A survey of the approaches and methods used to assess the economic effects of a Competition Authority's work

KONKURRENSVERKET
Swedish Competition Authority

Förord

I Konkurrensverkets uppdrag ingår att främja forskning på konkurrens- och upphandlingsområdet.

Konkurrensverket har gett professor Stephen Davies vid University of East Anglia, i uppdrag att redovisa för- och nackdelar med olika metoder och angreppssätt för att utvärdera effekterna av konkurrensmyndigheters arbete.

Till projektet har knutits en referensgrupp. Den har bestått av Sten Nyberg (Stockholms universitet) samt Arvid Fredenberg, Lena Fredriksson och Erik Hegelund från Konkurrensverket.

Det är författaren som svarar för alla slutsatser och bedömningar i rapporten.

Stockholm, maj 2012

Dan Sjöblom
Generaldirektör

Contents

Sammanfattning

Inledning

Denna rapport ger en översikt över de tillvägagångssätt och särskilda metoder som används för att bedöma de ekonomiska effekterna av en konkurrensmyndighets arbete, med särskild betoning på metodernas fördelar och nackdelar.

Syften med utvärdering

Utvärdering genomförs av konkurrensmyndigheter (CA) och/eller oberoende akademiker/konsulter. Utvärdering kan anta många olika former, som kan delas in i fyra kategorier

- Ansvar: det finns en skyldighet att kvantifiera konkurrenspolitikens totala fördelar, som kan uppskattas i förhållande till något på förhand angivet mål.

- Uppskattning av ett specifikt politikområde eller ett särskilt ingripande i syfte att kontrollera kvaliteten på konkurrensmyndighetens beslutsfattande; detta kan användas för att bidra till att göra framtida interna prioriteringar och resursfördelning.

- Uppskattning i den akademiska litteraturen har ofta bidragit vid teknisk utveckling av utvärderingsmetoder som används av konkurrensmyndigheter.

- Att värdera skadestånd och böter kräver ofta en utvärdering av den skada som orsakats av de berörda företagen, och därför vinsterna med ingripanden från konkurrensmyndigheter.

Utvärderingsmetoder

Denna rapport fokuserar på tre kvantitativa huvudmetoder:

- Simulering, baserad på ekonometrisk modellering i oligopolteori
- Eventstudier, med bevismaterial från kapitalmarknaden
- Skillnad-i-skillnader, baserat på statistisk analys av före och efter

Vi diskuterar också andra metoder mer kortfattat, inklusive: aggregerade internationella ekonometriska jämförelser; uppföljningsundersökningar av parterna/konkurrenterna/kunderna/leverantörerna; expertkommentarer och domstolsbeslut; och undersökningar av kollegers/praktikers åsikter.

Ex-ante versus ex-post

Ex post-utvärdering genomförs retrospektivt med användning av information om 'vad som faktiskt hände sedan', det vill säga efter att ingripandet ägde rum. För att utvärdera effekten, jämförs sedan 'vad händer sedan' med vad som skulle ha hänt, om ingripandet inte hade ägt rum. Ex-ante-utvärdering kan genomföras antingen före eller efter händelsen, men använder endast information som skulle ha varit tillgänglig vid tiden för policybeslutet. I det fallet projicerar utvärderaren framåt och jämför vad som skulle hända med och utan ingripandet.

Politikområden

I princip kan utvärdering tillämpas på alla konkurrenspolitikens områden, men i praktiken har fusioner och karteller tilldragit sig mest uppmärksamhet. Enligt författarens åsikt krävs mer forskning om utvärdering av andra områden, i synnerhet artikel 102 (missbruk av dominerande ställning) och advokatverksamhet.

Det kontrafaktiska fallet

En huvudfråga i alla utvärderingar är att identifiera ett lämpligt kontrafaktiskt fall - vad skulle hända om ett ingripande inte skett. Detta tilldrar sig ibland otillräcklig uppmärksamhet i tidigare studier.

Utvärderingar för ansvarighet

Första delen av rapporten (kapitel 2-5) ägnas åt utvärderingar av en konkurrensmyndighets aktiviteters allmänna effekter på konsumenternas välfärd. Den fokuserar på de rapporterade metoderna hos de myndigheter där sådant arbete verkar mest avancerat, åtminstone enligt vad som kan bedömas av allmänt tillgängliga publikationer.

Storbritanniens Office of Fair Trading (OFT)

Varje år publicerar OFT en rapport över positiva effekter, som kvantifierar värdet av de välfärdsvinster som OFT:s ingripanden ger, jämfört med kostnaderna för dem för den brittiske skattebetalaren. De utmärkande dragen i OFT:s metod är att:

- Uppskattningar uttrycks i termer av fördelar för konsumenter
- Uppskattningar som regel är ex-ante, men med några ex post-fall
- Uppskattningar är avsiktligt 'försiktiga', och presenteras som 'punkt'-uppskattningar, tillsammans med en rad rimliga världen
- De flesta, men inte alla, politikområden är inkluderade i utvärderingen

För åren 2008-2011 bedömer OFT att andelen fördelar för konsumenterna på grund av dess aktiviteter i förhållande till kostnaderna är 10:1, nästan det dubbla jämfört med målet på 5:1. De flesta av dessa fördelar härrör från korrigering/förbud av

annars konkurrenshämmande fusioner, uppbrytning av karteller och effekten av marknadsanalyser.

Vid bedömning av effekten av sina beslut om **fusioner**, stödjer sig OFT starkt på simulering. Man kalibrerar modellerna med värdena låg, medelhög och hög branschelasticitet (som ofta är en nyckelinput i modellen), och man använder mediet som den valda punktuppskattningen. Besparingar antas räcka i två år. Off-model-justeringar kan också behövas för att infoga något särdrag i fusionen eller marknaden som inte fångats upp av modellen. För **ingripanden mot karteller**, baseras där så är möjligt uppskattningarna på information från case-team. Där det inte är möjligt gäller "tumregler som överensstämmer med internationell bästa praxis och aktuell akademisk forskning": (i) i brist på ett ingripande skulle kartellen ha existerat i ytterligare sex år; (ii) endast priset för de överträdande parterna ökas av kartellen; (iii) kartellen skulle ha fortsatt att ta ut överpriser med 15 procent. Vanligen utvärderas väldigt få **artikel 102**-fall. Metoden liknar den för artikel 101, men med två skillnader: (i) standardprisökningen är 10 procent, (ii) varaktigheten överlåts åt handläggarens gottfinnande i det enskilda fallet. **Marknadsanalyser** är det enda område där effektutvärderingar ibland baseras på ingående ex post-analys.

OFT gör också försök att partiellt utvärdera sitt **konsumentskyddsarbete**. Metoden använder data om antalet klagomål före och efter ett ingripande, och minskningar av antalet klagomål omvandlas till ett ekonomiskt värde genom att varje klagomål värderas till en andel av inköpsvärdet av den aktuella produkten eller tjänsten.

Jämförelse av Storbritannien, EC och USA

Om man sätter de rapporterade uppskattade fördelarna för dessa tre jurisdiktioner bredvid varandra står det omedelbart klart att de inte är direkt jämförbara. Det beror på att de olika myndigheterna använder olika antaganden och konventioner i sina uppskattningar.

Det är också påfallande att uppskattningarna varierar avsevärt över tid. En orsak är att storleken på marknaderna där ingripanden görs kan variera dramatiskt, till exempel när stora ingripanden görs på en marknad ett särskilt år kan det leda till ojämnheter i tidsserien. Antalet avslutade fall kan också variera avsevärt över tid. Detta kan bero på de långvariga förfaranden som är förenade med vissa typer av utredningar (särskilt karteller). För att minska omfattningen av svängningar i siffrorna över tid rapporterar brittiska myndigheter tre års rörliga medeltal. En tredje orsak till variationerna är ändringar i de antaganden som används av konkurrensmyndigheterna.

Jämförelse av antaganden

För att uppskatta effekten av ett enskilt ingripande krävs information om marknadens storlek, den avlägsnade eller undvikta prisökningen och hur lång tid

det höjda priset skulle ha gällt om ingripandet inte gjorts. Av dessa är den mest utmanande delen av utvärderingen att uppskatta storleken på prisökningen och hur lång tid ingripandet har effekt.

För **karteller** är den första bästa uppskattningen av en kartells överpriser att använda en uppskattning som har fåtts fram under utredningen. Till skillnad från OFT antar andra myndigheter att deras ingripande bara leder till tio procents minskning av priset. OFT:s antagande om 15 procent verkar överensstämma bättre med empiriska bevis i den akademiska litteraturen. Vad gäller en kartells förväntade livslängd, om den inte hade upptäckts, använder konkurrensmyndigheterna antaganden med vidare spridning. OFT försöker stödja sig på bevis från existerande empiriska studier för att avgöra vad som vore det lämpligaste antagandet, men dessa studier bygger givetvis bara på upptäckta karteller, som mycket väl kan skilja sig från oupptäckta (det antagande som krävs här är hur länge kartellen skulle ha fortsatt om den förblivit oupptäckt). Om man vänder sig till ekonomisk teori för insikter är min slutsats att kartellerna är så heterogena att ett fallspecifikt tillvägagångssätt, som nyligen antogs av EC, verkar överensstämma bäst med ekonomisk teori.

För **ingripanden mot fusioner** är det ganska vanligt att använda simulering för att uppskatta hur priset skulle ha kunnat ändras om fusionen hade fortsatt utan konkurrensmyndighetens ingripande i specifika fall. Alternativt när effekten av en fusion på priset uppskattats under utredningen används det ibland. Till exempel antog EC tidigare att besparingar för konsumenter som blev resultatet av ett beslut om korrigering av en fusion skulle bli så mycket som tio procent av marknadsomsättningen, under det att andra myndigheter ibland har använt ett mycket lägre standardvärde på en procent. Om man vänder sig till den akademiska litteraturen råder ingen uppenbar samsyn om hur mycket konkurrenshämmande fusioner vanligen höjer priset. Av den orsaken verkar simulering fall-för-fall vara det lämpligaste tillvägagångssättet. Konvergensen är större mellan olika konkurrensmyndigheter när det gäller hur länge den prishöjning som fusionen leder till antas ska vara: för den lägre gränsen vanligen ett till två år; men för den högre gränsen verkar den fall-för-fall-metod som används av EC lämplig.

För de andra politikområdena är **missbruk av dominerande ställning** den största utmaningen för bedömningen. Om fallspecifik information inte finns görs flera olika antaganden om standardprisökningen och överträdelsens varaktighet. Emellertid gör bristen på empiriskt akademiskt arbete om missbruk av dominerande ställning det svårt att värdera dessa antaganden. En del konkurrensmyndigheter värderar ibland också **konsumentskydd** eller advokatverksamhet. Exempelvis har FTC rapporterat antalet klagomål av konsumenter och andelen av FTC:s ingripanden för konsumentskydd som riktar sig mot ämnet konsumenters klagomål till FTC. Andra metoder, till exempel undersökningar av konsumenttillfredsställelse, används också i andra länder.

Begreppsfrågor i utvärderingar av ansvarighet

Denna genomgång väcker några allmänna frågor för vidare diskussion och undersökning:

- Uppskattningar är uppenbarligen mycket känsliga för de antaganden som används. En fall-för-fall-metod är därför önskvärd när det är möjligt, men när det inte är det, finns det skäl att eftersträva gemensam praxis när det gäller standardantaganden, och dessa bör baseras på teori och bevis.

- Det är rimligt att tolka resultaten av dessa utvärderingar som typiskt **försiktiga uppskattningar**. Under det att detta är lämpligt särskilt för självutvärdering finns det en risk för att konkurrensmyndigheter kan 'underskatta sig själva' när de rapporterar till regeringen och yttervärlden om de fördelar som de ger.

- Effektutvärderingar beaktar endast de fall då konkurrensmyndigheten hade beslutat att ingripa, och **utesluter fall då det inte skedde något ingripande.** Till exempel i fusioner finns ingen effekt förenad med fusioner som konkurrensmyndigheten rätteligen gav tillstånd till utan ingripande, kanske av effektivitetsförbättrande skäl.

- När effektutvärderingar anskaffas av ansvarsskäl finns det en risk att konkurrensmyndigheten kommer att välja de 'enkla alternativen' i sina ingripanden i stället för att driva svåra fall där det vore mer önskvärt att använda mer resurser och skapa prejudikat, även om dessa inte ger 'mätbara fördelar'. En relaterad risk är att själva utvärderingsprogrammet kan introducera ett ytterligare **motiv för förvrängande gottfinnande** i konkurrensmyndighetens beteende. Av dessa orsaker kan det vara bättre att ibland använda en extern revisor för att genomföra bedömningen

Det finns utan tvekan avsevärda icke observerade effekter av konkurrenspolitik - särskilt **avskräckande av konkurrenshämmande beteende** (se nedan).

Ex post-utvärdering av konkurrensmyndigheter

Ovanstående utvärderingar är huvudsakligen ex-ante, men konkurrensmyndigheter genomför också ex post-utvärderingar, vanligen mer på ad hoc-basis.

Dessa utvärderingar är av två huvudtyper. **Kvalitativa studier** undersöker vanligen marknaden efter ett ingripande för att fastställa om förutsägelserna vid tiden för ingripandet har slagit in eller inte - till exempel marknadsinträde och expansion. **Kvantitativa studier** försöker oftare fastställa orsakssamband mellan ingripande och till exempel en förändring av marknadspriset. När utvärderingen

beställs av konkurrensmyndigheten kan en blandning av dessa två typer av metod användas.

Det mest anmärkningsvärda publicerade exemplet är en rad beställda granskningar för de olika brittiska konkurrensmyndigheterna och relevanta departementen. Olika metoder har använts: bara djupintervjuer; intervjuer med marknadsaktörer i kombination med simuleringar; DiD och en undersökning av marknadsaktörer. Uppföljnings-frågeformulär och/eller intervjuer ger utan tvivel ofta ovärderliga insikter men är oundvikligen behäftade med ett antal potentiella begränsningar: låg svarsfrekvens, fördomar hos svarande, parterna har ofta korta gemensamma minnen, och enligt deras synsätt kan ingripanden ofta tas över av andra efterföljande och viktigare händelser.

Andra ex post-granskningar fokuserar ibland på intressenters uppfattning om konkurrensmyndighetens arbete i allmänhet (t.ex. Storbritanniens Competition Commission (CC) och EU-kommissionen (EC)). EC beställer också en del av sitt utvärderingsarbete av externa experter (se nedan anförda fallstudier).

Utvärderingsmetoder

Andra delen av rapporten (kapitel 6 och 7) fokuserar mer specifikt på de alternativa metoder som används för att utvärdera effekt i särskilda fall.

Simulering och strukturmodeller

Stimulering omfattar vanligen tre stadier: (i) en explicit formell modellering av typen av konkurrens på marknaden, som ofta innefattar en strukturmodell som härrör från oligopolteori, parat med en särskild modell för efterfrågesystemet; (ii) kalibrering av modellen med numeriska värden för modellens parametrar, härledda från direkta observationer eller från ekonometrisk uppskattning av efterfrågesystemet (t.ex. befintliga marknadsandelar, priser och externa uppskattningar av efterfrågeelasticiteter); (iii) användande av modellen för att bedöma hur jämvikten skulle förändras med och utan ett ingripande av en konkurrensmyndighet. Simulering kan vara antingen ex-ante eller ex post, men de flesta tillämpningar är ex-ante i den meningen att de baseras på data som är tillgängliga vid tiden för ingripandet - även om de faktiskt genomförs retrospektivt.

- Tidigare litteratur

Den tidiga utvecklingen av simulering var fast rotad i den akademiska litteraturen och de flesta av de fruktbärande bidragen avsåg simulering av fusioner. I ökande grad har de senaste 20 åren simulering förekommit i specifika fusionsutredningar av konkurrensmyndigheterna och rådgivare till parterna.

