantify. For example, in the light bulb case, speculation would be needed on the probability that collusion would have occurred, and the extent of that collusion.

Nevertheless, we believe that more work can be done in this area, and that it should be included in UNCTAD's methodology, using parallels which can be drawn with abuse and cartel enforcement, and merger control. Given that some anticompetitive regulations allow firms to coordinate prices or prevent entry, methodologies applicable to cartels and abuse/market foreclosure should also be applicable here.

Of course, there will be problems in identifying appropriate counterfactuals. For instance, to quantify the impact of advocacy to drop a piece of draft legislation which would have had

[15] An International Competition Network(ICN) (2002) study reports that, amongst those countries that felt able to quantify, almost one third reported that they devoted 20-30% of their budget to advocacy.

anticompetitive consequences, we would need to know what would have happened had the anticompetitive proposal been enacted. For enforcement activities, such as cartels, there is always the possibility of using a similar market as counterfactual, but in this case it would have to be a similar jurisdiction, something that would be even harder to find. It may be equally difficult to establish that a proposed piece of legislation was prevented as a result of CA advocacy rather than other political considerations. It would also be difficult to identify what version of the proposed legislation would have been accepted in the absence of the successful CA advocacy.

Nevertheless, given the particular importance of advocacy for a relatively young CA, operating within an emerging economy, we make the following proposal:

(i) In the ex-ante evaluation, informed opinion could be sought on how many pieces of legislation were changed in the year due to CA advocacy. Then, for each market concerned, impact could be calculated using the same rules of thumb as above for cartels/abuse: 10% reduction in price over 3 years (discounted). This would then be scaled down by an informed estimate of share of benefit that could be reasonably attributed to the CA in the policy decision – perhaps 30%, 50% or 70%.

(ii) In the ex-post evaluation, there should be one advocacy activity included – either annually or biennially.

3 Evaluation of individual decisions

There has been a recent surge in the number of studies that retrospectively evaluate the effect of various anti-competitive practices and competition enforcement decisions. These studies can help improve decision-making by identifying and correcting previous mistakes, evaluate the predictive power of analytical tools used in enforcement, improve transparency, and increase public support for competition policy.

Conducting ex-post evaluations of competition authority decisions is often hindered by the lack of in-house expertise, or limited resources. However, thanks to the recent attention, there is now a considerable body of knowledge on how to conduct these evaluations. Even less experienced institutions can build on this knowledge to deliver such studies. For example the OECD *Reference guide on ex-post evaluation of competition agencies' enforcement decisions*[16] is a comprehensive source of such references. The main objective of this report is to provide assistance and guidance for competition agencies that are still new to the area of ex post evaluations.

The OECD reference guide provides a broad discussion of all possible methods used in ex-post evaluation. All of these methods have got their pros and cons depending on their application. Below we focus on one of the methods, difference-in-differences, which is the most widely accepted robust empirical method for such evaluations. We will also offer more specific guidance for the three main case types: mergers, cartels, and abuse of dominance.

3.1 Objectives of ex-post evaluation

A recent European Commission document[17] listed the main objectives of ex-post evaluations are to:

- *Improve the effectiveness of competition policy decisions and the enforcement practice of the CAs, and improve the quality of competition law*: One of the main reasons for engaging in ex-post evaluations is to find out if the CA was right in its decision to intervene or to refrain from intervention. Of course the evaluation of a single case cannot tell us much about the effectiveness of enforcement practices and the appropriateness of the law. Instead, a large number of such evaluations would be required in order to be able to make inferences on the quality of enforcement in general.
- *Set internal priorities*: Ex post evaluations can direct the CA towards the more problematic cases. Also, it could help identify cases where insufficient resources were allocated by the CA and avoid similar mistakes for the future. hints of which industries are the most likely to
- *Defend legitimacy and improve advocacy*: it is key that the CA is able to demonstrate to the public that its activities benefit the economy and the welfare of consumers.

Whilst all of these are important for any competition authority, advocating for the legitimacy of competition enforcement is particularly vital for young authorities. Demonstrating how much consumers gained from specific cases could boost public support for the mandate of the CA, and could help convince politicians that competition policy is a worthy cause. It can also be used to justify an expanding budget.

[16] http://www.oecd.org/daf/competition/reference-guide-on-ex-post-evaluation-of-enforcement-decisions.htm

[17] European Commission (2015) Ex-post economic evaluation of competition policy enforcement: A review of the literature, http://ec.europa.eu/competition/publications/reports/expost_evaluation_competition_policy_en.pdf

3.2 Selecting the cases

The selection of cases to evaluate is usually strongly influenced by three main factors: (1) the importance of the cases, (2) the complexity of the case, and (3) data availability.

High profile cases are typically the ones where there were many consumers affected, the harm was very large, or the industry was strategically important. The evaluation of these cases can offer important ammunition for the CA to use in advocacy activities as the harm avoided (provided that the intervention was adequate and successful) is likely to be very high. Strategically important industries, or industries where a large number of consumers are likely to be affected by the anti-competitive practice, are also more likely to attract media attention and publicity for the CA.

Complex cases are typically the ones where it is difficult to establish whether a behaviour or a merger is anti-competitive or not. For example, with cartels we know that they offer very little (if any) beneficial effects, therefore it is typically straightforward to establish that they had net anti-competitive effects. On the other hand, mergers can have pro-, as well as anti-competitive effects, and the two needs to be weighed against each other carefully. The same is true for abuse of dominance cases. In these situations it is useful to do ex-post evaluations following investigations, especially if it was not easy to predict whether the pro-, and anti-competitive effects dominate. These marginal cases are the ones where the ex-post study could be most informative for the CA's internal purposes (to help fine-tune policy, or readjust priorities) but because of the uncertainty of the outcome (the study might show that the original decision of the authority was mistaken), it may be less preferred by CAs who want to use the ex-post evaluation for advocacy purposes to gain public support.

Data availability inevitably determines which industries are analysed. A comprehensive study on US merger retrospective by John Kwoka 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.

When selecting cases to evaluate, one has to try to ensure that the case is equally high priority on each criterion. Moreover, other considerations can play also a role. The analysis is somewhat easier and cleaner for end-consumer and retail markets. 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 would require a number of assumptions and additional thinking on how to apply known methodologies in a sensible way.

Finally, 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 could 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).