- Fördelar och nackdelar

Simuleringens största styrka är dess uttryckliga användning av teori för att identifiera det kontrafaktiska fallet. Detta underlättar 'förening' av den analys som genomfördes vid tiden för ingripandet med efterföljande utvärdering av politikens effekter, och ger i sin tur en tydlig möjlighet att utvärdera de antaganden som gjordes vid tiden för ingripandet.

Emellertid är simulering väldigt känsligt för modelleringsantaganden. Ibland är det en styrka, t.ex. genom att det avslöjar hur känsliga förutsägelser är för det kontrafaktiska fallets exakta art, men ibland är denna känslighet inte till någon hjälp, eftersom det inte finns några starka teoretiska skäl att välja, t.ex. efterfrågekurvans funktionella form. Simulering passar bättre för vissa typer av oligopolmodeller (och därför marknader) än andra. Det är lättast att använda när konkurrens åstadkoms genom pris och kvantitet med uteslutande av innovation, ompositionering etc., och möjliga förändringar i beteende (relevant för koordinerade effekter). Köpkraft har också visat sig svårt att införliva tillfredsställande, och simulering av budgivningsmarknader är fortfarande i sin barndom. Således är utvärdering baserad på simulering kraftigt vinklad mot vissa typer av marknader, vilket kan leda till ensidighet vid urval av sampel. En annan potentiell källa till ensidigt urval härrör från simuleringens stora beroende av data. Många av de fruktbärande studierna bygger på disaggregerade datauppsättningar av hög kvalitet konstruerade från scannerkällor; men dessa hämtas givetvis oftast från en relativt liten uppsättning konsumtionsvaror (som ofta säljs i livsmedelsbutiker). Slutligen är bevisen blandade för hur väl simulering förutsäger faktiska utfall.

- Back-of-the-envelope (enkel) simulering

Eftersom fullständig simulering är mycket data- och analystidskrävande använder konkurrensmyndigheter ofta förenklade versioner - som vanligtvis kan approximeras av Cournot-modellen eller enkla differentieringsmodeller

Eventstudier

En eventstudie utnyttjar data om finansmarknader för att mäta effekten av en händelse på ett företags marknadsvärde. Om finansmarknader är effektiva kommer effekten av någon händelse på ett företags diskonterade resultat omedelbart att kunna iakttas genom förändringarna i dess aktiekurs. Metoden innebär mätning av eventuella onormala intäkter förenade med en händelse (t.ex. tillkännagivande av en fusion, och dessa identifieras som skillnaden mellan den observerade rörelsen i aktievärderingen och den som skulle ha inträffat om händelsen inte inträffat.

- Tidigare litteratur

För **fusioner** granskar eventstudier vanligen effekten av fusionens tillkännagivanden och konkurrensmyndighetens beslut på värderingen av de sammangående företagen och deras konkurrenter. Med stöd i stabila resultat från den övervägande delen av oligopolteorin, kommer en konkurrensfrämjande fusion att skada parternas konkurrenter, under det att en konkurrenshämmande fusion kommer att gynna konkurrenterna. Detta framhäver metoderna i ett antal studier av effekterna av fusionsregleringar i de viktigaste jurisdiktionerna de senaste 20 åren. Eventstudier av **karteller** granskar vanligen effekten av gryningsräder och konkurrensmyndighetens efterföljande beslut på aktieägarnas lagervärderingar. Metoden har mindre ofta använts på fall av **missbruk av monopol**.

- Fördelar och nackdelar

Det centrala antagandet i en eventstudie är att finansmarknaderna är effektiva. Om de är det ger denna metod opartiska uppskattningar som är både "objektiva" och snabba.

Emellertid ifrågasätts detta antagande - på allmän nivå - av ett antal kommentatorer. Mer specifikt måste, när det tillämpas på fusioner, ett antal frågor ställas: till exempel kan en ökning av parternas aktiekurs spegla antingen konkurrensfrämjande effekter (effektivitetsvinster) eller konkurrenshämmande effekter (uteslutning, dominerande ställning eller hemliga överenskommelser); avslöjar på liknande sätt ett ras i värderingarna av konkurrenterna konkurrensfrämjande (effektivitet) förväntningar eller förutser det konkurrenshämmande (uteslutnings-)effekter? Det finns också en möjlighet att fusionen kan tolkas som en signal om att andra företag på samma marknad kommer att inleda fusionsaktiviteter inom en snar framtid, som skulle öka marknadens värdering av konkurrenter, och leda till samma tecken på förändring som med resultat av hemliga överenskommelser och/eller dominerande ställning. För- och nackdelarna är liknande för bedömning av ingripanden mot karteller, fast de teoretiska förväntningarna är mindre tvetydiga. Effekten av nyheter om utredningen och det slutgiltiga beslutet minskar nästan alltid värderingen av parterna. På praktisk nivå kommer det trots det allmänna antagandet att eventstudier är lätta att använda och data lätta att komma åt från ekonomiska databaser finnas många omständigheter när lämpliga data inte är tillgängliga - i synnerhet för onoterade bolag och/eller för konglomerat och multinationella företag, för vilka den aktuella marknaden kanske endast utgör en liten del av dess totala verksamhet.

Skillnad-i-skillnader

I metoden skillnad-i-skillnader, (DiD), är grundidén att utvärdera 'resultatet' före och efter en händelse (eller ibland före, under och efter) på den aktuella marknaden jämfört med resultatet på en annan liknande (kontroll-)marknad, som är opåverkad

av händelsen. Den är vanligen ekonometrisk, och priset spåras över tid för att inkludera för- och efterperioderna i behandlingen (t.ex. fusion) och jämfört med detsamma på kontrollmarknaden. Nästan per definition genomförs DiD-analys bara ex post. Fördröjningen innan den genomförs speglar en kompromiss mellan att välja en tillräckligt lång period efter händelsen för att få en bättre uppfattning om de långsiktiga effekterna och undvika att välja en tidsperiod som är så lång att den skulle äventyra chanserna att hitta en användbar kontroll för att efterlikna det kontrafaktiska fallet.

- Tidigare litteratur

Metoden har tillämpats ganska omfattande på fusioner, ingripanden i fusioner och karteller; det finns få exempel där DiD har tillämpats på fall med missbruk av dominerande ställning, avskräckning eller andra effekter av offentligt ingripande.

- För- och nackdelar

Metoden skillnad-i-skillnader är tilltalande eftersom den använder observerade data om vad som faktiskt händer jämfört med en kontrollgrupp, där händelsen inte inträffar. Således är det kontrafaktiska fallet inte beroende av icke testbara, kanske restriktiva, teoretiska antaganden. Emellertid är det då avgörande att den valda kontrollgruppen nära motsvarar vad som skulle ha hänt på den behandlade marknaden, om inte ingripandet skett. I praktiken finns det en risk att valet av kontrafaktiskt fall begränsas av 'vad som finns där ute', dvs. det bästa av en uppsättning alternativ, av vilka inget är helt lämpligt. För att bara nämna ett exempel är den valda kontrollgruppen ofta konkurrenterna till parterna i ett fusionsfall, men så som visas av den övervägande delen oligopolteori ändrar också konkurrenter priserna efter en fusion, och därför är kontrollgruppen inte oberoende av de behandlade (dvs. sammangående) företagen.

- Före och efter-approximering till skillnad-i-skillnader

En av de oftast använda utvärderingsmetoderna är att genomföra en enkel jämförelse före och efter av till exempel pris. Detta är i strikt mening inte någon skillnad-i-skillnader eftersom det inte finns någon kontrollgrupp. Undermeningen är i själva verket att det inte behövs någon kontrollgrupp, eller att det kontrafaktiska fallet bara är status quo: om fusionen eller ingripandet inte skett skulle marknaden helt enkelt ha presterat lika före och efter. Resultatet av en sådan metod ska definitivt tolkas med försiktighet - man skulle åtminstone sträva efter kvalitativa källor (t.ex. intervjuer/frågeformulär) för att underbygga valet av ett kontrafaktiskt fall som inte förändras.

Att välja mellan metoder

Mycket av diskussionen om fördelarna och nackdelarna med olika metoder avser hur träffande de implicit beskriver det kontrafaktiska fallet. Simulering placerar genom sin karaktär valet av kontrafaktiskt fall på ett iögonenfallande sätt mitt på scenen: en specifik oligopolmodell måste identifieras, fast detta förstås inte innebär att rätt modell väljs. I DiD spelar kontrollgruppen rollen av kontrafaktiskt fall. Här är valet av kontrafaktiskt fall mindre teoretiskt drivet, och metodens styrka beror på huruvida det faktiskt finns en kontrollgrupp (med adekvata data) som är tillräckligt lik. I praktiken kan datas lämplighet ibland dra uppmärksamheten från hur väl detta villkor uppfylls. Under det att det är mindre vanligt att tänka på det kontrafaktiska fallet i den typiska eventstudien, finns det fortfarande implicit där - fångat av det jämförande aktieprisindex som utvärderaren använder för att beräkna onormala intäkter. Här finns en kompromiss mellan att använda ett allmänt index, som sannolikt är mindre känsligt för marknadsspecifika yttre händelser, och mer skräddarsydda sektorspecifika index, som kanske inte är verkligen oberoende av händelsen i fråga.

I en idealvärld skulle de alternativa utvärderingsmetoderna jämföras genom att samla ett väldigt stort, slumpmässigt sampel av fall, och försöka tillämpa alla metoder på alla fall i samplet. Under det att det är sannolikt att vissa konkurrensmyndigheter och rådgivningskonsulter ibland genomför parallella bedömningar under genomförandet av ett särskilt fall (samtidigt experimenterande med till exempel en eventstudie och en simulering), förekommer de givetvis inte i den publicerade litteraturen. Mer allmänt vore det svårt att försöka sig på en sådan uppgift i ett stort sampel av fall, och det har i själva verket aldrig gjorts hittills.

Nästan som en utvikning från ämnet föregriper avsnitt 6.4 hur en sådan stor jämförelse av sampel skulle kunna genomföras genom att på nytt studera en existerande databas med uppskattningar av överpriser i karteller för ett mycket stort sampel av olika karteller observerade i olika länder och vid olika tidpunkter. Eftersom dessa uppskattningar härletts med en mängd olika metoder är syftet att identifiera huruvida vissa metoder i sig med större sannolikhet genererar större uppskattningar av omfattningen av överpriser. Under det att jämförelsen avslöjar vissa anmärkningsvärda skillnader skulle ytterligare arbete, som mer noggrant kontrollerar urvalet av sampel, krävas innan några definitiva slutsatser kan dras.

Exempel på fallstudier av utvärderingsmetoder

Sex representativa fallstudier används för att illustrera ovanstående metoder.

Fall 1: Två fusioner i Storbritannien (skillnad-i-skillnader)

Denna studie genomfördes av konsultfirman LEAR för Storbritanniens Competition Commission (CC). Två olika fusionsbeslut togs i betraktande: (i)

GAME/ Gamestation, en fusion mellan två återförsäljare av videospel, och (ii) Waterstone's/Ottakars, en fusion mellan två specialiserade bokdetaljhandlare på 'huvudgatan'.

Dessa två fall illustrerar vissa av de allmänna särdragen i skillnad-i-skillnader-metoden. Kraven på data var starka. Emellertid underlättade det väsentligt att data från tiden före fusionen redan fanns att tillgå från de ursprungliga undersökningarna. Icke desto mindre visade det sig även i dessa två exempel rika på data omöjligt för ett mycket kunnigt team som inte led av alltför stora tids-/kostnadsbegränsningar att utföra några test av huvuddimensionen service-kvalitet. Valet av och tillgången till alternativa kontrafaktiska fall är också en huvudfråga. I detta fall som i många andra fall med lokala detaljhandlare, är existensen av vissa lokala marknader där bara en av parterna var närvarande före fusionen mycket värdefull. Å andra sidan är användningen av konkurrenters priser som en kontroll potentiellt problematisk.

Fall 2: Ost i Mauritius (back-of-envelope skillnad-i-skillnader)

Detta fall används som ett mycket nytt exempel på en intern ex post-effektutvärdering av en ung konkurrensmyndighet för dess ingripande 2010 mot ett missbruk av monopol. Missbruket var en praxis att erbjuda retroaktiva rabatter på processad cheddar i block av märket Kraft i utbyte mot bästa butiksutrymme i livsmedelsaffärer. Utvärderingen genomfördes ovanligt snabbt, bara ett år efter ingripandet av Mauritius Competition Commission (CCM). Emellertid var det två år efter att utredningen inletts och författarna ansåg att effekten började göra sig kännbar redan medan utredningen pågick. Erforderliga data samlades in från olika statliga institutioner och från marknadsaktörerna själva.

Detta är ett exempel på vad vi kallar en mycket förenklad back-of-envelope skillnad-i-skillnader, i vilken CCM förutsatte att om den inte hade ingripit skulle det inte ha skett något inträde och priset skulle ha fortsatt att stiga i samma takt som före ingripandet. Sådana enkla antaganden är givetvis bara gissningar och kan ifrågasättas. Å andra sidan gör de att den erforderliga analysen blir mycket mindre data- och tidskrävande.

Fall 3: Pirelli/BICC-fusionen (eventstudie)

Denna studie beställdes av EU-kommissionen och genomfördes av Lear Consulting. Den avsåg Pirellis förvärv av sex produktionsanläggningar i Italien och Storbritannien i kraftkabelindustrin från BICC General. Data som användes togs från Datastream om konkurrenternas aktiekurser och de sammangående parternas kunder vid tre händelser: när fusionen först föreslogs, sedan vid EC:s tillkännagivande att man skulle inleda en fas 2-undersökning, och slutligen vid beslutet om godkännande.

På många sätt är detta ett klassiskt exempel på en eventstudie. Just den här branschen bör ge fruktbar mark för användning av en eventstudie eftersom många av de relevanta företagen är stora och börsnoterade. Med detta sagt var även i en så

storskalig, viktig bransch som den här bara hälften av alla relevanta företag börsnoterade, vilket väckte frågor om ensidighet i urvalet. Dessutom var många av de börsnoterade företagen i samplet multinationella och diversifierade, vilket väckte tvivel om värderingen av deras aktie skulle vara känslig för en fusion mellan två av deras konkurrenter i bara två länder i bara en av deras branscher. Resultaten tyder på att fusionen inte hade någon effekt på konkurrenterna och blandad effekt på kunderna. Emellertid visar noggrann granskning av storleken på uppskattningarna att de ofta är osannolikt höga, eller helt betydelselösa. Under det att det är sant att en betydelselös effekt är förenlig med att fusionen inte har någon negativ inverkan på konkurrensen, är den också förenlig med en mer cynisk tolkning att fusionen (vilken effekt den än har på konkurrensen) inte skulle ha någon märkbar effekt på den framtida lönsamheten i sådana stora multinationella diversifierade företag.

Fall 4: Antitrustböter i sydafrikanska bröd- och mjölkarteller (eventstudie)

Detta fall använder också en eventstudie för att uppskatta effekterna av böter på börsvärderingarna av företag som gjort sig skyldiga till överträdelser och som är inblandade i bröd- och mjölkarteller i Sydafrika (Darj et al., 2011). Det är ett andra exempel på utvärdering av en relativt ung konkurrensmyndighet utanför Europa. Det är också mycket mindre tekniskt, och var förmodligen mycket mindre tidskrävande än föregående fall. Detta fall är emellertid ett mycket bra exempel på hur en eventstudie kan användas framgångsrikt på ett enkelt sätt för att kvantifiera effekterna av ingripande på det berörda företaget - såsom är möjligt med böter. Det är viktigt att effekten av böter kan överstiga kostnaderna för själva böterna på grund av andra potentiella indirekta effekter på företagets framtida lönsamhet. Det föreligger ingen verklig komplikation avseende det kontrafaktiska fallet, som är ganska uppenbart i fall med böter.

Fall 5: Volvo-Scania (simulering)

Detta fall är ett exempel på en fullständig, tekniskt mycket avancerad simulering av en föreslagen fusion mellan Volvo och Scania på den europeiska lastbilsmarknaden. I centrum för metoden står ekonometrisk uppskattning av efterfrågesystemet av ekvationer för lastbilar. Den använder paneldata för 16 olika medlemsstater under två år. Priser är listpriserna för en basmodell och omsättning är omsättningen för modellsortimentet. Den ekonometriska modell som används är nested logit, och det är nödvändigt att inkludera variabler som speglar de olika egenskaperna hos olika lastbilsmodeller. Emellertid, och detta är en potentiell svaghet i uppskattningen, användes bara en verkligt teknisk egenskap - lastbilens hästkrafter. Effekterna av fusionen simulerades sedan genom att introducera en vald icke samarbets-oligopolmodell (dvs. ensidiga effekter) och analytiskt lösa (i) jämviktspriserna med antagande först om ingen fusion (dvs. före) och sedan (ii) med fusionen (dvs. efter).

Fallet illustrerar väl många av de vanliga styrkorna och svagheterna i simulering av fusioner. Det tvingar analytikern att uttryckligen ange vad som är det valda kontrafaktiska fallet, men det är också möjligt att simulera alternativa kontrafaktiska fall (oligopolspel). I sig själv kräver strängheten i simuleringsmodellen också att analytikern bildar sig en grundlig uppfattning om konkurrensens och efterfrågans natur på den berörda marknaden. Å andra sidan kan resultaten vara väldigt känsliga för alternativa antaganden, och datakraven kan vara prohibitiva - och kräver ofta approximationer (t.ex. här bristen på mått på produktegenskaper). Slutligen är simulering, när den genomförs på den tekniskt komplicerade nivå som används här, uppenbarligen tids- och resurskrävande och kräver mycket kunniga ekonometriker.

Fall 6: En genomgång av åtta brittiska fusionsbeslut (back-of-envelope simulering)

Det sista fallet använder också simulering, men denna gång mycket mer förenklat och mindre tidskrävande. Det är en utvärdering på bred basis av åtta olika fusioner, genomförd av konsultfirman Deloitte för ett konsortium av brittiska myndigheter (2007). Syftet var att genomföra en genomgång av fusionsbeslut enligt den relativt nya företagslagen 2002. Analysen baserades på två metoder. Den första var en kvalitativ undersökning som använder intervjuer med marknadsaktörer, kompletterad av frågeformulär och officiellt tillgängliga data. Den andra var att genomföra ex ante-simulering på så många av fallen som möjligt, som simulerade vad de sannolika följderna av fusionerna skulle bli, men som bara använde information som varit tillgänglig för myndigheterna vid tiden för deras beslut. Simuleringarna var begränsade till enkla back-of-envelope-metoder som kunde utföras snabbt och med minimala datakrav. I själva verket var simulering, givet databegränsningarna, bara möjligt i fyra av de åtta fallen. I vart och ett var produkten fullständigt homogen och kapacitetsbegränsningar var typiskt nog ett problem. Av den orsaken var den modell som användes en kapacitetsbegränsad Cournot-modell.

Studien illustrerar mycket väl vad som kan och vad som inte kan åstadkommas med back-of-envelope-simulering. Dess styrka ligger i snabbheten och enkelheten, och därigenom möjligheten till tillämpning på ett antal olika fall under en ganska kort analysperiod. Svagheten är att den ofta bara kan tillämpas på enkla marknader (här, på vilka konkurrensen kunde beskrivas med enkla ehuru kapacitetsbegränsade Cournot-modeller), eller på vilka enkla avvikelsetal kan användas för att modellera produktdifferentiering. Dessutom är en risk med simuleringsmodeller den falska känsla av precision och exakthet som de kan ingjuta. Författarna betonade mycket riktigt att sådana simuleringar bara bör betraktas som anordningar för att visa hur olika antaganden om marknaden och konkurrensens natur kan översättas till olika utvärderingar av den sannolika effekten av olika beslut. Med andra ord är deras uppgift bara att tillhandahålla ett av ett antal olika hjälpmedel för beslutsfattande.

Några viktiga inte tillräckligt utvecklade problem

Enligt författarens åsikt finns det tre brister, osäkerheter och olösta problem i tidigare litteratur och praxis som förtjänar att undersökas i framtiden.

Ensidighet i urval och avskräckning

Utvärderingsstudier tenderar oundvikligen att fokusera på ett sampel av fall som är: (i) dokumenterade och tillgängliga i litteraturen och bland dessa (ii) vanligen de där ett ingripande faktiskt har skett eller allvarligt övervägts, och (iii) för vilka utvärdering är genomförbar. Detta reser allvarliga tvivel huruvida ett sådant sampel kan ge en opartisk uppskattning av konkurrenspolitikens fulla fördelar.

Några potentiella källor till ensidighet i urval

Såsom betonats i denna rapport härrör en källa till ensidighet från det faktum att olika utvärderingsmetoder är mindre användbara för vissa typer av marknader och ingripanden. Men frågan är mycket större än så. Tidigare litteratur innehåller flera andra exempel på varför utvärdering är känslig för ensidighet i urval:

- I den empiriska litteraturen om **karteller** är det allmänt bekant att analyserade sampel kan vara ensidiga eftersom de hämtats uteslutande från upptäckta utredda fall, som kanske inte är representativa för den okända populationen av oupptäckta fall.

- När man bedömer hur sträng **politiken avseende fusioner** är, är det direkt vilseledande att fokusera på *obestridda* fusioner, för även om konkurrensmyndigheten är 'slapp' (benägen för typ 1-fel), kommer ett sådant sampel att inkludera många 'goda' fusioner tillsammans med eventuella felaktigt tillåtna 'dåliga' (prishöjande) fusioner.

- Det är allmänt erkänt att de fördelaktiga **avskräckande** effekterna av konkurrensmyndigheters ingripanden sannolikt är avsevärda. Av det följer att utvärderingar av fördelarna med en politik baserade bara på utredda fall kan vara en allvarlig underskattning.

I rapporten presenteras ett klassifikationsschema avsett att strukturera de olika dimensionerna av potentiell ensidighet i urval, och anvisningar för framtida undersökningar belyses.

Hur stor del av isberget ligger under vattenytan?

I avsnitt 8.1 i rapporten ställer jag den hypotetiska frågan: "Vilken information skulle behövas innan man kunde extrapolera från vad vi vet - baserat på undersökta fall - till vad vi inte vet - om de fall som är avskräckta eller oupptäckta eller inte undersökta?"

Om man inför tre villkorliga sannolikheter: avskräckningsgraden (ω), upptäcktsgraden (ϕ), och undersökningsgraden (σ) följer det att samplet av undersökta fall representerar bara en (kanske väldigt) liten andel, $(1- \omega)*(\phi)*(\sigma)$, av populationen av alla potentiellt relevanta fall.

Under det att frågan är hypotetisk är det lärorikt att som ett tankeexperiment spekulera om den typ av storheter som är involverade. För att göra det använder vi en ovanlig kvalitativ studie av avskräckning, beställd av OFT, som kan användas för att kalibrera ovanstående tre parametrar. När vi för in dessa i ovanstående uttryck antyder det att (i) det undersökta samplet fusioner bara är omkring tio procent av hela populationen potentiella fusioner och att upptäckta karteller är mindre än 4 procent av alla potentiella karteller; och (ii) det finns fem (15) gånger så många avskräckta som undersökta fusioner (karteller). Emellertid måste det betonas att dessa beräkningar är mycket ungefärliga och mycket spekulativa.

Det bör också betonas att dessa beräkningar bara avser *frekvensen* fusioner och karteller - de kvantifierar inte nödvändigtvis de relativa *volymerna* avskräckt skada eller missade möjligheter. Detta vore endast sant om den förväntade skadan av observerade och icke observerade fall vore identisk. I själva verket betyder ovanstående oro för ensidighet i urval att det finns mycket reella skäl att anta att detta inte är fallet, dvs. att de observerade fallen inte är ett slumpmässigt sampel av hela (delvis icke observerade) populationerna.

Rapporten granskar en del av den akademiska litteratur som mest sannolikt är till hjälp för att främja vår förståelse av arten av och storleken på de olika potentiella ensidigheterna i urval. Det är emellertid uppenbart att mycket framtida forskning, både teoretisk och empirisk, kommer att behövas.

Behovet av långsiktiga studier

Det är nästan en truism att påstå att effekterna av en förändrad politik eller ett ingripande (på vilket område som helst) sannolikt kommer att vara på lång sikt. Det är också ganska uppenbart att utvärdering av långsiktiga effekter är svårare än kortsiktiga utvärderingar - med tanke på risken för möjliga sammanblandningseffekter med tiden, som gör det svårt att urskilja effekter som orsakats av ingripandet från effekter orsakade av externa faktorer.

Detta stämmer förvisso för antitrust. Den akademiska litteraturen talar om flera sätt på vilka en specifik händelse kan utlösa en följd eller kedja av efterföljande händelser - vilka var och en kan utvärderas oberoende, men som i själva verket är tydligt vägberoende. Litteraturen om endogena fusioner väcker oss till exempel till insikt om att om fusion A godkänns, gör det en följande fusion B mer eller mindre sannolik. På liknande sätt kan i fall med fusioner av sviktande företag följderna av ett ingripande inkludera efterföljande förslag till alternativa fusioner av andra parter, såsom hände i fallet med Airtours. Det finns bevis från fallstudier som talar för att ibland när en viss konkurrenshämmande praxis förbjuds kommer företag att

försöka ersätta den med annan. Det har länge varit erkänt att horisontella fusioner ibland kan vara ett alternativ som företag väljer när kartellisering är omöjligt.

Sammanfattningsvis har vad som händer efter ingripandet betydelse, och 'efter' bör ibland tolkas som långsiktigt och inte för inskränkt. Under det att detta ibland erkänns i utvärderingslitteraturen är fall med långsiktig utvärdering ovanliga. Det är författarens åsikt att framtida forskningsprogram bör prioritera långsiktigare utvärderingar. Ett möjligt sätt att göra det är genom fallstudier av marknadens utveckling över låt säga decennier och som involverar en rad ingripanden, snarare än att de är begränsade till bara enstaka, oberoende utvärderingar av enstaka händelser.

Konkurrenspolitikens vidare inverkan

Största delen av denna rapport har fokuserat ganska specifikt på effekten av konkurrensmyndigheters ingripanden på konsumenternas välfärd. Emellertid bör vi också ägna oss åt den vidare inverkan på ekonomin som helhet och makroaggregat som tillväxt, investeringar, sysselsättning och produktivitet och innovation.

Det finns empiriska studier som utforskar dessa områden, och en del av litteraturen gås igenom i rapporten. Emellertid är det om sanningen ska fram en tunn litteratur av varierande kvalitet. Det finns två huvudsakliga inriktningar: arbeten som undersöker konkurrensens inverkan på produktivitet, tillväxt etc., och forskning som mer specifikt försöker mäta effekten av *konkurrensmyndigheters ingripanden* på samma faktorer. På det hela taget ger forskning av första typen en bred bekräftelse på att konkurrens bidrar till att driva ekonomisk tillväxt, genom minst tre olika kanaler: (i) genom att den utövar press på företagens ledning att minska ineffektiviteten; (ii) genom att främja innovation, och (iii) genom att styra en bättre fördelning av resurser till mer effektiva företag. Emellertid använder många studier av denna typ tvetydiga eller tvivelaktiga mått på graden av konkurrens.

När det gäller den andra typen - effekterna av konkurrensmyndigheters ingripanden - är en avgörande utmaning att hitta lämpliga mått på konkurrensmyndigheters ingripanden eller konkurrenspolitik. Det har gjorts framsteg i detta ämne och flera olika index över konkurrenspolitikens effektivitet har föreslagits. Dessa stödjer sig vanligen på institutioners och konkurrensmyndigheters egenskaper, till exempel graden av konkurrensmyndighetens formella självständighet, åtskillnad mellan dömande och beivrande funktioner inom konkurrensmyndigheten, omfattningen av konkurrensmyndighetens utredningsbefogenheter, nivån på aktiviteterna (sanktionernas storlek, budgetarnas och resursernas storlek etc.), och hur allvarliga de sanktioner är som konkurrensmyndigheten kan utdöma. Det finns numera en del bevis för att konkurrensmyndigheters ingripanden bidrar till ekonomisk tillväxt. Emellertid behövs det mer arbete på detta område innan vi kan hävda att detta resultat är säkert.

Slutsatser

Varför bör konkurrensmyndigheter utvärdera sitt eget arbete?

Jag anser att det finns minst tre starka skäl till att konkurrensmyndigheter bör utvärdera sitt eget arbete:

För det första, om konkurrensmyndigheter ska upprätthålla och till och med öka sina anslag från staten, är det viktigt att de kan visa att de ger valuta för pengarna. Det är min åsikt och erfarenhet att alla väl fungerande myndigheter kan göra just det utifrån en allsidig och offentlig utvärdering av sina aktiviteter, även med de mest försiktiga antaganden. Denna åsikt bygger på två (enligt min åsikt) mycket sannolika antaganden: (i) att konkurrens väsentligt ökar konsumenternas välfärd, och (ii) de flesta konkurrensmyndigheter verkligen främjar konkurrens genom sina olika aktiviteter.

För det andra bör regelbunden och relativt allsidig utvärdering utgöra en viktig beståndsdel i beslut om intern resursfördelning inom myndigheten. Vid fördelning av budgeten mellan olika aktiviteter (kontroll av fusioner, upptäckt av karteller, övervakning av dominerande företag), är det viktigt att veta vad som är den sannolika marginalprodukten av att spendera ytterligare en euro på varje område. Mer allmänt bör fördelningsbeslut - säg mellan upptäckt och avskräckning - präglas av en viss kunskap om utbytet.

För det tredje finns det i varje väl sysselsatt konkurrensmyndighet en inneboende risk att när väl ett beslut har fattats eller ett ingripande gjorts det sätts in i en pärm och aldrig mer öppnas. Ex post-utvärdering är förmodligen det bästa sättet att säkerställa att myndigheten blir medveten om 'vad som händer härnäst' på en marknad. Av särskild betydelse är frågan om hur marknader reagerar på och presterar efter ingripanden. Dessutom ger utvärdering en värdefull kontroll av huruvida antagandena, analysen och beslutsfattandet som gjorts vid tiden för beslutet rättfärdigas av vad som händer under efterföljande år.

Emellertid är två varningar på sin plats här. För det första bör vi inte bli slaviskt fästa vid de exakta uppskattningar som erhålls i alla utvärderingar. Utvärdering är alltid ett oprecist ämne, oundvikligen baserat på antaganden som är ungefärliga och ibland inte går att verifiera. Utvärdering finns för att vägleda snarare än att diktera. För det andra kan en övertro på uppskattningar av effekter snedvrida fördelningsbeslut inom en myndighet. Risken är att vissa aktiviteter, särskilt avskräckning, kan negligeras i relativt stor utsträckning eftersom effekten av dem är svårare att mäta.

Hur bör konkurrensmyndigheter genomför utvärdering?

Två teman löper genom hela denna rapport - att erhålla uppskattningar av den samlade effekten av en myndighets aktivitet, och sedan mer specifikt de metoder

som kan användas för att utvärdera enskilda fall. Det är min åsikt att konkurrensmyndigheten bör göra båda delarna - regelbundet (kanske varje år) utvärdera sin totala påverkan och även ägna resurser åt djupa ex post-utvärderingar av åtminstone några specifika fall.

Det finns ett antal problem och beslut som skulle behöva lösas när man tar fram utvärderingar av den totala påverkan - vilka tumregler som ska användas för kartellers överpriser och varaktighet etc., vilka aktivitetsområden som ska utvärderas. Här kan befintliga metoder som tillämpas av OFT, DGCOMP, FTC och DoI användas, men i brist på internationellt överenskommen bästa praxis behöver det inte vara nödvändigt.