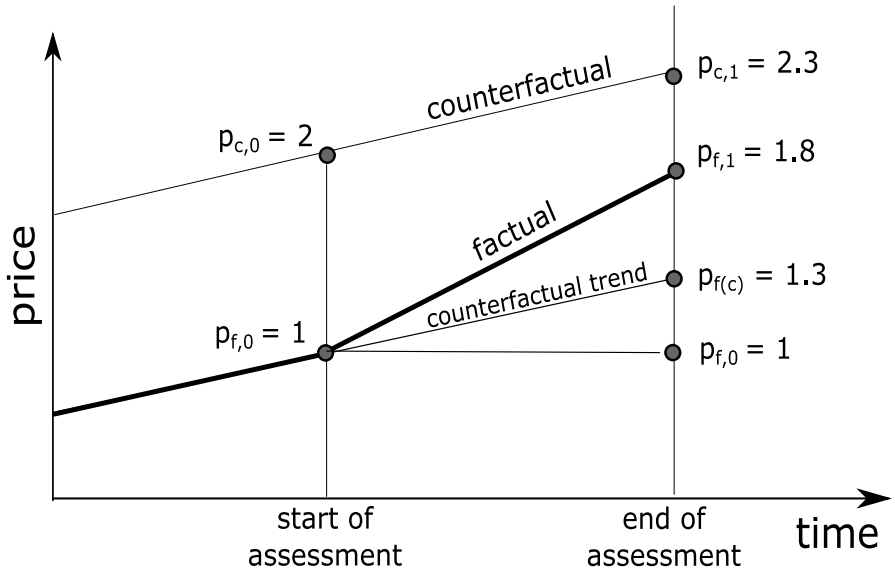
3.3 A gentle introduction to the difference-in-differences method

The simplest forms of causal analyses are before and after price comparisons. In the context of ex post analysis this means comparing prices before, during, and after the behaviour (for example post-merger price compared to the pre-merger price). The main problem with this methods is that the estimated difference is likely to be a biased estimator of the effect of the evaluated behaviour, simply because prices could have been affected by factors other than just the analysed behaviour.

Following the usual notation in ex post evaluations, we refer to one group of individuals (firms, products) as the Factual (also often referred to as the Treatment) group. This is the group that is directly involved in the evaluated behaviour (e.g. members of a cartel, the merging firms, or the dominant abusive firm). Another group of firms or products, as similar as possible to the Factual group, is used as Counterfactual. The main requirement of the Counterfactual group is that should be as similar to the Factual group as possible, except for the analysed behaviour. For example, in ex-post merger studies, the Factual would be the merging firms, and the Counterfactual would be a group of very similar firms that were not involved in the merger. Constructing the Counterfactual is one of the most important parts of impact evaluation. A badly designed Counterfactual could lead to biased estimates and thus misguided conclusions on the effect of competition enforcement.

To illustrate some of this terminology Figure 1 shows how prices change for these two groups during the period of assessment. Assume that the Factual (treatment) group is the firms that were members of a cartel, or the firms that merged. The Counterfactual (control) group is a group of other firms that are very similar to the factual group (i.e. they are affected by the same supply and demand shocks) but for the fact of the cartel or the merger (at least one should assume that they were not affected by the analysed behaviour).

Figure 1: Interpretation of estimators



At the start of the assessed period (e.g. at the time the merger is approved, or at the start of a cartel), the price in the factual and counterfactual groups was 1 and 2 respectively. At the end of the assessment the average price in the Factual group was 1.8, and the price in the Counterfactual was 2.5. A before-after analysis would compare the end-of-assessment Factual price $p_{f,1} = 1.8$, with the pre-assessment Factual price $p_{f,0} = 1$, and would conclude that the analysed behaviour increased prices by $1.8 - 1 = 0.8$. This would be a biased estimate and the slopes of the Factual and the Counterfactual explain why.

As Figure 1 shows, there is a price-increasing trend in the market before the merger and the Counterfactual group suggests that this trend continues after the start of the assessment. If the only difference between the Control and the Treatment is the fact of the merger or cartel, then the

Factual price-trend would be exactly the same as the Counterfactual trend. To take this trend effect out of the estimate, the difference-in-differences method compares the difference across Factual and Counterfactual in the differences between pre-, and post-assessment prices. Therefore the difference-in-differences estimator would be $(p_{c,1} - p_{c,0}) - (p_{f,1} - p_{f,0}) = 0.5$. In Figure 1 $p = 1.3$ shows what the price would have been in the Factual market, in the absence of the analysed behaviour (cartel or merger).

The difference-in-differences method uses variation over group of individuals and over time to mitigate the issue of selection bias. The Counterfactual group is used as benchmark to evaluate what would have been the situation without treatment. The identification relies on the assumption that the Counterfactual group is similar to the Factual group and that the two groups have similar behaviour over time.

3.3.1 Characteristics of the Counterfactual

The choice of the right Counterfactual is key to delivering unbiased impact estimates. Below we highlight three important characteristics of a good Counterfactual.

Parallel trends

This assumption states that the output variable of interest (typically price) follows parallel trends over time for the Factual and Counterfactual groups, i.e. they are affected by the same external demand and supply shocks to the same extent. A violation of this assumption will result in biased price-effect estimates. We can compare trends over time, if we have data on more than two time-periods, by estimating a model with dummy variables for each time period (e.g. year) pre-merger, differentiated for the treatment and control groups.

Adding time periods to estimate trends for the Factual and Counterfactual groups before and after the merger, cartel, or abuse is good practice to validate the hypothesis that the distance between the two groups is constant over time, but the researcher has to be aware that longer time series bring in issues of serial correlation, which may lead to biased standard errors (typically underestimated standard errors).

No spill-over effects

A spill-over effect occurs where the analysed behaviour affects both the Factual and Counterfactual groups. This nibbles away the difference between the Factual and Counterfactual groups, hence underestimating the effect. For example, a cartel typically affects not only its members but non-participating rivals as well. Previous literature suggests that businesses that are rivals to cartel members are also likely to charge a higher price. Similarly, a merger might trigger rival firms to also increase their prices. Because of this, estimates will be downward biased – i.e. they will not reflect the full effect of the cartel or the merger.