Rapporten behandlar de specifika utvärderingsmetoderna och antyder att det inte finns någon idealisk metod. Simulering, eventstudier, skillnad-i-skillnader, före-efter och mer kvalitativa undersökningar har sina styrkor och svagheter, och var och en är mer lämpad för vissa fall än för andra. Det kanske viktigaste budskapet i rapporten är att tillräcklig tid och uppmärksamhet bör ägnas formulering av ett lämpligt kontrafaktiskt fall (eller alternativa kontrafaktiska fall). Ibland kommer det kontrafaktiska fallet att diktera vilken utvärderingsmetod som bör användas.

En tredje allmän rekommendation är att myndigheten bör använda en portföljmetod för sin utvärderingsverksamhet. Mycket talar således för att använda en blandning av ex ante- och ex post-utvärderingar – så som de andra myndigheter som oftast diskuteras i denna rapport gör. Likaledes bör en del av utvärderingen använda 'back-of-envelope'-metoder, under det att en del kan vara mer grundlig och tekniskt avancerad. Det finns fördelar med att anlita utomstående (konsulter och akademiker) för att göra en del, men inte hela undersökningen. Det uppenbara problemet med intern utvärdering är misstanken att den kommer att betraktas som självrättfärdigande och bristande i objektivitet. Å andra sidan är det viktigt att myndigheten utvecklar sin egen tekniska skicklighet.

Vilka är utmaningarna?

Förutom de tre områden där jag anser att ytterligare arbete behövs bör det betonas att utvärderingsprojektet fortfarande till stor del är 'pågående arbete'. Utvärdering är fortfarande underutvecklat på vissa politikområden, till exempel advokatverksamhet, utbildning och artikel 102. Inte mycket är känt vad gäller jämförelser om de olika metodernas förmåga att förutsäga. Här är utmaningen att bredda utvärderingens omfattning utan att oskäligt öka kostnaderna. Som på många politikområden spelar internationellt samarbete och spridning av erfarenheter och kunskaper en viktig roll.

Executive summary

Introduction

This report provides a survey of the approaches and specific methods which are used to assess the economic effects of a Competition Authority's work, paying particular attention to the advantages and disadvantages of those methods.

Purposes of evaluation

Evaluation is undertaken by either Competition Authorities (CAs) and/or independent academics/consultants. It can take many forms, which can be grouped into four categories.

- Accountability: there is an obligation to quantify the aggregate benefits of competition policy which can be assessed against some pre-specified target.

- Assessment of specific areas of policy or a particular intervention, with the purpose of checking on the quality of the CA's decision-making; this may be used to help set future internal priorities and resource allocation.

- Assessment in the academic literature has often helped in the technical development of the evaluation methodologies used by CAs.

- Estimating damages and fines often requires an evaluation of the harm caused by the firms concerned, and therefore, the gains resulting from CA intervention

The methodologies of evaluation

This report focuses mainly on three key quantitative approaches:

- Simulation, based on econometric modelling of oligopoly theory
- Event studies, using evidence from the capital market
- Difference-in-differences, based on statistical analysis of before and after.

We also discuss other methodologies more briefly, including: aggregate international econometric comparisons; follow up surveys of the parties/rivals/customers/suppliers; expert commentaries and court decisions; and surveys of peer/practitioner opinions.

Ex-ante versus ex-post

Ex-post evaluation is conducted retrospectively using information on 'what actually happened next', i.e. after the intervention took place. To evaluate impact,

'what happens next' is then compared to what would have happened, had the intervention not taken place. Ex-ante evaluation may be conducted either before or after the event, but only employs the information that would have been available at the time of the policy decision. In that case, the practitioner projects forward comparing of what would happen with and without the intervention.

Policy areas

In principle, evaluation can be applied to all areas of competition policy, but in practice mergers and cartels have attracted most attention. In the author's opinion, further research is required on evaluation of other areas, notably Article 102 (abuse of dominance) and advocacy.

The counterfactual

A key issue in any evaluation is identifying an appropriate counterfactual – what would have happened, absent an intervention. This sometimes attracts insufficient attention in previous studies.

Evaluations for accountability

The first part of the report (Chapters 2-5) is devoted to evaluations of the overall impact on consumer welfare of the activities of a competition authority. It focuses on the reported methodologies of the authorities in which such work appears to be most advanced, at least as can be judged by publications in the public domain.

The UK's Office of Fair Trading (OFT)

Each year the OFT publishes an annual positive impact report, which quantifies the value of the welfare benefits of its interventions compared to its cost to the UK taxpayer. The key defining features of its methodology are that:

- Estimates are in terms of consumer benefits
- Estimates are generally ex-ante, but with some ex-post cases
- Estimates are deliberately 'conservative', and are presented as 'point' estimates, alongside a range of plausible values
- Most, but not all, areas of policy are included in the evaluation

For the years 2008-2011, OFT estimates that the ratio of the consumer benefits from its activities to its costs is 10:1, nearly double its target of 5:1. Most of these benefits derive from remedying/prohibiting otherwise anti-competitive mergers, breaking cartels and from the impact of market studies.

In estimating the impact of its **merger** decisions, it relies heavily on simulation. It calibrates the models with low, medium and high values of the industry elasticity (which is often a key input in the model), and it employs the medium as the chosen

point estimate. Savings are assumed to last for two years. Off-model adjustments may also be necessary to accommodate any features of the merger or market not picked up by the model. For **cartel interventions**, where possible, its estimates are based on information from case teams. Where this is not available, it applies "rules of thumb that are consistent with international best practice and recent academic research": (i) absent intervention, the cartel would have existed for a further 6 years; (ii) only the price of the infringing parties is increased by the cartel; (iii) the cartel would have continued to overcharge by 15 per cent. Typically, very few **Article 102** cases are evaluated. The methodology is similar to Article 101, but with two differences: (i) the default price rise is 10 per cent, (ii) duration is left at the discretion of the case officer in the particular case. **Market studies** is the one area where the impact evaluations are sometimes based on detailed ex-post analysis.

OFT also attempts a partial evaluation of its **consumer protection work**. The methodology employs data on the number of complaints pre- and post-intervention, and reductions in complaints are converted into a financial estimate by valuing each complaint at a proportion of the purchase value of the product or service concerned.

Comparison of UK, EC and USA

On setting the reported estimated benefits for these three jurisdictions alongside each other, it is immediately clear that they are not directly comparable. This is because the different authorities use different assumptions and conventions in their estimations.

It is also very noticeable that estimates vary considerably over time. One reason is that the sizes of the intervened markets can vary dramatically, e.g. when a very large market is intervened in a particular year, this can lead to lumpiness in the time series. Also, the number of closed cases may vary significantly over time. this may be due to the lengthy procedures associated with some types of investigations (especially cartels). To soften the extent of oscillation of figures over time, the UK authorities report three-year moving-averages. A third cause of the variance is due to changes in the assumptions employed by the CAs.

Comparison of assumptions

To estimate the impact of an individual intervention, information is required on the size of the market, the price increase removed or avoided and the length of time the increased price would have prevailed absent the intervention. Of these, the most challenging part of the evaluation is to estimate the magnitude of price increase and the length of time over which the intervention impacts.

For **cartels**, the first best estimate of cartel overcharge is to employ an estimate which has been thrown up during the investigation. Unlike the OFT, the other authorities assume that their intervention leads to only a 10 per cent reduction in

price. The OFT's assumption of 15 per cent seems more in keeping with the wider empirical evidence in the academic literature. On the expected duration of a cartel, had it not been detected, the CAs employ a wider dispersion of assumptions. It is tempting to draw on the evidence from existing empirical studies in deciding what would be the most appropriate assumption, but, of course, these studies are based only on detected cartels which may well differ from undetected ones (the required assumption here is how long would the cartel have continued had it remained undetected). Turning to economic theory for some intuition, my conclusion is that such is the heterogeneity of cartels, that a case-specific approach, as recently adapted by the EC, would seem to be most in line with economic theory.

For **merger intervention** it is fairly common to use simulation to estimate how price might have changed had the merger proceeded absent the CA's intervention in specific cases. Alternatively, when the price impact of a merger was estimated during the investigation, this is sometimes used. Failing this, default assumptions are made on the price impact. For example, in the past the EC assumed that the consumer savings resulting from a corrective merger decision would be as much as 10 per cent of the market turnover, while other authorities have sometimes used a far lower 1 per cent default. Turning to the academic literature, there is no obvious consensus about by how much anti-competitive mergers typically raise price. For this reason, simulation case-by-case seems to be the most appropriate approach. There is more convergence between the CAs concerning the assumed duration of the merger-generated price impact: for a lower bound, typically one or two years; but for the upper bound, the case-by-case approach used by the EC seems appropriate.

For the other areas of policy, **abuse of dominance** pose arguably the greatest challenge for assessment. If case-specific information is unavailable, various different assumptions are made about the default price rise, and the duration of the infringement. However, the lack of empirical academic work on abuse of dominance makes it difficult to assess these assumptions. Some of the CAs also sometimes assess **consumer protection**, or advocacy. For example, the FTC has reported 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.

Conceptual issues in accountability evaluations

This review raises some general issues for further discussion and research:

- Estimates are clearly very sensitive to the assumptions used. A case-by-case approach is therefore desirable when possible, but when not, there are grounds for seeking common practice on the default assumptions, and these should be firmly based on theory and evidence.

- It is reasonable to interpret the results of these evaluations as typically **conservative estimates**. While this is appropriate, especially for self-evaluation, there is a risk that CAs may be 'selling themselves short' when reporting to Government and the outside world on the benefits that they bring.

- Impact evaluations only consider the cases where the CA had decided to intervene, and **exclude cases where there was no intervention**. For example in mergers, no impact is associated with mergers that the CA rightly authorised without intervention, perhaps for efficiency-enhancing reasons.

- When impact estimates are acquired for accountability reasons, there is a danger that the CA will pursue the 'easy options' in its enforcement activities, rather than pursuing difficult cases where it might be desirable to devote more resources and establish precedents, even if these do not yield 'measurable benefits'. A related danger is that the evaluation programme itself may introduce an additional **motive for distortionary discretion** in CA conduct. For these reasons it may be better to sometimes use of an external auditor to conduct the assessment.

There are undoubtedly considerable unobserved impacts of competition policy – in particular **deterrence of anti-competitive behaviour** (see below).

Ex-post evaluation by the CAs

In the main, the above evaluations are ex-ante, but the CAs also undertake ex-post evaluations, usually on a more ad hoc basis.

These evaluations are of two broad types. **Qualitative studies** typically survey the market post intervention to identify whether the predictions at the time of the intervention have proven to be true or not – for example market entry and expansion. **Quantitative studies** more often attempt to establish causality between the intervention and, for example, a change in the market price. When the evaluation is commissioned by the CA then a mixture of these two types of method may be used.

The most notable published example is a series of commissioned reviews for the various UK CAs and relevant Government Departments. Various methodologies have been used: in-depth interviews alone; interviews with market participants coupled with simulations; DiD and a survey of market actors. Follow-up questionnaires and/or interviews undoubtedly often provide invaluable insights but are inevitably prone to a number of potential limitations: low response rates, respondent bias, the parties often have short corporate memories, and in their view, interventions can often be overtaken by other subsequent and more important events.

Other ex-post reviews sometimes focus on the perception of stakeholders regarding the CA's practice in general (e.g. the UK's Competition Commission (CC), and the European Commission (EC)). The EC also commissions some of its evaluation work to external experts (see the case studies cited below).

The techniques of evaluation

The second part of the report (Chapters 6 and 7) focuses more specifically on the alternative techniques used to evaluate impact in particular cases.

Simulation and structural models

Simulation typically involves three stages: (i) An explicit formal modelling of the nature of competition in the market, often involving a structural model derived from oligopoly theory, coupled with a particular model of the demand system; (ii) Calibration of the model with numerical values for the model's parameters, derived from direct observation, or from econometric estimation of the demand system (e.g. existing market shares, prices and extraneous estimates of demand elasticities); (iii) Using the model to assess how the equilibrium would change with and without a CA intervention. Simulation may be either ex-ante or ex-post, but most applications are ex-ante in the sense that they are based on data available at the time of the intervention – even if they are actually conducted retrospectively.

- Previous literature

The early development of simulation was firmly rooted in the academic literature and most of the seminal contributions related to merger simulation. Increasingly over the last 20 years, simulation has featured in specific merger investigation by the CAs and advisers to the parties.

- Pros and cons

The major strength of the simulation approach is its explicit use of theory to identify the counterfactual. This facilitates the 'joining-up' of the analysis undertaken at the time of the intervention with 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, 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, because there are no strong theoretical reasons to choose, e.g. the functional form of the demand curve. Simulation is better suited for some types of oligopoly models (and therefore markets) than others. It is easiest to use where competition is effected through 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. Thus, 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 its 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.

- Back-of-the-envelope (simple) simulation

Because full-fledged simulation is extremely demanding of data and research time, CAs often employ simplified versions – typically which can be approximated by the Cournot model or simple models of differentiation.

Event studies

An event study draws on financial market data to measure the effect of an 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 its share price. The methodology entails measuring any abnormal returns associated with an event (e.g. the announcement of a merger, and these are identified as the difference between the observed movement in stock valuation and that which would have occurred absent the event.

- Previous literature

For **mergers**, event studies typically examine the effect of the announcements of the merger and the CA's decision on the valuation of the merging firms and their rivals. Drawing on a robust result from most oligopoly theory, a pro-competitive merger will harm the parties' rivals, while an anti-competitive merger will benefit the rivals. This underlines the methodologies of a number of studies of the impact of merger regulation in the main jurisdictions over the last 20 years. Event studies on **cartels** typically examine the impact of dawn raids and the subsequent CA decisions on members' stock valuations. The methodology has been less frequently applied to cases of **monopoly abuse**.

- Pros and cons

The central assumption of an event study is that financial markets are efficient. If they are, this methodology yields unbiased estimates which are both "objective" and quick.

However, this assumption is questioned – at the general level - by a number of commentators. More specifically, when applied to mergers, a number of questions

must be raised: for example an increase in the share price of the parties may reflect either pro-competitive effects (efficiency gains), or anticompetitive effects (exclusion, dominance, or collusion); similarly, does a fall in the rivals' valuation reveal pro-competitive (efficiency) expectations, or does it anticipate anticompetitive (exclusion) effects? There is also the possibility that the merger may be interpreted as a signal that other firms in the same market will engage in merger activities in the near future, which would increase the market valuation of rivals, resulting in the same sign of change as with collusive and/or dominant outcomes. The pros and cons are similar for assessing cartel enforcement, although the theoretical expectations are less ambiguous. The effect of news of the investigation and the ultimate decision nearly unambiguously reduce the valuation of the parties. On a practical level, in spite of the general presumption that event studies are easy-to-use and data are easily accessed from financial databases, there will be many circumstances when appropriate data are not available – notably, for unquoted firms and/or for conglomerates and multinationals, for whom the market concerned may constitute only a small part of its aggregate activities.

Difference-in-differences

In the difference-in-differences (DiD) methodology, the basic idea is to evaluate 'performance' before and after an event (or sometimes before, during and after) in the market concerned relative to performance in another similar (control) market, unaffected by the event. Typically it is econometric, and price is tracked over time to include the pre- and post-periods in the treatment (e.g. merger) and compared against the same in the control market. Almost by definition, DiD analysis is only conducted ex-post. The time lag before it is undertaken reflects a trade-off between choosing a sufficiently long post-event period to gain a better grasp of long-term effects, and avoiding the choice of a time period which is so long that it would compromise the chances of finding a practicable control to emulate the counterfactual.

- Previous literature

The methodology has been applied fairly extensively to mergers, merger interventions and cartels; there are fewer examples where DiD has been applied to abuse of dominance cases, deterrence, or other effects of public intervention.