The Factual and Counterfactual are sufficiently similar

The only difference between the two groups is the incidence of the analysed behaviour (merger, cartel, or abuse). A violation of this assumption can result in biased price-effect estimates, because the price-effect estimates measure not the impact of the merger, but the difference in price changes between two different markets. There are now widely used formal methods for selecting counterfactuals that adhere to this assumption (propensity score matching, see below).

3.3.2 Choice of Counterfactual

The validity of the difference-in-differences estimates hinges on the choice of the Counterfactual. The two most frequently chosen Counterfactuals in literature are competitors' prices and prices in local markets not affected by the merger.

Competitors' prices

A simple Counterfactual group could compose of the merging/colluding/dominant firms' competitors (i.e. firms in the same market but not involved in or affected by the analysed behaviour). This is of course based on the assumption that the merger/cartel/abuse had no impact on competitors' prices. There is a wide range of theoretical and empirical evidence refuting this assumption.[18] This would imply that using competitors' prices as Counterfactual may provide biased price-effect estimates. For example if a merger resulted in a price increase and it had a spill-over effect on the rivals, the rivals' prices would have also increased (although to a lesser extent[19]). Similarly, cartels often have an impact of the prices of non-cartel-members. Any such estimates will be biased downwards. Nevertheless the price effect estimate can still be interpreted as a lower bound to the unbiased effect.

Local markets

The Counterfactual can also be constructed using local markets, i.e. different geographical markets for the same product but not affected by the analysed behaviour. This method explores the local variation in prices and assumes that there are geographical markets for the same product that have similar characteristics to the Factual market but are unaffected by the analysed behaviour.

There are various methods for selecting the Counterfactual. The simplest of these is when all local markets where data was available are included in the study. More sophisticated methods have been developed for improving the choice of Counterfactual (i.e. to select local markets that are sufficiently similar and unaffected by the merger). Here we briefly introduce two of these methods:

- **Propensity score matching (PSM)**: PSM is a method that allows the selection of *a Counterfactual group of products or markets that is most similar* to the Treatment group. In many applications, especially in health studies, PSM uses variables that predict the reception of treatment. These variables are applied to decide which individuals are equally likely to receive the treatment (i.e. who are the most similar to the Treatment/Factual group). The technical implementation of the method is more involved, which falls beyond the scope of this report, but most statistical software packages include PSM by default.
- **Synthetic control**: Another way to compose an unbiased Counterfactual is by constructing a synthetic control.[20] This is based on the idea that a combination of potential counterfactuals is a more efficient Counterfactual than a single one. This is particularly useful when the number of potential Counterfactuals is small or when it is difficult to find a single Counterfactual that is unaffected by the analysed behaviour. Instead of using a single Counterfactual, the synthetic control method proposes a weighted combination of possible Counterfactual. The weights are derived by using the observable characteristics of the Factual and Counterfactual groups, and

[18] For a discussion in the literature for merger, see Borenstein (1990), Clougherty and Duso (2009), and Ashenfelter et al. (2013).

[19] See for example Deneckere and Davidson (1985, RAND).

[20] See for example Abadie et al (2010).

determining which combination of Counterfactuals gives the closest match to the Factual when all characteristics are considered.

Two additional caveats have to be made about the use of local markets as control. Firstly, local market price variation often does not exist (when prices are set at national level). Second, as Chone et al. (2012) show, even local markets that are not directly affected by the merger might still fall under the merger's umbrella effect, therefore their choice must always be based on careful consideration.

3.3.3 The dataset

Ex-post evaluations are typically data intensive, and often demanding on resources. On the other hand, data availability has largely improved over the past years and also, there are numerous public or commercial data sources that can be used to perform econometric exercises. The other option is for the investigator team to collect their own data. If possible, it is prudent policy to require the parties to the analysed case to provide data for the ex-post evaluation – however, these sort of commitments may be legally difficult to impose/enforce.

The choice of the data to use in the analysis is also crucial. The design of the evaluation framework as well as the specifically adopted methodological tools are heavily influenced by the structure and quality of the data. Low quality data most often only allows low quality analysis, and poor, unreliable results. What makes it even more challenging is that because of the costly nature of compiling a reliable dataset, it is often very difficult to go back and adjust (expand) the dataset if it turns out to be necessary during the analysis.

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).

Data on the dependent variable (price)

Various sources can be used to acquire price data. For example, price data is most widely available from proprietary sources. However, this is often the most costly way to acquire data. Industry regulators could be a free source, as they typically disclose industry-relevant price data, but of course availability is limited to the regulated industry. Data could also be collected by the team conducting the ex-post evaluation from various websites. Automated data mining and web scraping techniques have now become widely used, mainly because of their flexibility and inexpensive nature.

Data on independent variables

Below we provide a list of the typical independent variables that are used in models estimating the impact of competition cases.

Cost variables: The obvious candidate for this purpose is information on the cost of the relevant product. A typical approach is to include information of the price of one or more input used in the

production of the relevant product. If the merger happens in the retail market then the model could include wholesale prices as a control for costs.

Firm characteristics: This could be the firms in the Factual or in the Counterfactual groups. Sometimes it may be relatively simple to find information on the characteristics of these firms (firm size, market share, etc.). One way to control for some (the time-invariant part) of the unobserved heterogeneity is by including a simple set of fixed effect variables (dummy variables for each individual).

Product characteristics: In the case of differentiated products it is advisable to control for product characteristics. Fixed effect dummies can also be included to control for some of the unobserved product heterogeneity.

Market characteristics: This is especially useful if the data exploits cross-market price variation. Again, some market characteristics are relatively easy to observe (market structure, size of the market – population, etc.). Fixed effect dummies can also be included to control for some of the unobserved market heterogeneity.

3.3.4 The exogeneity of the behaviour

A central assumption to the difference-in-differences method is that the analysed event is exogenous, i.e. the incidence of the merger, cartel, or abuse of dominance is not correlated with unobserved factors that also affect price. This assumption is violated if there were 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.

In practice it is not trivial how to deal with such endogeneity, and much of the previous literature simply assumes 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, we can set up three general rules of thumb:

1. Try and identify what triggered the merger or the cartel and think how this would have affected the prices. Controlling for these causes might help eliminate the omitted variable bias. 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.
2. Include a set of fixed effect dummies for each individual/geographical area/etc to eliminate some of the unobserved heterogeneity.
3. In exceptional cases there might be suitable instruments. This may vary from case to case but intuitively we think that in the case of large diversified mergers a possible set of instruments could be gained from characteristics of parts of the merging firm that are not directly involved in the merger.