- Pros and cons

The difference-in-differences approach is appealing because it uses observed data on what actually happens in comparison with a control, where the event does not occur. Thus the counterfactual is not dependent on untestable, maybe restrictive, theoretical assumptions. However, it is then crucial that the chosen control group should closely correspond to what would have happened in the treatment market, absent the intervention. In practice, there is a danger that the choice of

counterfactual is constrained by 'what is out there', i.e. the best of a set of alternatives, none of which is entirely appropriate. To cite just one example, quite often, the chosen control is the rivals to the parties in a merger case, but, as shown by most oligopoly theory, rivals will also change price after a merger, and so the control will not be independent of the treatment (i.e. merging) firms.

- The before and after approximation to difference-in-differences

One of the most commonly used methods of evaluation is to conduct a simple before and after comparison of, say, price. This is strictly not a difference-in-differences since there is no control group. In effect the implication is that there is no need for a control, or that the default counterfactual is one of merely the status quo: absent the merger or intervention, the market would simply have performed identically before and after. Clearly, the results from such a methodology should be interpreted with caution – at the least, one would look for qualitative sources (e.g. interviews/questionnaires) to substantiate the choice of a no-change counterfactual.

Choosing between methodologies

In effect, much of the discussion about the pros and cons of different methodologies relates to how appropriately they implicitly describe the counterfactual. Simulation, by its nature, places the choice of counterfactual conspicuously at centre stage: a specific oligopoly model must be identified, although, of course, this does not mean that the correct model is selected. In DiD, the control plays the role of the counterfactual. Here, the choice of counterfactual is less theoretically driven, and the strength of the methodology rests on whether there really is a control (with adequate data) which is sufficiently similar. In practice, data expediency may sometimes distract attention from how closely this condition is met. While it is less common to think of the counterfactual in the typical event study, implicitly it is still there – captured by whatever comparator share price index the practitioner uses to compute abnormal returns. Here there is a trade-off between using a general index, which is less likely to be sensitive to market-specific exogenous events, and more customised sector-specific indexes, which may not be truly independent of the event at issue.

In an ideal world, the alternative evaluation methodologies would be compared by assembling a very large, random sample of cases, and attempting to apply all the methodologies to all cases in the sample. While it is likely that some CAs and advising consultancies do sometimes conduct parallel assessments during their conduct of a particular case (experimenting simultaneously with, say, an event study and a simulation), these do not appear in the published literature of course. More generally, attempting such a task across a large sample of cases would be difficult, and, in the event, has not occurred to date.

Almost as a digression, section 6.4 anticipates how such a large sample comparison might be conducted by revisiting an existing database of estimates of cartel overcharge for a very large sample of different cartels observed across different countries and at different points in time. Since these estimates were derived using a variety of different methodologies, the purpose is to identify whether certain methodologies are inherently more likely to generate larger estimates of the extent of overcharge. While the comparison does reveal some noteworthy differences, further work, controlling more carefully for sample selection, would be required before definitive conclusions can be drawn.

Example case studies of evaluation methodologies

Six representative case studies are used to illustrate the above methodologies.

Case 1: Two UK Mergers (difference-in-differences)

This study was conducted by the LEAR consultancy for the UK's Competition Commission (CC). Two different merger decisions were considered: (i) GAME/ Gamestation, a merger between two specialist retailers of video games, and (ii) Waterstone's/Ottakars, a merger between two specialist 'High Street' book retailers.

The two cases illustrate some of the generic features of the difference-in-differences methodology. The demands on data were intense. However, it helped considerably that the pre-merger data were already available from the original investigations. Nevertheless, even in these two data-rich examples, it proved impossible for a very able team who were not overly time/cost constrained to undertake any tests of the key service-quality dimension. The choice and availability of alternative counterfactuals is also a key issue. In this case, as in many other local retail cases, the existence of some local markets where only one of the parties was present pre-merger is extremely valuable. On the other hand, the use of competitors' prices as a control is potentially problematic.

Case 2: Cheese in Mauritius (back-of-envelope difference-in-differences)

This case is used as a very recent example of an in-house ex-post impact evaluation by a young competition authority for its intervention in 2010 against an abuse of monopoly. The abuse was the practice of offering retroactive rebates on Kraft branded block processed cheddar cheese in exchange for supermarket premium shelf space. The evaluation was conducted unusually soon, only one year, after the Competition Commission of Mauritius (CCM) intervention. However, this was two years after the investigation-launch and the authors judged that the impact was beginning to be felt even whilst the investigation was underway. The required data were collected from various governmental institutions as well as from the market players themselves.

This is an example of what we call a very simplified back-of-envelope difference-in-differences, in which the CCM assumed that, had it not intervened, there would have been no entry, and price would have continued to rise at the same rate as it had prior to the intervention. Clearly, such simple assumptions are really only guesses and open to question. On the other hand, they render the required analysis much less data and time-demanding.

Case 3: Pirelli/BICC merger (event study)

This study was commissioned by the European Commission (EC) and undertaken by LEAR Consulting. It concerned Pirelli's acquisition of six manufacturing plants in Italy and the UK in the power cables industry from BICC General. The data employed were taken from Datastream on the stock prices of the rivals and customers of the merging parties at three sub-events: when the merger was first proposed, then at EC's announcement that it was going to instigate a Phase 2 investigation, and finally at the clearance decision.

In many ways, this is a classic example of an event study. This particular industry should provide fertile ground for the use of an event study since many of the relevant firms are large and quoted on stock exchanges. Having said this, even in such a large scale important industry as this, only half of all relevant firms were publicly quoted, raising concerns about selection bias. Moreover, many of the quoted sample firms were multinational and diversified, raising doubts whether their stock valuations would be sensitive to a merger between two of their rivals in just two countries in only one of their lines of business. The results suggest that the merger had no impact on rivals and a mixed impact on customers. However, close inspection of the magnitudes of estimates show that they are often either implausibly high, or completely insignificant. While it is true that an insignificant effect is consistent with the merger having no adverse competitive impact, it is also consistent with a more cynical interpretation that the merger (whatever its competitive impact) would have no noticeable effect on the future profitability of such large multinational diversified firms.

Case 4: Antitrust fines in South African bread and flour cartels (event study)

This case also employs an event study to estimate the impact of fines on the stock market valuations of infringing firms involved in bread and flour cartels in South Africa (Darj *et al.*, 2011). It is a second example of evaluation from a relatively young competition authority outside of Europe. It is also far less technical, and was presumably much less time-consuming than the previous case. This case, however, is a good example of how an event study can be used successfully in a straightforward way to quantify the effects of intervention on the firm concerned – as is possible with fines. Importantly, the impact of a fine may exceed the simple cost of the fine itself because of other potential indirect effects on the firm's future profitability. There is no real complication regarding the counterfactual which is fairly obvious in the case of fines.

Case 5: Volvo-Scandia (simulation)

This case is an example of a full-fledged technically quite advanced simulation of a proposed merger between Volvo and Scandia in the European truck market. At the heart of their methodology, is the econometric estimation of the demand system of equations for trucks. This uses panel data for 16 different Member States over two years. Prices are the list prices of a base model and sales are total sales for the model range. The econometric model used is nested logit and it is necessary to include variables reflecting the different characteristics of different models of trucks. However, and this is a potential weakness of the estimation, only one truly technical characteristic was employed – the truck's horsepower. The effects of the merger were then simulated by introducing a chosen non-cooperative oligopoly model (i.e. unilateral effects) and analytically solving for (i) the equilibrium prices assuming first no merger (i.e. before) and then (ii) with the merger (i.e. after).

This case illustrates well many of the standard strengths and weaknesses of merger simulation. It forces the analyst to specify quite explicitly what is the chosen counterfactual, but it is also possible to simulate alternative counterfactuals (oligopoly games). In itself, the rigour of the simulation model also requires the analyst to form a detailed understanding of the nature of competition and demand in the market concerned. On the other hand, results can be very sensitive to alternative assumptions, and data demands can be prohibitive – often requiring approximations (e.g. here the paucity of measures of product characteristics). Finally, when conducted at the level of technical complexity employed here, simulation is clearly time- and resource-consuming and requires highly skilled econometricians.

Case 6: A review of 8 UK merger decisions (back-of-envelope simulation)

This final case also uses simulation but this time much more simplistic and less time consuming. It is a broad-based evaluation of eight different mergers, conducted by the consultants Deloitte for a consortium of UK authorities (2007). The purpose was to conduct a review of merger decisions under the relatively new Enterprise Act of 2002. The assessment was based on two approaches. The first was a qualitative survey using interviews with market participants, supplemented by questionnaires and publicly available data. The second was to conduct ex-ante simulation on as many of the cases as possible, simulating what the likely consequences of the mergers would be, but using only information that was available to the authorities at the times of their decisions. The simulations were confined to back-of-envelope simple methods which could be effected quickly and with minimal data requirements. In fact, given data limitations, simulation was possible for only four of the eight cases. In each, the product was effectively homogenous and capacity constraints were typically an issue. For this reason, the model used was a capacity-constrained Cournot model.

The study illustrates quite well what can, and what can not, be achieved by back-of-envelope simulation. Its strength is in speed and simplicity, and thus the capability of application to a number of different cases in a fairly short period of research time. Its weakness is that it can often only be applied to simple markets (here, in which competition could be described by simple Cournot models, albeit capacity constrained), or in which simple diversion ratios can be used to model product differentiation. Moreover, a danger with simulation models is the spurious sense of precision and accuracy it might encourage. The authors correctly stressed that such simulations should be viewed only as devices for showing how different assumptions about the market and the nature of competition might translate into different evaluations of the likely impact of different decisions. In other words, their role is to provide only one of a number of different aids to the decision-making.

Some important under-developed issues

In the author's opinion, there are three gaps, uncertainties and unresolved issues in the previous literatures and practices which merit future research.

Selection bias and deterrence

Inevitably, evaluation studies tend to focus on a sample of cases which are: (i) documented in the public domain, and amongst these, (ii) usually the ones where an intervention has actually occurred or been seriously contemplated, and (iii) for which evaluation is feasible. This raises various doubts about whether such a sample can provide an unbiased estimate of the full benefits of competition policy.

Some potential sources of selection bias

As stressed throughout this report, one source of bias derives from the fact that different evaluation methodologies are less practicable for some types of markets and interventions. But the issue is far wider than just this. Previous literatures include numerous other instances of why evaluation is susceptible to selection bias:

- In the empirical literature on **cartels,** it is widely understood that the samples analysed may be intrinsically biased because they are drawn exclusively from detected investigated cases, which may not be representative of the unknown population of undetected cases.

- When evaluating the rigour of **merger policy**, it is misleading to focus on *unchallenged* mergers, because even if the CA is 'lax' (inclined to Type I errors), such a sample will include many 'good' mergers alongside any incorrectly permitted 'bad' (price increasing) mergers.

- It is widely acknowledged that the beneficial **deterrent** effects of competition enforcement are likely to be considerable. It follows that any evaluation of the benefits of policy based only on investigated cases may be a serious underestimate.

In the report, a classification scheme designed to structure the various dimensions of potential selection bias is presented, and directions for future research highlighted.

How much of the iceberg lies below the waterline?

In section 8.1 of the report, I pose the hypothetical question: "What information would be needed before one would be able to extrapolate from what we know – based on investigated cases - to what we do not know – about those cases which are deterred, or undetected or un-investigated?"

Introducing three conditional probabilities: the deterrence rate (ω), the detection rate (ϕ), and the investigation rate (σ), it follows that the sample of investigated cases represents only a (perhaps very) small proportion, $(1- \omega)^*(\phi)^*(\sigma)$, of the population of all potentially relevant cases.

While this question is hypothetical, it is instructive, as a thought experiment, to speculate about the sort of magnitudes involved. To do this, we draw on a rare qualitative study of deterrence, commissioned by the OFT, which can be used to calibrate the above three parameters. Inserting these into the above expression suggests that (i) the investigated sample of mergers is only about 10 per cent of the full population of potential mergers and that detected cartels is less than 4 per cent of all potential cartels; and (ii) there are five (fifteen) times as many deterred as investigated mergers (cartels). However, it must be stressed that these calculations are very approximate and extremely speculative.

It should also be stressed that these calculations relate only to the *frequencies* of mergers and cartels – they do not necessarily quantify the relative *amounts* of deterred harm or missed opportunities. This would only be true if the expected harm of cases of observed and unobserved cases were identical. In fact, the above concerns regarding selection bias mean that there are very real reasons for supposing that this will not be the case, i.e. that the observed cases are not a random sample of the full (partly unobserved) populations.

The report surveys some of the academic literature which is most likely to be helpful in furthering our understanding of the nature and magnitudes of the various potential selection biases. However, it is evident that much future research, both theoretical and empirical, will be required.

The need for longer term studies

It is almost a truism to state that the effects of any policy change or intervention (in whatever area) are likely to extend into the long-run. It is also fairly obvious that the evaluation of long-term effects is more challenging than short-run evaluations - given the exponential rise of possible compounding effects with time, which will make it difficult to distinguish effects that were caused by the intervention from effects triggered by exogenous factors.

This is certainly true for anti-trust. The wider academic literature suggests various ways for how a specific event might trigger a sequence or chain of subsequent events – each of which might be evaluated independently, but which are in reality clearly path-dependent. For example, the literature on endogenous mergers alerts us to the possibility that, if merger A is cleared, this makes a subsequent merger B more or less likely. Similarly, in failing firm merger cases, the consequences of intervention may include subsequent alternative merger proposals by other parties as happened in the case of Airtours. There is case study evidence which suggests that sometimes when some anti-competitive practice is prohibited, firms will attempt to replace it with others. It has long been recognised that horizontal mergers may sometimes be an alternative that firms pursue once cartelisation is impossible.

In short, it matters what happens after the intervention, and 'after' should sometimes be interpreted as long-term and not too narrowly. While this point is occasionally acknowledged in the evaluation literature, the instances of long-run evaluation are rare. It is the author's opinion, that the future research programme should prioritise longer-run evaluations. One potential way of doing this is through case studies regarding the evolution of markets over, say, decades and involving a series of interventions, rather than being confined to just one-off independent evaluations of single events.

The broader impact of competition policy

Most of this report has focussed rather specifically on the impact of enforcement on consumer welfare. However, we should also be concerned with the wider impact on the economy as a whole and the macro aggregates of growth, investment, employment as well as productivity and innovation.

There are empirical studies which explore these areas, and some of the literature is reviewed in the report. However, in truth, it is a thin literature of variable quality. There are two broad strands: works that examine the impact of *competition* on productivity, growth, etc., and research that attempts to measure more specifically the impact of *competition enforcement* on the same factors. Taken overall, research in the first strand offers broad confirmation that competition helps fuel economic growth, and this is through at least three different channels: (i) exerting pressure on firm management to reduce inefficiencies; (ii) by fostering innovation, and (iii) by

guiding a better allocation of resources to more efficient firms. However, many studies in this strand employ ambiguous or questionable measures of the degree of competition.

Turning to the second strand - the effects of competition enforcement – a key challenge is to find appropriate measures of competition enforcement or competition policy. There have been recent advances on this issue with various indexes of the efficacy of competition policy being suggested. These typically rely on institutional and enforcement characteristics such as the degree of formal independence of the CA, the separation of adjudicatory and prosecutor functions within the CA, the scope of investigative powers of the CA, the level of enforcement (size of sanctions, size of budgets and resources, etc.), and the seriousness of sanctions that the CA can impose. There is now some evidence that competition enforcement does contribute to economic growth. However, more work is needed in this area before we can claim any robustness for this result.

Conclusions

Why should CAs evaluate their own work?

I believe that there are at least three strong reasons why CAs should evaluate their own work:

First, if competition authorities are to maintain and even increase the funding they receive from government, it is important that they are able to show that they provide value for money. It is my view, and experience, that any well-functioning authority will be able to do just this from a comprehensive and public evaluation of its activities, even making the most conservative assumptions. This view is based on two (in my opinion) very likely assumptions: (i) that competition considerably enhances consumer welfare, and (ii) most competition authorities do indeed foster competition through their various activities.

Second, at a more micro level, regular and fairly comprehensive evaluation, should provide an essential ingredient in internal resource allocation decisions within the authority. When allocating its budget between its different activities (merger control, detection of cartels, monitoring of dominant firms), it is important to know what is the likely marginal product of spending an additional euro in each area. More generally, allocational decisions – say between detection and deterrence - should be informed by some notion of the potential pay-offs.

Third, in any busy competition authority, there is an inherent danger that once a decision or intervention has been made, it is consigned to a file which is then never again opened. Ex-post evaluation provides probably the best means for ensuring that the authority becomes aware of 'what happens next' in a market. Of particular importance is the question of how markets react to, and perform, after

interventions. Moreover, evaluation provides a valuable check of whether the assumptions, analysis and decision making taken at the time of the decision are vindicated by what happens in subsequent years.

However, there are two caveats. First, we should not become slavishly attached to the precise estimates derived in any evaluation. Evaluation will always be an imprecise subject, based inevitably on assumptions which are approximate and sometimes unverifiable. Evaluation is there to guide rather than dictate. Second, an over-reliance on impact estimates can distort allocational decisions within an authority. The danger is that certain activities, especially deterrence, may go relatively neglected because their impact is more difficult to measure.

How should CAs conduct evaluation?

Two broad themes run though this report – deriving estimates of the aggregate impact of an authority's activity, and then more specifically the techniques which can be used to evaluate individual cases. It is my view that the competition authority should do both – regularly (perhaps annually) evaluating its overall impact and also devoting some resources to in-depth ex-post evaluations of at least some specific cases.

There are a number of issues and decisions which would need to be resolved in deriving aggregate impact evaluations – which rules of thumb to use for cartel overcharge and duration etc., which areas of activity to evaluate? Here, the existing practices of OFT, DGCOMP, FTC and DoJ might be adopted, but, in the absence of any internationally agreed best practice, this need not be essential.

Turning to the specific techniques of evaluation, the report suggests that there is no such thing as an ideal technique. Simulation, event studies, difference-in-differences, before-after and more qualitative surveys all have their strengths and weaknesses, and each is more suited to some cases than others. Perhaps the strongest message to draw from this report is that sufficient time and attention should be devoted to formulating an appropriate counterfactual (or alternative counterfactuals). Sometimes, the nature of the counterfactual will dictate which evaluation technique should be used.

A third broad recommendation is that the authority should adopt a portfolio approach to its evaluation activities. Thus, there is much to be said for employing a mix of ex-ante and ex-post evaluations - as do the other authorities most frequently discussed in this report. Similarly, some of the evaluation should employ 'back-of-envelope' techniques, while some might be more thorough and technically advanced. There are advantages in commissioning outsiders (consultancies and academics) to do some, but not all, of the research. The obvious problem with in-house evaluation is the suspicion that it will be seen as self-justifying and without objectivity. On the other hand, it is important that the authority should develop its own technical capabilities.

What are the challenges?

Beyond the three areas where I think further work is generally required, it should be stressed that the evaluation project is still very much 'work in progress'. Evaluation in some areas of policy is still under-developed, e.g. advocacy, education and Article 102. Little is known, in a comparative sense, about the predictive performance of the different techniques. Here, the challenge is to widen the scope of evaluation without unduly increasing the cost. As is true for many policy areas, there is an important role for international cooperation and diffusion of experience and skills.

1   Introduction

The aim of this report is to provide a survey of the approaches and specific methods which are used to assess the economic effects of a Competition Authority's work, paying particular attention to the advantages and disadvantages of those specific methods.[1] This first introductory chapter sets the scene by providing initial background definitions and classifications. These include the purposes and scope of evaluation; a taxonomy of methodologies; the difference between ex-ante and ex-post evaluation; and the central importance of the choice of the counterfactual in any evaluation exercise.

1.1   The purposes of evaluation

Evaluation is undertaken both by the Competition Authorities (CAs) themselves and independent academic economists and lawyers. It can take many forms, but for presentational convenience these can be grouped into the following four broad categories.

1.1.1 Accountability

Increasingly, there is an obligation on CAs around the world, and to their Governments, to quantify the aggregate benefits of competition policy (measured perhaps by increased consumer surplus). The aggregate estimate is then sometimes assessed against some pre-specified target to judge whether the CA has met its required objectives. These impact evaluations typically involve using some of the methodologies described below.

1.1.2 Assessment of specific policies

Sometimes, the purpose is to evaluate a specific area of policy, e.g. merger control or cartel enforcement, or of the success of a particular intervention, e.g. a prosecuted cartel, or prohibited merger. Again, this may be undertaken internally by the CA itself or commissioned to an outside consultant or academic. Here the objective is perhaps to check on the quality of the CA's own decision-making – the validity of its assumptions, rigour of its analysis and data collection etc. It may also be used to help set future internal priorities and allocate resources.

---

[1] This report draws on some of the author's own recent work, such as a critical review of the methods used by the UK's Office of Fair Trading (Davies 2010), and two academic articles: Davies and Ormosi (2010 and 2012).

1.1.3 Assessment in the broader academic literature

In the academic literature, policy evaluation appears in a multitude of guises. Sometimes it is designed as an independent check on the performance of the CAs. More often however, the immediate objectives of the research may be more academic, focusing primarily on the development/testing of theory and/or empirical techniques, with the specific anti-trust case providing a well-documented and experimental setting. Nevertheless, many of these studies have helped in technical developments of the evaluation methodologies used by CAs.

1.1.4 Estimating damages and fines

In any private damages case, or indeed when a CA sets a fine (where applicable), this entails an evaluation of the harm caused by the firms concerned, and therefore, by implication, the gains resulting from the CA intervention which removes that harm. Expert opinions in such cases invariably draw on methodologies similar to those described in this report.

The question of who undertakes the evaluation, and how deeply, raises important issues of the potential distortion of incentives and priorities within the CA. We will return to this later in the report.

1.2 The methodologies of evaluation

There is a rich variety of evaluation methodologies which have been used in the literature to date. In this report, we shall mainly focus on what we refer to as the three key quantitative approaches: simulation, event studies and differences-in-differences. These will be discussed at greater length in Chapters 6 and 7, but as they are also frequently referred to in Chapters 2-5, brief preliminary definitions will be helpful here.

1.2.1 Simulation of structural models

Simulation models typically entail modelling the nature of competition in a market, calibrating the parameters using real world information (sometimes estimated econometrically) and then assessing how the intervention will change the equilibrium relative to what would have happened without the intervention. In the academic literature, these models are often highly technical and complex, but they are also often used for evaluation by CAs in more simplified forms.

1.2.2 Event studies

Event studies use the financial markets' assessment of the impact of an event. In the context of say merger enforcement, the events in question are the initial announcement of the merger, the CA's announcement of investigation, and then its subsequent decision. The effects are quantified by comparing movements in stock prices – both of the parties and their immediate rivals – with movements in more general stock price indices.

1.2.3 Difference-in-differences

The difference-in-differences (DiD) methodology involves a comparison of, say, prices before and after an event, (e.g. a merger or a dawn raid), relative to some other real world control, i.e. a similar market without the event, or within the same market for firms not involved in the event. Again, there are simplified equivalents of this involving just before-after comparisons without a control market, sometimes in multivariate form with control variables included.

We shall interpret each of these categories fairly broadly to encompass much simpler methods which can be seen as special forms of one or the other of the categories. Like any taxonomy of this sort, it is bound to entail a certain degree of arbitrariness, but this is a price worth paying for the expositional convenience it brings. There are a number of alternative taxonomies used in the recent literature, e.g. Bergman (2008), Buccirossi *et al.* (2006), OXERA (2009).[2] The first two are sufficiently similar to our own for us to merely refer the interested reader to the original papers, but OXERA (2009) is sufficiently different to merit a short summary. The authors group the existing methods and models into three broad types:

- Comparator based, which includes (i) DiD, (ii) cross-section (across firms, markets or countries) econometrics or averages, and (iii) time series econometrics or interpolation for during, before and/or after.

- Market structure based using industrial organisation (IO) models, especially structural models.

- Financial analysis based, including event studies but also other more descriptive techniques from the accountancy/finance literatures.

---

[2] OXERA (2009) is a commissioned study report for the European Commission on quantifying antitrust damages. There are obvious parallels (although by no means always exact) between the savings that CAs achieve for consumers from removing anti-competitive practices and claims for damages against defendants responsible for inflicting losses on their customers and rivals and/or suppliers.

This taxonomy is more exhaustive than our own, but the three categories correspond fairly closely to our own categories.

Other methodologies

There is also a variety of other, mainly qualitative (with the exception of the first), approaches which we shall refer less frequently to below.

1.2.4 Aggregate international econometric studies

International econometric studies are more aggregate studies, employing cross-country or panel data in order to identify the effects of competition policy on macro or sector aggregates, e.g. price-cost margins, GDP, productivity. These are discussed and well summarised by Bergman (2008), who provides some of the key references. Necessarily they confront daunting measurement issues, often relying on the construction of subjective indexes of the 'severity of competition policy' which are allowed to differ across countries. Other econometric problems are familiar in any international comparisons based on production functions or related concepts, e.g. identification, simultaneity and the requirement that the underlying functional forms are stable across countries. In general, this report does not cover studies of this very aggregate type, but the final section does include a slightly more detailed survey.

1.2.5 Follow-up surveys

It is quite common for CAs to undertake or commission reviews of their previous cases, (especially mergers). For example the UK Competition Commission (CC), together with the Office of Fair Trading (OFT) and the relevant Government Department, regularly conducts (or commissions) retrospective merger reviews, reflecting upon market developments following merger interventions. These reviews often use follow up questionnaires and/or interviews with the interested parties and related firms. These can provide invaluable insights and are discussed further below in Chapter 5 on ex-post evaluation by the CAs.

1.2.6 Expert commentaries on specific cases

Book collections of expert (economic) commentaries on specific, often well known anti-trust case studies, such as Kwoka and White (2004) and Lyons (2009) for the US and EU respectively, can also be viewed as contributing to the evaluation literature, but for obvious reasons, it is difficult to generalise from a set of heterogeneous non-

randomly selected small samples, especially with respect to the evaluation methodologies used (if any).

1.2.7 Court decisions

Some studies have assessed the quality of CA decision-making by the frequency of court appeals and/or the success rates in those appeals (see Bergman (2008, p.389-91) for a brief survey). For certain purposes this may be a valuable extra source of (presumably well-informed and objective) evaluation, but obvious limitations include the likelihood of selection bias, and the fact that court decisions will sometimes involve judgement on the correctness of legal process rather than economic substance.

1.2.8 Surveys of peer/practitioner opinions

Certain high-profile annual reviews provide an alternative approach, based on peer review evaluation of the performance of different CAs around the world, e.g. the Global Competition Review and OECD country reviews. These enable international comparisons over time at the aggregate level, but lie outside our current remit, being based on subjective opinion rather than quantitative methodologies.

1.3 Level of sophistication or technical complexity

Depending on the purpose, the budget, and who undertakes the evaluation, the sophistication of evaluations can vary considerably. Broadly speaking, there are three degrees of technical complexity:

- **Full blown, 'state of the art' applications:** these involve sometimes very detailed Industrial Organisation theory and advanced econometric methods. Many of the academic case studies fall into this category, as do some ex-post case evaluations conducted by the CAs or commissioned reports.

- **'Back-of-envelope'**: these are much simpler versions, especially of simulation or differences-in-differences, where the practitioner employs easy and quick to compute approximations to the above. A typical 'back-of-envelope' evaluation might require only days or hours to conduct, while the full blown equivalent might require months (or even years in some academic cases).

- **Default**: these are employed, especially in the accountability exercises described in Chapters 2-4, where a large number of cases need to be evaluated fairly quickly and cheaply and hard data are sometimes sparse.

1.4   Ex-ante versus ex-post

Evaluation may be either ex-ante or ex-post. Although the meaning of this dichotomy is superficially obvious, these terms are sometimes misunderstood. By definition, 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 coordinated effects, if cleared.

In practice, however, both ex-ante and ex-post evaluations are typically conducted retrospectively, and the meaning of the distinction relates mainly to the nature of the data used in the evaluation. Thus, an ex-post evaluation uses the information of 'what actually happened next', i.e. after the intervention took place. To evaluate impact, 'what happens next' is then compared to what would have happened, had the intervention not taken place.[3] When the perspective is ex-ante, but conducted after the event, this generally only employs the information that would have been available at the time of the policy decision. In that case, the practitioner projects forward comparing of what would happen with and without the intervention.

In general, ex-ante evaluation is simpler to conduct given that it makes less demand on data – employing information which should be available at the time of the policy decision. Ex-post evaluation, on the other hand, can only be conducted some years after the intervention when accurate data becomes available on what actually did happen. It is also confounded by the likelihood that what happens next will not be exclusively due to just the intervention. This is discussed further at greater length in the specific context of merger remedies by Davies and Lyons (2007 pp.106-7).

Usually, differences-in-differences, event studies and qualitative studies are only applied in ex-post form, while simulation may be used for both ex-ante and ex-post evaluations.

1.5   Policy areas

Most of this report will be largely confined to the evaluation of competition policy, defined to include all activities undertaken by competition law enforcement agencies. However, at the end of this section we briefly discuss the related areas of de-regulation and liberalisation which are more typically the territory of other areas of government including specific sector regulators.

---

[3] This comparison is not without difficulties. For example, when retrospectively evaluating the impact of a particular merger intervention, should actual post-merger prices be compared with those projected from the simulation model, or should these projections be adjusted for any post-intervention exogenous shocks, say a demand shock? Estimating the separate impact of such exogenous shocks is typically not straightforward

Table 1.1 provides an assessment of the extent of the existing literature for each component part of competition enforcement, distinguishing between the academic literature and CAs' evaluation reports.[4] It also classifies by the extent to which the three key methodologies have been applied in these parts.

Table 1.1   Extent of evaluation literature by broad area of competition policy

	Methodology	Academic Literature	CAs
Enforcement:			
Merger control	Simulation	Extensive	Extensive
	Event Studies	Extensive	Some
	DiD etc	Extensive	Some
Cartels/Article 101	Simulation	Some	Few
	Event Studies	Some	Few
	DiD etc	Some	Some
Abuse/Article102	Simulation	Few	Very few
	Event Studies	Few	Very few
	DiD etc	Few	Very few
Non-enforcement:			
Advocacy	None	None	None
Compliance	None	None	None
Consumer Education	None	None	None

Source: Author's assessment

Thus, most of the previous literature has concentrated on mergers and cartels. The academic literature is most developed and extensive for mergers, and includes highly influential contributions on simulation, but also event studies and DiD. In turn, these have had a strong impact on the evaluations undertaken by the CAs themselves, and it is commonplace for the major CAs to apply simulation models in their own evaluations. There is also an extensive literature on cartel overcharge,

---

[4] This assessment is probably uncontroversial and is broadly consistent with the emphases in literature surveys by other commentators, for example, Bergman (2008) and Werden (2008).

with ad-hoc versions of DiD being more frequent. CAs sometimes employ the latter in their evaluations of the consumer benefits from cartel-busting.

Since mergers and cartels feature extensively throughout the remainder of this report, the remaining preliminary comments here are confined to the other areas where evaluation has been much less common.

1.5.1 Article 102 (Abuse of dominance)

Our judgement that evaluation of cases involving abuse of power (Art 102) has been far less common than for mergers and cartels is inevitably impressionistic, but the following fragments of evidence strongly suggest that it is valid. Davies's review (2010, Table 4.2) of the OFT's Impact Estimates, reveals that for 2006-2009, while the OFT were able to evaluate 20 mergers, only one case of Article 102 enforcement was assessed. OXERA, in its report for the EC (OXERA, 2009, p.16) suggests that "There have been relatively few cases of exploitative abuse of dominance found by competition authorities or courts, either at the EU level, or in the Member States". In the US, Werden (2008, p.446) suggests that "Non-merger civil enforcement accounts for relatively few cases and for far less consumer savings than either criminal or merger enforcement".[5] Similarly, within the academic literature, relevant studies are relatively few. This is probably symptomatic of a more general scarcity of empirical IO work in the broad area of Article 102. For example, Slade (2008) concludes her survey of the empirical literature on the effects of vertical restraints (p.28) by suggesting that "Perhaps the most important lesson that can be learned … is how scant that evidence is, especially when compared to the amount of theoretical research."