3.3.5 Robustness checks

The above discussion shows that the researcher often has a large amount of discretion when designing ex-post evaluation studies. Decisions have to be made on the choice of Counterfactual, the time period used in the estimates, the data used, and the exact model specification. To demonstrate that the findings of the study are robust to these choices, it is advisable to do some 'robustness checks', and re-run the estimation using a different Counterfactual, a different span of data, or a different model specification. The three most widely used robustness checks are:

- *Placebo treatment*: In this case the model is re-estimated using a different (false) date of the analysed event (for example a false date of cartel formation). Placebo treatments are typically used in DiD models in order to show that the effect estimated is not spurious. The interpretation is typically simple, if the placebo treatment returns equally significant (and similar in magnitude) estimates for the effect of treatment, then one should be suspicious that the initial estimate was biased by issues such as autocorrelation.
- *Changing the Counterfactual*: Another way to test the robustness of results is by changing the Counterfactual to an alternative – often equally suitable – counterfactual. If the results remain in the same ballpark as the original estimate even with different controls, it would be interpreted as confirmation of the robustness of the original estimate.
- *Changing model specifications*: The goal of this exercise is typically to demonstrate that the results are not sensitive to the exact choice of model specification. In practice this most often means running the estimation with a different set of dependent variables.

3.3.6 How to interpret the findings?

One reason why some authorities might be reluctant to conduct ex-post evaluations is because of the concern: what if it shows that they made an error. We need to emphasise that this is a badly misguided apprehension. Finding for example that prices increase post-merger 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. It is precisely for these reasons that ex-post studies are invaluable, and why they should not stop at just estimating price-effects. The authority is in the best position to take the estimates from the ex-post evaluation and re-visit the decision to find out 'what went wrong'.

3.4 Evaluating the impact of a merger decision

The ex-post evaluation of merger decisions is especially important given that merger control is an ex-ante policy instrument – i.e. the competition authority has to assess the market effect of a transaction before the transaction takes place. In this respect even if it is based on the best available evidence, there is an inevitable uncertainty in the suitability of any intervention that follows. Because of this, it is crucial that merger control decisions are subjected to rigorous ex-post evaluation exercises in order to assess how well they are achieving their task of filtering out anti-competitive mergers. If done correctly, lessons learned from these ex-post studies can be channelled back into policymaking to affect how mergers are assessed in the future.

To improve the predictive power of ex-post merger studies, it is important to tailor the ex-post study to the CA's decision. If the CA unconditionally approves a merger it may be that it expected market

forces to self-correct for the increase in market concentration. As these mechanisms may take longer than a year to get in full swing, it is advisable in these cases to look at price data that spans longer than 1 year following the merger. A dataset that does not span longer than a year after the merger may only pick up short-term price increases (and would imply that the CA's decision was wrong). On the other hand, longer spanning data would be more able to verify if market self-correction did take place.

3.4.1 Methods used

In the general section we highlighted the difference-in-differences method as the most widely accepted method for evaluating the impact of competition decisions. Below, we briefly mention two other methods that are often seen in the ex-post evaluation of mergers: event studies, and merger simulations. Technically speaking, neither of these methods are ex-post. Event studies evaluate how the stock market reacted to news about a merger (i.e. how the stock market expected the merger to change the valuation of a given firm). Merger simulations typically use ex-ante data to estimate the parameters of the relevant demand system following the merger.

Merger simulations

By simulation, we refer to evaluation based on the following stages: (i) An explicit formal modelling of the nature of competition in the market (the nature of oligopoly, homogenous or differentiated products, existence of capacity constraints, unilateral or coordinated action, market symmetry, etc.). This stage often involves a structural model derived from a game theoretic perspective, coupled with a particular model of the demand system, e.g. logit, nested logit or random utility. (ii) Calibrating the model with real world information derived from direct observation, or from full-blown econometric estimation of the demand system to acquire estimates of model parameters (e.g. existing market shares, prices and extraneous estimates of demand elasticities). Finally, (iii) using it to assess how the equilibrium will change with and without an event/intervention (for example comparing the pre-merger equilibrium with the hypothetical non-intervened equilibrium).

Simulation may be either ex-ante or ex-post counterfactual analysis. The latter is backward-looking – what outcome would have happened, had, say, a cartel not actually existed; while ex-ante looks forward – anticipating whether or not, say, a merger would have had coordinated effects, if cleared. When the perspective is ex-ante, but conducted after the event, a decision must be made as to whether the counterfactual estimates should draw on all information available, or merely the information that would have been available at the time of the policy decision. Of course, the latter requires a reworking of the original model, now simulating the impact of that shock, as well as the original intervention.

The major strength of the simulation approach is the explicit use of theory to identify the counterfactual. This facilitates the 'joining-up' of the analysis undertaken at the time of the intervention with any subsequent evaluation of the effects of the policy, and, in turn, provides a clear opportunity for evaluating the assumptions made at the time of the intervention.

However, as is well documented, simulation is very sensitive to modelling assumptions. Sometimes this is a strength, e.g. in revealing how sensitive predictions are to the precise nature of the counterfactual, but sometimes this sensitivity is unhelpful, deriving from alternative specifications between which there are no strong theoretical reasons to choose, e.g. the functional form of the demand curve. Equally important, simulation is better suited for some types of oligopoly models (and therefore markets) than others. The trusted and well-tried workhorses are the Cournot homogeneous product model, and logit type models of product differentiation. Invariably, the

emphasis is on price and quantity to the exclusion of innovation, repositioning etc, and possible changes in conduct (relevant to coordinated effects). Buyer power has also proved difficult to incorporate satisfactorily, and simulation of bidding markets is still in its relative infancy. This raises the strong likelihood that evaluation based on simulation is heavily skewed towards certain types of markets, potentially leading to a sample selection bias.

Another potential source of selection bias derives from the heavy demands on data. Many of the seminal studies are based on high quality disaggregated datasets constructed from scanner sources; but, of course, these are typically drawn from a relatively small set of consumer good products (often sold through supermarkets). Finally, there is mixed evidence on how well simulation predicts actual outcomes.