There are two obvious explanations for this. The first is, simply, that CAs bring relatively few cases of Article 102 abuse to fruition, and the second is that evaluation is unusually difficult to conduct in these cases. Werden (2008, pp. 442-3) suggests that, certainly, the latter is true: "In assessing the effects of antitrust enforcement, cases involving exclusionary conduct present the greatest challenge. The effects of potentially exclusionary conduct are apt to be subtle and can be experienced long after the conduct itself." He points to the difficulties in establishing the extent to which rivals are harmed and the impact on consumers, and explains that delicate trade-offs are involved where the practice may entail an element of efficiency enhancement. He also identifies predatory behaviour cases as particularly problematic, given the need to quantify short-run and long-run impacts of opposite directions.

---

[5] He defines non-merger civil enforcement as "single-competitor exclusionary conduct, vertical restraints and agreements among competitors other than hard-core cartels or mergers".

1.5.2 Advocacy

Typically, competition authorities allocate a significant proportion of their budget to advocacy activities. An International Competition Network (ICN) (2002) study on competition advocacy reports that, amongst those countries that felt able to quantify the resources they devoted to advocacy, almost one third reported that this was as high as between 20 and 30 per cent of their budget. Despite this, there has been little research attempting to quantify the impact of competition advocacy. The reasons why are entirely understandable, once one attempts to define the term. The ICN defines advocacy as:

> "those activities conducted by the competition authority related to the promotion of a competitive environment for economic activities by means of non-enforcement mechanisms, mainly through its relationships with other governmental entities and by increasing public awareness of the benefits of competition."

Inevitably, much of this activity is general and intangible in nature, with specific effects that are not at all amenable to measurement. In the circumstances, it is inevitable that most of the work that has been done is predominantly qualitative. For example, the OFT (2010b) conducted a survey of UK government officials across various government departments asking how far its competition related advice was taken into consideration and influenced policymakers. Judged on the responses received from 43 respondents, it seems that advocacy did indeed have a significant impact – leading to important changes in policy approaches in half of the cases, and changes in objectives in one quarter. However, this study did not attempt to quantify this impact. Other authorities who have also carried out detailed surveys to assess the impact of their advocacy activities include the Federal Trade Commission's Bureau of Competition (FTC) and the Italian Competition Authority.[6]

Evenett (2006), in a thought-provoking commentary questions the real value of such qualitative studies – if somewhat over-critically. However, it would be wrong to merely dismiss qualitative work since it can often establish that advocacy can sometimes have very real pay-offs, even if they are difficult to measure.

The OFT study also describes three specific case studies which illustrate quite well that (i) advocacy can have important positive impacts, but that (ii) they are difficult to measure. The first case study regarded the advice given to the Ministry of Justice, warning against licensing regulations that would have posed a very 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. In

---

[6] The FTC study is referred to in Majoras (2005). For the Italian study see Arisi and Esposito (2007).

principle, it would be possible to quantify the impact, although this particular study did not do so, I suspect, because of cost and time constraints. None of the cases would be simple to evaluate and each would inevitably involve speculation. 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. An obvious question is whether the methods used for assessing merger and cartel enforcement could be applied also to measuring the impact of advocacy. A particular problem would be to identify the right counterfactual. For instance, in measuring the impact of advocacy that resulted in the dropping of a piece of draft legislation which would have had 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.[7] It would also be difficult to identify what version of the proposed legislation would have been accepted in absence of successful advocacy efforts.

None of this is to deny that some sort of evaluation would be possible and, in my opinion, desirable. However, it would be wise, in these circumstances, if the evaluation was run under a number of different scenarios.

1.5.3 Education: compliance by firms

Encouraging business compliance with competition law is a potentially important dimension of CA activity. However, attempts to measure the efficacy of compliance programmes per se are rare.[8] In one example, OFT (2010a) presents some survey findings designed to understand what motivates business compliance and includes examples of their compliance activities. Responding firms mentioned reputational damage, financial penalties, individual sanctions (e.g. risk of criminal proceedings) as key motivations for compliance. Compliance is obviously linked to deterrence, which we discuss more generally in later sections.

1.5.4 Education: consumers

Much consumer education relates to consumer protection law and falls outside our remit. However, insofar as better educated consumers should be less frequently exposed to asymmetric information, reduced market imperfections might enhance

---

[7] It may be equally hard to show that a proposed piece of legislation was prevented as a result of advocacy rather than other political considerations.

[8] The scarcity of evidence also reflects the fact that in many countries the CA is not allowed to carry out compliance activities.

competition between firms. Again, this is an under-researched area, but some CAs are beginning to work on evaluation in this area.[9]

1.5.5 Related policy areas – deregulation and liberalisation

The main focus of this report is on 'competition policy proper', relating to the enforcement of competition law. However, there are other policy areas which should also be briefly mentioned. In particular, liberalisation and privatisation (and potentially those advocacy activities of CAs that may trigger liberalisation) should be, and is, also evaluated – quite often with more emphasis on the long-term impacts. These are discussed below in section 8.2.

1.6   The central role of the counterfactual

A central issue running throughout most antitrust analysis is the choice of the counterfactual, i.e. what would have happened had some event, practice, policy intervention not occurred? This is not just a concern for academic economists. It also occupies the attention of the courts and lawyers, for example when attempting to quantify damages.[10]

Equally, it follows that any particular quantitative methodology used to evaluate a specific case must necessarily entail a choice of counterfactual, even if it is sometimes only implicit. The choice of counterfactual has both conceptual and empirical dimensions – which counterfactual is theoretically most tenable, and how do we calibrate it with plausible estimates of key parameters?

We return to this issue in section 6.4 below when comparing different methodologies.

1.7   Structure of the report

The remainder of this report has seven chapters arranged in three broad parts. The first part includes Chapters 2-5 relating to 'evaluation for accountability'. It describes those evaluations, which are designed to be comprehensive, in that they cover all (or most) of a CA's interventions in a given year. Chapter 2 begins with the UK's Office of Fair Trading (OFT) which has been very active in this area. Chapter 3 provides a comparison of such evaluations for the three jurisdictions most prominent in this area: the US, EC and UK. Chapter 4 considers some of the

---

[9] See OFT (2011c) p.30 fn.45, and OECD (2002)

[10] For a useful cross-discipline of the use of counterfactuals in antitrust and mergers, see Colley and Marsden (2010)

conceptual issues and problems raised by such evaluations. Chapter 5 turns to the use of ex-post evaluation by the CAs.

The second part of the report turns more specifically to the individual quantitative methodologies used in evaluations of all types. Chapter 6 provides a survey of the previous literatures on each of the three main methodologies (simulation, event studies and differences in differences) and discusses their relative advantages and disadvantages. Chapter 7 presents six specific cases to illustrate how these methodologies have been used for evaluation.

Finally, Chapter 8 points to some of the areas in which the author believes further research is most required, and Chapter 9 concludes.

2   The OFT's Annual Positive Impacts

This and the following three sections focus on (mainly) aggregate evaluations of a CA's activity - what was referred to as the 'accountability' purpose in section 1.1. The term 'aggregate' indicates that the evaluation is comprehensive across the broad range, or most of, the CA's activity. Thus, this will typically involve scores of cases in a given year, rather than just focusing on a particular case intervention.

CAs have been publishing information relevant to their impact for a long time. However, traditionally, this was limited to descriptive statistics and reports on the number of cases pursued, and/or the number of successful cases. Increasingly, this reporting obligation has begun to extend to more sophisticated measures, where the CA has quantified and reported on the aggregate benefits of competition policy (measured perhaps by increased consumer surplus). This aggregate estimate is then assessed against some pre-specified target to judge whether the CA has met its required objectives.

This section begins by describing the impact evaluations undertaken by the UK's Office of Fair Trading on an annual basis – the most recent example being year 2010/11. [11] OFT is highlighted for two reasons: (i) the author has worked with and advised the OFT's impact evaluations over a number of years, and is therefore very familiar with them at first hand; (ii) of all the authorities around the world, the OFT has been the most transparent and prolific in publishing details of its work in this area. Chapter 3 will then introduce the other jurisdictions who conduct and publish evaluations most similar to the OFT: DoJ (Department of Justice) and FTC (The Federal Trade Commission's Bureau of Competition) in the US and the European Commission (EC). These are the authorities for whom methods of impact evaluation are most developed, at least as well as can be judged by documents in the public domain. Some other jurisdictions around the world have also attempted overall assessments which estimate the consumer benefits of competition enforcement, but these are only sporadic and less regular. Davies (2010, Appendix B) provides a brief survey of these. Chapter 4 then enumerates some of the most important conceptual issues raised in accountability studies of this sort. Chapter 5 describes the use ex-post evaluations by the CAs, which are usually less comprehensive and more case-specific.

Each year the OFT publishes an annual report which quantifies the value of the welfare benefits of its interventions compared to how much it costs the UK taxpayer. While this evaluation would ideally assess the benefits of all the cases it handles, for practical reasons, it is only possible for a large number, but not all, cases. There are six defining features of the methodology:

---

[11] The most recent at the time of writing is OFT (2011c). See also Davies's evaluation (2010).

- Estimates are in terms of consumer benefits, reflecting the fact that competition enforcement should be guided by the consumer welfare standard.

- Estimates are generally ex-ante, but with some ex-post cases.

- Estimates are deliberately 'conservative'.

- It is assumed that no intervention can have a negative impact.

- Estimates are presented as 'point' estimates, rather than as a range of plausible values.

- Most areas of policy are included, but, to date, consumer education and deterrence are two major exclusions, pending further work.

Table 2.1   Estimated average annual consumer savings and OFT costs 2008-11

	Benefits (£mn)	Partial Costs (£mn)*	Benefit/Costs
Competition enforcement (Art 101&102)	83 (25%)		
Merger control	90 (28%)		
Market studies, reviews of orders & undertakings	117(36%)		
Consumer protection enforcement	36 (11%)		
Totals	*326*	*34*	*9.59*

* Partial costs are the costs attributable to these activities (no disaggregation across areas published). Including costs attributable to OFT's other activities, the total rises to £48mn. If total costs are used as the denominator, the ratio falls to 6.8, however, this would imply, incorrectly that the other activities led to zero benefits.

Source: OFT (2011a) Table 1.1.

Table 2.1 summarises the OFT's estimates of consumer savings for 2008-2011. As can be seen, in aggregate the ratio of benefits to costs (9.59) is nearly double the target of 5:1 it has agreed with the UK's Treasury. Nearly all of these benefits derive from remedying/prohibiting otherwise anti-competitive mergers, breaking cartels and market studies. Very briefly, the methods and assumptions it uses in each of these areas is as follows.

2.1.1 Mergers

The OFT relies very heavily on simulation. For any particular merger the model is chosen, as appropriate, from three candidates – Cournot for homogeneous products and PCAIDS (Proportionally-calibrated Almost Ideal Demand System) or ALM

(Antitrust Logit Model) for differentiated product industries.[12] For vertical mergers, simulation is done for each stage separately. The estimates of consumer benefit resulting from this methodology represent 'conservative point-estimates', in the sense that 'any assumptions made to run the model are conservative.' 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. Savings are assumed to last for two years. Thereafter, market correction (e.g. entry) is assumed to wipe out any anti-competitive consequences. 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 OFT uses a default assumption. This is that the consumer savings as a proportion of turnover are equal to the mean of the lower bound ratios across all the mergers which could be simulated in that year.

2.1.2 Article 101: Cartels

Where possible, estimates are based on information from case teams, but where this is not available, OFT applies "rules of thumb that are consistent with international best practice and recent academic research":

- *Future cartel duration prevented*. Absent intervention, the cartel would exist for a further 6 years,

- *Affected turnover*. This is assumed to be only the turnover of the infringing parties. This is conservative, to the extent that a high cartel price allows competing outsiders to free-ride by raising their own price up towards the cartel's price.

- *Extent of the avoided price-raise*. This is estimated if possible during the investigation. Where not, 15 per cent is assumed.

2.1.3 Article 102: Abuse of dominance

There have been typically very few Article 102 cases evaluated. The methodology is similar to Article 101 evaluations, but with two differences: (i) the default price rise

---

[12] Increasingly in recent years this 'simulation' is based on estimates undertaken at the time of the investigation of the likely Upward Pressure on Prices (UPP) or related techniques (OFT 2011d).

is 10 per cent, (ii) duration is left at the discretion of the case officer in the particular case.

2.1.4 Market studies

Again, the use of conservative assumptions is stressed, but this is the one area where the impact evaluations are not always ex-ante. It is explained that subsequent monitoring of market developments may lead to revised assumptions and updated impact estimates, and where possible, use is also made of ex-post impact estimates.

2.1.5 Consumer protection

At the heart of the methodology here are data on the number of complaints against targeted traders, pre- and post-intervention. Any reduction in 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. That proportion is calculated using a standard formula which has been previously estimated from survey data on consumer detriment relative to purchase price. This is then grossed up to allow for additional benefits of (i) consumers who had not complained against the trader concerned, and (ii) consumers of other firms who were also engaging in similar practices who, it is assumed, are now persuaded to also cease the practice.

3   Comparative Survey of UK, EC and USA

Apart from the OFT, the other major other jurisdictions in which impact estimations are regularly reported are the USA (DoJ and FTC) and the European Union, by the European Commission (EC). Table 3.1 displays the reported estimated benefits (also shown as a percentage of GDP) for these three jurisdictions. The figures are expressed in terms of consumer benefits or consumer savings, which reflects the CAs' approach to the general objectives of competition policy. It must be stressed that the purpose of this table is not to compare the performance of these authorities, but to illustrate the sort of magnitudes and to demonstrate some of the difficulties in interpreting these estimates. The figures were collected from CAs' respective annual reports[13], and to avoid excessive repetition, the detailed references of individual items are not referenced separately in the following.

One immediately noticeable feature of these figures is the variability over time in some instances. One reason for this is that the sizes of the intervened markets and the scope of the cases can vary dramatically, e.g. when a very large market is intervened in a particular year, this can lead to lumpiness in the time series. Also, the number of cases may vary significantly over time. Fluctuations in the annual number of closed cases may be due to the lengthy procedures associated with some types of investigations (especially cartels). For example a relatively large number of on-going cartel investigations in the US DoJ in 2010 would mean high agency activity, but only low measured impact in 2010, followed by higher reported impacts in subsequent years. The EC displays a more evenly distributed impact (over time), which probably implies smaller variance in the relevant market sizes and/or shorter procedures. To soften the extent of oscillation of figures over time, the UK authorities report three-year moving-averages. Another cause of the variance in the estimates over time is changes in the assumptions employed by the CAs, as discussed further below. For example, the large increase in estimated impact in cartels in the EC between 2009 and 2010 can be explained by changes in the authority's changed assumption about the cartel life-span assumption (see Table 3.2).

---

[13] Unless otherwise mentioned, the sources are EC (2010), USDoJ (2012), USFTC (2011) and OFT (2011a).