Event studies

An event study draws on financial market data to measure the effect of an economic event on the market valuation of a firm. If financial markets are efficient, then the effect of any event on a firm's discounted profits will be instantaneously observable through the changes in the prices of its shares. The methodology entails measuring any abnormal returns associated with an event (e.g. the announcement of a merger. Abnormal returns are identified as the difference between the observed movement in stock valuation and those that would have occurred absent the event.

Central to the event study is the assumed rationality of markets, the efficient market hypothesis (EMH), according to which share-prices instantly reflect the value to investors of all the relevant information available to them. It builds upon information that is generated by the interaction between a large number of self-interested, independent, rational market agents. This information then can be thought of as the best estimate, given the set of all available information. If the EMH holds, then the change in the market's valuation of a company will always reflect an unbiased estimate which is both "objective" and quick. This therefore enables a quicker assessment than from using more direct measures such as product prices. The methodology may be particularly attractive to CAs as it tackles the issue of information asymmetry between them and the firms involved in the event. This makes event studies more appealing than the analysis of accounting data, which typically suffers from the potential bias that such information is produced by the interested parties. It is also argued that event studies are undemanding of data – the necessary data are easily accessible for listed firms (but see below).

However, the plausibility of the EMH assumption is open to question, and many commentators are sceptical. Werden (2008) suggests that the presumption that 'the instant analysis of uninformed investors is more accurate than the painstaking work of enforcement agencies with access to confidential documents and data' is not supported by evidence. Another weakness of the event study method is that it tells us very little about what actually happened after the merger. Instead it focuses on the expectations of the market. For this reason it is very unlikely that they can be useful in the assessment of individual cases.

Finally, in spite of the general presumption that event studies are easy-to-use and data are easily accessed from financial databases, the practical reality is that there are many circumstances when appropriate data are not available. It is, of course, a necessary condition that the parties and their rivals are all quoted on the stock market, but this is often not the case in some markets especially where firms are small and rivals are scarce. Moreover, very often the parties are large conglomerates and/or multinational, and the market concerned may constitute only a small part of

its aggregate activities; where this is the case, it can often prove difficult to identify any effect on the firm's valuation resulting from an event in a small market in a particular country.

3.4.2 Taking a dynamic approach

The time-range of data used for ex-post evaluations is also important when assessing how well these studies evaluate the merger decision. It is possible that the immediate price-shock, caused by the merger, is self-corrected by the market within few years following the merger. It is also possible that in the immediate aftermath of the merger prices increase only in markets directly affected by the merger, but then spreads on to other markets later on. Both of these possibilities have to be accounted for when designing the ex-post study.

It seems unlikely that a CA would choose to intervene if it believed that market correction would offset the price increase within one or two years after the merger. It is also unlikely that the CA would refrain from intervention if it judged that localised effects of the merger would later spread on to other geographical areas. For these reasons it is vital that the ex post study uses data that spans sufficiently far following the consummation of the merger. On the other hand, longer spanning data is more likely to contain confounding effects (i.e. effects other than the merger).

In choosing the appropriate time-span of the data one should take into account what the CA's decision was. For example if the CA predicted that the initial price increase would be neutralised by subsequent entry then the data should span sufficiently long to pick up these effects. If data is available for a longer period post-merger, it should be relatively straightforward to check how the mean post-merger price changes as the data span is increasing.

3.5 Evaluating the impact of cartel decisions

Competition authorities have been less active in preparing detailed ex-post evaluations of the impact of specific cartel decisions. One of the main reasons is that cartels are unanimously considered to be harmful and therefore any decision that eliminates a cartel is considered consumer welfare enhancing. For this reason, the ex-post evaluation of cartels is not about deciding whether the intervention was justified, rather to estimate how much harm the cartel caused. From this, the ex-post study can extrapolate to express the amount of future harm avoided by the intervention. Estimating cartel harm is of course is the main question in actions for compensation, and as such is extensively looked at by private economic consultancies. For this reason the literature on estimating cartel damages provides useful assistance.[21]

3.5.1 Methods used

A well designed causal inference method (like difference-in-differences) should be preferred choice to estimate the price increase caused by the cartel, which is then used to conjecture about the price increase avoided by the CA stopping the cartel. However, other methods can be found in previous studies.

[21] Commission Staff Working Document – Practical Guide 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, SWD(2013) 205, 11.6.2013, http://ec.europa.eu/competition/antitrust/actionsdamages/quantification_guide_en.pdf, and Quantifying antitrust damages: Towards non-binding guidance for courts, Study prepared for the European Commission, http://ec.europa.eu/competition/antitrust/actionsdamages/quantification_study.pdf

For example the Mexican competition authority evaluated the impact of a collusive practice in the freight truck transportation market in Mexico.[22] The study looked at the price effect of the Fuel Price Adjustment Fee (FPAF), through which the National Chamber of Freight Truck Transportation (CANACAR, by its acronym in Spanish) and its members colluded to transfer increases in fuel prices directly to customers. Different from the other evaluations this study did not use a difference-in-differences method. Instead it estimated whether there was a structural break in the price data around the cartel. For this a simple Chow test was used. The Chow test can provide evidence that the periods before/during/after the cartel are significantly different – i.e. if there is a structural break in the data. To calculate the overcharge, the study then estimates a counterfactual (non-collusive) price using the same model specification. The overcharge is simply the difference between the non-collusive and collusive prices, divided by the non-collusive price.

3.5.2 Choosing the Counterfactual

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.

Whilst in merger ex-post evaluations the focus is on the post-merger price, in cartel evaluations it is on the price during the cartel. The CA's decision identifies the duration of the cartel, which should help in distinguishing the periods before, during, and after the cartel. The identification task is therefore to test the difference between the prices in the Factual and the Counterfactual during the cartel.

Two main types of Counterfactuals 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 pre-, or post-cartel prices are often used without answering these questions. There is a growing body 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. 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.

[22] https://www.oecd.org/daf/competition/cofece-assessment-freighttruck-april2015.pdf

3.6 Evaluating the impact of an abuse of dominance decision

Previous studies mostly avoided the ex-post evaluation of abuse of dominance cases. There are two obvious explanations for this. Firstly, competition authorities bring relatively few abuse of dominance cases.[23] However, the ones that are investigated and intervened tend to be very high profile (e.g. Microsoft, or Intel cases in the EU), which should warrant the ex-post evaluation of the impact of these interventions. Second, evaluation is unusually difficult to conduct in these cases. Exclusionary behaviour is likely to have a wide range of pro-, and anti-competitive effects. Finding, and quantifying the magnitude of the net negative effect, and establishing the extent to which rivals are harmed and the impact on consumers is not a trivial task. Moreover, exclusionary abuse is likely to have its effect long after the behaviour itself. Take the example of predatory behaviour, where one would need to quantify short-run and long-run impacts of opposite directions.

Otherwise, the treatment of abuse of dominance cases in an ex-post study should be very similar (methodologically) to cartels. The evaluation would aim at estimating the price-increase caused by the abuse and assume that the CA intervention eliminated this price increase.

Choosing the counterfactual also carries similar challenges as cartel evaluations: for example prices after the abuse stopped might not reflect the competitive price. Moreover, competitors are less likely to be a good choice for Counterfactual, because by definition we are talking about a dominant firm, therefore any competitor will be characteristically very different (smaller). Similarly large firms (or the same dominant firm) in other geographical markets where there is no evidence of abuse might be a better Counterfactual.

3.7 Conclusion on ex-post evaluations

To summarise, we highlight a few key rules of thumb:

- 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 a decision error.
- 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.

[23] It is fairly common in new competition regimes to address excessive prices under abuse of dominance rules. In our discussion we refer to abuse of dominance as exclusionary abuse (i.e. abuse that leads to the foreclosure of a rival).

4 Evaluating the distributional effects of competition enforcement

4.1.1 Literature Review

Recently, the question of how competition policy can reduce inequality has attracted increased attention from academics and policymakers alike. This is not surprising, since the Great Recession, the question of growing inequality increasingly dominated the political debate. It was only a matter of time that these issues were picked up by competition academics and professionals. As a result we have seen some valuable papers on the topic, one of which is a 2016 working paper by Jonathan Baker and Steven Salop, who have provided a list of suggestions on how to strengthen competition enforcement to further the objective of increased equality.[24] These suggestions are ambitious but the paper remains sober and highlights some important controversies. Much of this literature is focused on how market power can contribute to inequality and accordingly how competition policy can foster wealth redistribution by eliminating market power. This argument is based on the premise that people in the upper tail of income distribution are more likely to own business assets, and therefore supernormal rent from market power is more likely to benefit the wealthiest. On the other hand, the lower tail of the income distribution pay higher prices resulting from market power but they are unlikely to benefit from these supernormal business profits.

To provide some empirical backing to these ideas, we consider two recent working papers, produced by the European Commission[25] and the OECD[26], on how competition enforcement can contribute to greater income equality.

The European Commission uses a macroeconomic Dynamic Stochastic General Equilibrium (DSGE) model and simulations to identify whether increasing mark-ups affect the wealthier (non-liquidity constrained) households more than the less wealthy (liquidity constrained) households. The model is a fairly complex one and for the purposes of this paper we do not go into details on it. However, the intuition is simple: all households bear the cost of a mark-up increase but only the richer non-liquidity constrained households benefit from the greater profits since they own most of the assets, i.e. stakes in the firms with market power. To simulate the model the paper uses parameter values which are consistent with empirical observations, and the simulations assume that the Commission's actions eliminate some of the price mark-ups. The magnitude of this effect is derived from previous estimates of the Commission on the consumer benefits of its enforcement actions. The simulations show that the increase in consumption (as a result of eliminating anti-competitive mark-ups) is proportionately more for the poorer, liquidity constrained, households than for non-liquidity constrained households: the liquidity-constrained households increase their consumption four times more than the non-liquidity constrained households after 5 years.

The OECD paper uses a different approach, but is still based on similar intuition, namely that the rich, who are more likely to be shareholders of businesses with market power, are more likely to benefit from this market power than the poor. The paper looks at 8 countries and compares measures of market power with measures of inequality (the share of the wealth of the top10

[24] Baker, J. B., & Salop, S. C. (2015). Antitrust, Competition Policy, and Inequality. Geo. LJ Online, 104, 1.

[25] Adriaan Dierx, Jukka Heikkonen, Fabienne Ilzkovitz, Beatrice Pataracchia, Anna Thum-Thysen and Janos Varga (2015) Distributional macroeconomic effects of EU competition policy – a general equilibrium analysis, http://www.cresse.info/uploadfiles/2015_pa14_p1.pdf

[26] OECD (2015) Market power and wealth distribution, DAF/COMP(2015)10, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP(2015)10&docLanguage=En

richest). It reports that market power might account for between 10-24% of the income of the richest decile of the population, therefore eliminating this would reduce wealth inequalities.

In a recent review of relevant literature, Begazo and Nyman (2015),[27] adds an extra dimension. They focus on how anti-competitive practices affect the bottom 40% of the population, but allow households to have more than one economic function: they are employees, and sometimes producers, as well as consumers. This adds some extra insights, for example a price fall due to the elimination of anti-competitive practices (or trade barriers) would benefit households as consumers but if lower prices mean lower employment then it could have negative effects on them as employees and (if applicable) producers. Their review shows that that tackling anticompetitive behaviour (and reducing government restrictions on competition) can have a positive distributional impact by lowering consumer prices, whilst also contributing to raising returns to small producers. More competitive markets can also bolster job growth and improve income distribution.

An alternative literature, more micro based, focusses more on a case by case by approach, in which markets are looked at individually. For example, there is a growing body of literature on how demand elasticities differ depending on household income. In a large meta-study on the effect of food prices on food consumption (reviewing 136 studies reporting 3495 own price food elasticities from 162 different countries) Green et al. (2013) found that demand for food was more responsive to price changes among households with lower incomes.[28] On the other hand, focusing on a variety of products in single country (Mexico), Urzua (2013) comes up with more nuanced results.[29] He estimates a structural demand model to demonstrate how the welfare loss resulting from market power varies across the income distribution. The author finds that that in the urban sector the negative impact of monopoly power grows (in relative terms) as households become poorer. In the limit, the poorest households have a relative welfare loss about 19.8 per cent higher than the one suffered by the richest. For the rural sector the redistributive impact is even more serious, since the first decile has a relative welfare loss of about 26.4 per cent compared to the ninth decile, and of 22.7 per cent compared to the tenth decile. This difference can be explained by the difference in elasticities across the various demographic groups. He estimates price elasticities of demand for poor and rural households' demand and finds that for some products (corn tortilla, processed meat, medication) demand is less responsive to a price increase in poorer households. One explanation why these estimates might differ from the Green et al. meta-study is that these are products that are more likely to be a necessity for poor households. Therefore in general it appears to be true that wealthier households are less sensitive to a price increase. However, when it comes to necessities, a price increase leaves the poor with little choice, especially if the nearest substitute is a more expensive product.

4.1.2 Proposed Methodology

We draw two main conclusions from this literature. First, at the aggregate level, competition policy does indeed have the potential to reduce income inequality, and the intuition is obvious: in the main, the poor lose relatively most from market power, whilst the rich may on balance benefit from market power because any losses they incur as consumers are balanced by the fact that they are the

[27] Begazo, T., & Nyman, S. (2016). Competition and Poverty.

http://documents.worldbank.org/curated/en/2016/04/26211459/competition-poverty

[28] Green, R., Cornelsen, L., Dangour, A. D., Turner, R., Shankar, B., Mazzocchi, M., & Smith, R. D. (2013). The effect of rising food prices on food consumption: systematic review with meta-regression. Bmj, 346, f3703.

[29] Urzúa, C. M. (2013). Distributive and regional effects of monopoly power. Economía Mexicana Nueva Época, 22(2), 279-295.

main beneficiaries of monopoly profits. This general effect is reinforced because the poor are all employees or run small family businesses. Second, at a more micro level, the beneficial effects may vary systematically across different markets, depending on the income and price elasticities.

In our opinion, this literature remains seriously under-developed. All of the papers reviewed above tackle a dauntingly difficult empirical question, and all have made important contributions. However, they simplify in being based, either explicitly or implicitly, on static models, which tend to ignore any potential dynamic gain from the profits generated by temporary market power. The aggregate studies assume that all profit is paid out as dividends to shareholders (or is spent on rent-seeking). Moreover, the EC model, which is probably potentially the most directly relevant for our purposes is too technically complex and data-intensive to be practicable in a developing country context.

With these serious qualifications, we turn to the issue that UNCTAD have asked us to consider: is there a methodology for evaluating how a competition authority's enforcement activities have helped reduced inequality?

The first part of our answer is that there is no such methodology in the literature at the time of writing. The second part of our answer is to propose a 'rule of thumb' methodology which is in the spirit of most of the previous literature, but would be relatively easy to apply in practice. It would enable us to quantify the impact of competition policy in aggregate and identify the specific markets in which interventions would have the greatest positive impact on inequality.

Assumptions

1. Household welfare will be measured by consumer and producer surpluses. This is the general convention in the competition policy literature. Thus, households receive 'utility' in two forms: consumer surplus from consumption and producer surplus (profits) from owning firms.
2. The economy includes two classes of household by income: RICH and POOR.
3. RICH own a disproportionate share of firm stocks and shares, and therefore receive a disproportionately large share of producer surplus – both directly in the form of dividends and indirectly in the form of increases in firms' present value valuations.
4. On average, consumer products have an income elasticity of less than unity. As such, POOR account for a disproportionately large share of consumption.

Assumptions 1 and 2 are made merely for simplicity of exposition. It would be easy to generalise to three economic functions as in Begazo and Nyman, namely consumers, producers, and employees, and to n>2 income classes. This would not change the thrust of our analysis. Assumptions 3 and 4 are easily verifiable empirically.

In algebraic terms, let POOR account for λ% of aggregate consumption and receive θ% of profits, while the corresponding percentages for RICH are $(1-\lambda)$ and $(1-\theta)$ respectively. From assumptions (3) and (4), we know $\lambda>\theta$.

Now consider how to measure the impact of competition policy in a given period, say a year. Suppose that we estimate that interventions by the CA increase consumer surplus by ΔCS whilst reducing profits by $-\Delta PS$. Assuming that POOR's share of the increased surplus is the same as its share of all CS and ditto for its share of PS, then it follows that POOR gains more than RICH, i.e. POOR's share of utility increases if:

$$\lambda\Delta CS - \theta\Delta PS > (1-\lambda)\Delta CS - (1-\theta)\Delta PS$$

that is:

$$(2\lambda-1)\Delta CS > (2\theta-1))\Delta PS$$

Two key conclusions follow immediately:

I **Assuming that competition interventions satisfy the law in that they improve consumer welfare, ($\Delta CS>0$), then policy will always improve equality if $\lambda>0.5$ and $\theta<0.5$.** Note that these are only sufficient, but not necessary, conditions. Nevertheless, they will nearly always be satisfied, assuming that the POOR account for the majority of consumption, but receive hardly any profits. Indeed, in most cases θ will be close to zero, and equality will improve so long as the POOR account for at least half of consumption.

II **Competition Authority intervention will have the greatest impact on inequality in those industries supplying necessities.** A necessity is a product with a low income elasticity, and where this is so, the POOR will account for a much larger share of consumption than of income.

4.1.3 Conclusions

The main message then is that competition policy is likely to have a positive impact in reducing inequality, and that this impact will be most beneficial for necessity products.

Given the uncertainties in measurement, we would caution against quantifying these effects. However, if this is considered to be desirable, then a rough quantification would be possible if one is able to estimate the impact of an intervention on consumer and producer surplus, and to make a rough estimate of the share of consumption of the POOR. Assuming POOR is defined as the 40% poorest households in the population, their share of consumption would be at least 40%, often much more, except for luxury goods (with income elasticity >1).

5 Evaluation for accountability: A case study of Tunisia

In describing our methodology, we were asked to refer to a practical example. UNCTAD originally asked that the example should be Kenya but, for practical reasons, later changed this to Tunisia.

After a period of internal investigation within UNCTAD, it was suggested that we might use the calendar year 2012. However, UNCTAD personnel subsequently found that the case was deficient in two respects: (i) there were absolutely no data for mergers, and (ii) only very brief qualitative data were available for cartels and abuse of dominance.

Without these data, even the most rudimentary of evaluations is impossible. As a second best therefore we use the small amount of information that UNCTAD was able to provide on the caseload of the Tunisian Authority in 2012. This is summarised in Table 7 and allows us to discuss what would have been possible if more data were available:

Mergers: Although UNCTAD was unable to provide us with any information on mergers, we presume that there were a number notified to the CA and that, for some, after investigation, the CA will have either blocked or remedied the proposal. In those cases, impact could be estimated given information on turnover in the merger markets, and either case-team estimates of price increase averted or rules of thumb defaults of 3%.

Advocacy: No information is available on advocacy so we cannot offer any comment on this.

Competition cases: Here, we do have some information on the Authority's activities on **Cartels and Abuse of Dominance.** As can be seen from Table 7, there were 63 candidate cases, but of these 46 led to no impact, either because the cases turned out to be outside the Authority's jurisdiction, or because they were rejected due to defects in the case – either substantive or in form.

This leaves 17 cases where an impact can be claimed for the Authority. Most obviously, (i) there are the 6 cases where the Authority made an adverse finding; but also relevant are the (ii) 9 cases where the firms decided to abandon their practice – presumably because they feared an adverse finding; and (iii) 2 other cases where the Authority made a precautionary judgement which influenced the firms concerned.

Table 7 Tunisian Competition Authority's Case load 2012

	Number of cases	Impact?
Merger references	Not known	
Of which: Merger Investigations	Not known	
Advocacy	Not known	
Competition Cases	63	
Of which:		
Adverse findings	6	Impact
Abandoned cases	9	Impact
Precautionary	2	Impact
Lack of jurisdiction	14	None
Rejected for Substantive Defect	26	None
Rejected for Defect in Form	6	none

UNCTAD was also able to supply some information on the nature of the 6 adverse findings. Two of these (see Table 8) were effectively cartels and two were abuse; from the little information we have

seen, the other two probably straddled the two types of anti-competitive activity. Unfortunately, there was no information in the files on either the turnover of the firms, or of the likely beneficial effects on price of the Authority's interventions. Sadly therefore no estimate of savings can be made.

Table 8 Competition Cases: Adverse Findings

Abuse of economic dependence	2
Parallel pricing	1
Hard-core cartel	1
Cover up bids	2
Total	6

Likewise, we do not really have sufficient information on any of these six cases to identify which might be suitable for an ex-post evaluation, although we suspect that one or more might be good candidates.

An Hypothetical example

Obviously it is disappointing that UNCTAD was unable to provide us with the quantitative data needed to conduct a real world illustration. In the circumstances, the best we can do is to present an hypothetical example: what an evaluation *might* look like for Tunisia. We stress, however that this is hypothetical and the assumptions on which it are based are totally arbitrary and do not benefit from any more knowledge on Tunisia than we have presented above.

Suppose in a given year,

(i) The Authority blocked or remedied six proposed mergers, and prohibited six cartels or abuses of dominance.
(ii) The turnovers of the parties were, say, 5 million Tunisian dinars (2.5 million euros in 2012 exchange rates) in each of the 12 cases
(iii) The beneficial price effects of these interventions are assumed to be the defaults shown earlier in Table 6: 3% for mergers and 10% for cartels and abuses.
(iv) The beneficial effects are assumed to last for 3 years.

Given these assumptions, the CA's impact was

$$6*5*0.03*3+6*5*0.10*3 = 11.7$$

Discounting the savings in the 2nd and 3rd years, and assuming a discount rate of say 3%, the stream of impacts in present value terms are:

$$3.9+3.8+3.7 = 11.4$$

We have no information or knowledge about the Tunisian competition authority, but if its annual costs were 1 million dinars, this would mean a benefits/costs ratio roughly equivalent to UK's CMA (see Table 4 above.)

6 References

Ashenfelter, Orley and Daniel Hosken (2010) 'The effect of mergers on consumer prices: evidence from five selected case studies', Journal of Law and Economics, vol. 53, issue 3, pages 417 – 466.

Block, Michael Kent, Frederick Carl Nold and Joseph Gregory Sidak (1981) 'The deterrent effect of antitrust enforcement', Journal of Political Economy, 89, 3, 429-445.

Bolotova, Yuliya and John Connor (2006) 'Cartel Overcharges: Survey and Meta-Analysis', International Journal of Industrial Organization, 24, 1109–1137.

Borenstein, S. (1990) 'Airline Mergers, Airport Dominance, and Market Power', American Economic Review (Papers and Proceedings), 80, 400-404.

Bundeskartellamt (2010), "Effective cartel prosecution: benefits for the economy and consumers"

CMA (2015) "Positive Impact 14/15 - Consumer benefits from the OFT's work", CMA47

Competition Commission (2011) "Estimated benefits to consumers from the CC's actions in mergers and market investigations between April 2009 and March 2010"

Davies, Stephen W (2012), "A survey of the approaches and methods used to assess the economic effects of a Competition Authority's work", Swedish Competition Authority

Davies, Stephen W. and Bruce Lyons (2007) 'Mergers and Merger Remedies in the EU: Assessing the Consequences for Competition', Cheltenham, Edward Elgar.

Davies, Stephen W (2012), "Ex-post evaluation of cases of Abuse of Dominance" and "Deterrence: some thoughts on the deterrence multiplier", presentations at the OECD, October 2012

European Commission DGCOMP (2012), "Competition Management Plan", European Commission

Levenstein, M. and V. Suslow (2006) 'What Determines Cartel Success?', Journal of Economic Literature, 44, 43–95.

NMa (2010), "Outcome of NMa activities: a description of the calculation method" (Kemp R, Mulder M, and van Sinderen J), NMa Working Paper No. 1

NMa (2012), "NMa Annual Report 2012"

OECD (2012), "Evaluation of competition enforcement and advocacy activities: The results of an OECD survey", December 2012, DAF/COMP/WP2(2012)7/REV1

Ormosi P., (2012), "Evaluating the Impact of Competition Law Enforcement", Paper presented at OECD…. Stocktaking on evaluation, June 2012, DAF/COMP/WP2(2012)5

Smuda F., (2012), "Cartel Overcharges and the Deterrent Effect of EU Competition Law", ZEW Discussion Paper No. 12-050

US Department of Justice Antitrust Division (2012), "Congressional Submission FY 2013 Performance Budget"

US Federal Trade Commission (2012), "Performance and Accountability Report, Fiscal Year 2012"

Werden, Gregory J. (2008) 'Assessing the Effects of Antitrust Enforcement in the United States', De Economist, 156, 4, 433–451.

Werden, G., A. Joskow and R. Johnson (1991) 'The Effects of Mergers on Price and Output: Two Case Studies from the Airline Industry', Managerial and Decision Economics, 12, 341–352.