Table 3.1   CA estimates of annual consumer benefits from interventions[14]

	Cartels		Mergers		Other antitrust	
	Consumer saving	% of GDP ×10^{-4}	Consumer saving	% of GDP ×10^{-4}	Consumer saving	% of GDP ×10^{-4}
EC (billion EUR)						
2010	7.2 (7)	5.87	4.2-6.3(16)	4.28	NR (58)	–
2009	1.2 (6)	1.02	5.6 (16)	4.77	2.0 (54)	1.70
2008	1.7 (7)	1.36	5.5 (24)	4.41	4.3 (111)	3.45
2007	3.8[15] (8)	3.06	NR (23)	–	NR (133)	–
USDOJ[16] ($ billion)						
2010	0.05 (60)	0.03	0.19*(19)	0.13	0.19*(4)	0.13
2009	0.60 (72)	0.42	1.02 (12)	0.72	0.02 (2)	0.01
2008	0.02 (54)	0.01	0.48 (16)	0.33	0.05 (4)	0.03
2007	0.56 (40)	0.41	0.15 (12)	0.11	0.02 (2)	0.01
USFTC ($ billion)						
2010	N/A	–	0.59 (16)	0.40	0.51 (6)	0.35
2009	N/A	–	0.79 (13)	0.55	0.19 (7)	0.13
2008	N/A	–	0.36 (13)	0.25	0.03 (4)	0.02
2007	N/A	–	0.81 (20)	0.58	0.08 (11)	0.06
UK CC+OFT (£ billion)						
2009/10	0.083	0.57	0.235 (11)	1.62	0.083	0.57
2008/09	0.083	0.59	0.229 (15)	1.64	0.083	0.59
2007/08	0.083	0.57	0.309 (14)	2.13	0.083	0.57

NR denotes not reported

* Merger and other civil antitrust combined in 2010.

** Cartels and other antitrust combined for all years. Sources: see footnote 14

3.1.1 Comparison of assumptions

To estimate the impact of an individual intervention, information is required on: (i) the size of the market concerned, (ii) the price increase removed or avoided and (iii) the length of time the increased price would have prevailed absent the intervention. Of these, market size is the easiest to estimate. This information is normally available for the CA and can be easily recalled for the evaluation process. It is easy to see that when a conduct impacts on a large amount of commerce in the relevant markets, the implied transfer from consumers to manufacturers becomes substantial. As mentioned above, for this reason the size of the relevant market is likely to explain much of the variance in the estimated impact.

---

[14] The number of cartel convictions, merger interventions (or challenges), and other antitrust interventions are displayed in brackets. The number of challenged merger cases for the USDOJ was collected from the Hart-Scott-Rodino Annual Reports for Fiscal Years 2007-10. The number of cartel cases and the number of non-merger civil cases filed by the USDOJ was collected from USDOJ statistics (http://www.justice.gov/atr/public/workload-statistics.html). The number of challenged merger and non-merger interventions by the US FTC was collected from the FTC's Performance and Accountability Reports. For the EC, the number of interventions was collected from EC merger statistics (http://ec.europa.eu/competition/mergers/statistics.pdf) and EC cartel statistics (http://ec.europa.eu/competition/cartels/statistics/statistics.pdf) and the number of abuse of dominance cases closed was collected from Global Competition Review reports.

[15] Using an average overcharge of 20%-34% from Bolotova and Connor (2006).

[16] The number of cases refers to the total number of criminal cases filed.

The more challenging part of the evaluation is to estimate the magnitude of price increase from the investigated conduct and the length of time it would have prevailed absent the intervention. Depending on the case type, CAs often use assumed default values for these two factors in their evaluations. The following briefly describes the methodologies used in the three jurisdictions. Unless otherwise stated, the figures reported were collected from the respective CAs methodology documents.[17]

Table 3.2 summarises the assumptions used in the evaluation of the impact of *cartel investigations*. As for the OFT, these assumptions are only used if there is no hard figure on overcharge thrown up during the investigation. On affected consumers (i.e. size of market), it is often assumed that the cartel would only affect the infringing parties' turnover.[18] Unlike the OFT, the other authorities assume that the CA intervention leads to a 10 per cent reduction in price. As explained above, the OFT has recently revised its default assumption to a reduction in price of 15 per cent following intervention. This seems more in keeping with the wider empirical evidence in the academic literature. Much of this evidence – meticulously collected and organised in a meta-study by Bolotova and Connor (2006) – suggests that the median cartel-induced price increase lies between 17 and 30 per cent.

Table 3.2   Assumptions used in cartel cases[19]

	EU 2009	EU 2010	USDOJ 2009	USDOJ 2010	OFT 2008-11
Assumed life span (yrs)	5	1/3/6	1	1	6
Gain from cartel	10%	10%	10%	10%	15%[20]
Affected consumers	Affected market size	Affected market size	Volume of commerce	Volume of commerce	Turnover affected goods
Social discount rate[21]	3.5	3.5	N/A	N/A	3.5
Estimated impact (bn)	1.2 (EUR)	7.2 (EUR)	0.6 ($)	0.05 ($)	0.25[*] (£)

* Including cartels and other commercial agreements, and abuses of dominant position.

Sources: European Commission DGCOMP (2011a), the USDoJ (2012) and OFT (2011a).

[17] Sources: European Commission DGCOMP (2011a), the USDoJ (2012) and OFT (2011a).

[18] Of course the effects will typically will be much wider than just the sales of the cartel members, hence this is likely to be an underestimate of the total impact.

[19] Among other countries, the NMa assumes a 10 per cent overcharge.

[20] 10% before 2010.

[21] This allows a discounting of future estimated cartel gains.

The assumptions on the expected future life-span of cartels show much wider dispersion, not least probably because of the significantly smaller amount of research that has been done on this matter. It seems tempting to rely on empirical studies of cartel duration such as Block, Nold et al. (1981) and Levenstein and Suslow (2006). However, these studies were conducted on cartels that were detected, which are likely to be different from undetected ones and the assumption here has to be made on the cartel remaining undetected. For this reason it may seem more reasonable to turn to what economic theory has to offer on the matter. Works on the incentives created by leniency programmes have shown that it is the less stable cartels that are more likely to apply for leniency.[22] For these cartels it would be reasonable to assume a shorter future life-span (i.e. that the cartel would not have survived much longer even absent intervention). In other cases however the CA detects cartels ex officio. Block *et al.* (1981) argue that the detection probability in ex-officio cases increases with mark-ups. The rationale behind this is that higher mark-ups are more likely to be spotted by customers or the CA, and therefore there is a higher chance of the ex-officio triggering of investigations. As higher mark-up cartels are likely to be more stable, a longer life-span assumption may be more fitting in these cases. It therefore seems more appropriate to apply a case sensitive assumption for the expected duration of the cartel absent intervention. For example Harrington (2008) shows that the quality of leniency programmes (i.e. the amount of immunity leniency programmes award to firms) has an effect on how stable cartels are. This would suggest that different assumptions should apply depending on the given on the rigour of the given competition regime. Other characteristics, such as the type of the industry, specific market conditions, and entry conditions also have an impact on cartel stability. Given this potential heterogeneity, a case-dependent approach recently adapted by the EC would seem to be most in line with economic theory.[23]

In *merger cases, as for the OFT,* simulation is fairly often used to estimate how prices, demand, and market shares might have changed had the merger gone ahead absent the CA's intervention. When the price impact of a merger was estimated during the merger procedure this can be (and often is) used in impact evaluation. In other cases default assumptions are made on the price impact, which are summarised in Table 3.3.[24] Previously the EC assumed that the future customer savings resulting

---

[22] See Harrington (2008) for a discussion on this together with relevant references.

[23] The EC classifies cartels (based on economic theory and evidence) into three categories: "unsustainable", "fairly sustainable" "very sustainable", and assumes a future cartel life of 1, 3, and 6 years respectively.

[24] Amongst some of the other jurisdictions, in the Netherlands the NMa uses the turnover of the relevant firms as a basis and assumes a one per cent price increase (Kemp and Sinderen (2008)). In Portugal, the PCA assumes mergers to lead to a 5.3 per cent price increase, which will last for two years and is discounted at 3.5 per cent (Weinberg (2007)). The Competition Commission (CC) in the UK – being a Phase II body – has more information available on the cases it looks at, therefore it does not adopt a single approach in each case, rather, it seeks to capture what the team conducting that investigation believed was the likely effect of the merger.

from corrective merger decisions corresponds to 10 per cent of the size of the relevant market(s) on which the concentration would have significantly impeded effective competition. This has now been changed to a practice whereby price effects are simulated on a case-by-case basis.

Table 3.3   Assumptions used in merger cases

	EU2009 (bn eurs)	EU2010 (bn eurs)	USFTC2010 (bn $)	USDOJ2010 (bn $)	OFT2008-11(£bn)
Affected consumers	N/A	Size of relevant market	Volume of commerce	Volume of commerce	Turnover of affected goods
Price effect	N/A	Simulated	1%	1%	Average of simulated
Consumer impact	10% of relevant market size	N/A	N/A	N/A	N/A
Duration of price impact (yrs)	1	Significant/high/very high	2	1	2
Estimated impact	5.6	4.2-6.3	0.59	0.19[*]	0.36

* Merger and other civil antitrust combined.
Sources: European Commission DGCOMP (2011a), the USDoJ (2012) and OFT (2011a).

The suitability of these methods is more difficult to assess than in the case of cartels. Firstly, in merger cases the CA has to decide ex-ante whether a merger is anticompetitive and to find a suitable intervention. This means that impact estimates would need to establish if the CA had been right to intervene in the first place. It is unlikely that any CA would admit to a wrongful intervention in their accountability reports, therefore, absent intervention, the intervened mergers in the evaluation are always assumed to have a negative impact. Secondly, price increase estimates depend largely on the severity of the merger control regime. In a lax regime, only cartels with large and positive price effects are intervened, therefore the average estimated price impact of intervened mergers will be larger than with a stricter CA.[25] Thirdly, it is not clear what value, if any, should be given to mergers that the CA correctly did not intervene (i.e. mergers with a negative price change).

In the academic literature, there are some studies that investigate the price impact of intervened US mergers.[26] Ashenfelter and Hosken (2008) for example looked at five selected cases and found estimated price increases to be between 3 and 7 per cent. In earlier studies Werden et al. (1991) reported a 5.6 per cent price increase, and Borenstein (1990) estimated a 9.5 per cent average increase. The literature is much scarcer for other jurisdictions. However, even if we had a more comprehensive idea about the average price impact of mergers, its use as a best

[25] Carlton (2009) also discusses this possibility.

[26] Weinberg (2007) surveys the relevant US literature.

practice or default assumption in this context could be questioned for various reasons. For example, as mentioned above, a systematic bias in the CAs decision making might mean that mergers with small positive price effects (when the CA is too lax) or mergers with small negative price changes (when the CA is too strict) are not picked up by the CA and hence would not appear in the evaluation. Relying on an average merger price effect assumption would therefore also require some knowledge on whether the CA is too lax or too strict. Mergers are also likely to have very different price effects depending on the economic environment and therefore assumptions based on estimates in one jurisdiction may lead to biased evaluations in the other.

For these reasons it seems most reasonable for the evaluation of mergers to rely, whenever is possible, on a case-by-case approach and use price-effect figures from the simulations conducted during the investigation. Simulations are becoming more common, especially in the assessment of those mergers that the CA judges to be potentially more harmful and an estimated price effect would therefore be available from the investigation. In more simple cases the impact of the merger is more likely to be closer to zero and therefore excluding them from evaluations (because there is no case-specific price-impact estimate) would only have a marginal effect on the estimated aggregate impact of merger control. A variation of this approach is used by the OFT, where, if simulation is not appropriate for the case, consumer savings as a proportion of turnover are assumed to be equal to the mean lower bound of the same ratio across all simulated mergers over the previous three years.

The assumption on the duration of the merger-generated price impact shows more convergence (being either one or two years). On the lower bound, as Davies (2010) points out, it seems unlikely that a CA would choose to intervene if it believed that post-merger self-correction within the market would occur within the following one or two years. On the upper bound, the case-by-case approach used by the EC seems appropriate. This method categorises all cases into one of three groups: "significant", "high" and "very high" and assigns them duration period in years as the minimum time it would take to restore competition to its pre-merger state.

Turning to *abuse of dominance* cases, as Werden (2008) points out, these pose arguably the greatest challenges for assessment (see section 1.5 above). Similarly to other case types, if case-specific information is not available, assumptions are made about the default price rise, and the duration of the infringement (see Table 3.2). The lack of empirical IO work on abuse of dominance cases makes it difficult to assess these assumptions.

In addition, some of the CAs also conduct an assessment of other activities such as consumer protection, or advocacy. As described above, the OFT 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.[27] 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.

Table 3.4   Assumptions used in other antitrust (non-cartel) cases

	EU 2009	EU 2010	USFTC 2010	USDOJ 2010	OFT 2008-11
Affected consumers	N/A	N/A	Volume of commerce	Volume of commerce	Turnover affected goods
Price effect	N/A	N/A	1%	1%	15%[28]
Consumer impact	10% of relevant market	10% of relevant market	N/A	N/A	N/A
Duration of price impact (yrs)	1	1	2	1	6
Estimated impact (bn)	2 (EUR)	0 (EUR)	0.52 ($)	0.19[*] ($)	0.25[**] (£)

* Merger and other civil antitrust combined.

** Including cartels and other commercial agreements, and abuses of dominant position.

Sources: European Commission DGCOMP (2011a), the USDOJ (2012) and OFT (2011a).

---

[27] This is problematic if the number of complaints increases due to increased consumer awareness.

[28] 5% before 2010.

4   Conceptual Issues in Accountability Evaluations

Some general issues arise from the above discussion. Firstly, some of the figures reported above show that estimates are very sensitive to the assumptions used. A case-by-case approach would therefore be desirable when possible. For the remaining cases, where default assumptions are made, it would be desirable if there was theoretically and evidence-based **common practice**. This would allow a better comparison of estimates both over time (say to inspect the effect of a policy change) and cross-jurisdiction.

Although evaluations are becoming increasingly more sophisticated, they usually emphasise only the price effects of interventions – the price increases avoided as a result of interventions, taking into account the size of the affected market. Of course, it is to be hoped that the total benefits should extend beyond just price and include the effects on quality, choice and innovation, but these are usually very difficult to measure accurately. For this reason, it is reasonable to interpret the results of these evaluations as typically **conservative estimates**. Some activities of the CA are also excluded, such as competition advocacy, or consumer education. The tendency for producing conservative estimates seems justified as a definitive measure of ex-post impact is often impossible (for example, we will never know what would have happened had a prohibited merger not been prohibited). Also, interventions may have an impact well into the very long-term and onto wider socio-economic factors, and the cost of conducting a comprehensive ex-post evaluation across all cases would be disproportionately large. These longer-term and wider effects are discussed later in the report.

Although practicability issues may limit the scope of impact evaluations and make it difficult to include longer-term or wider impacts, there are some areas which seem unjustifiably omitted from these reports. Impact evaluations only look at cases where the CA had decided to intervene and therefore exclude those **cases where there was no intervention**. For example in mergers, no impact is associated with mergers that the CA rightly authorised without intervention. This may be particularly important for cleared efficiency-enhancing mergers, where the impact on consumer benefits might be increased by some measure of the positive impact that arises from the unconditional approval of the merger. Cases where parties settle before the CA reaches a final decision are also excluded from these evaluations. These settlements are clearly the result of effective enforcement and should therefore be taken in to account, for example as a result of the deterrent effect of the CA's work.

Because of the size-difference in the relevant affected markets across cases, the aggregate savings computed may be very sensitive to one or two extreme observations. Also, potential errors in particular cases could make year-to-year fluctuations volatile. As Davies (2010) argues, **using moving-averages** helps to smooth impact estimates over the years. In the UK both the OFT and the CC applies

rolling-average figures, and so does the Netherlands Competition Authority (NMa).

When impact estimates are acquired for accountability reasons, an important source of bias could be caused by the fact that given finite resources, coupled with a need to substantiate its impact, it is rational for any CA to pursue the 'easy options' in its enforcement activities; i.e. to cherry pick easier cases at the expense of more difficult cases (for which the probability of 'success' is lower or where there is greater uncertainty). A related danger is highlighted by Neven and Zengler (2008), who suggest that the evaluation programme itself may introduce an additional **motive for distortionary discretion** in CA conduct: "Faced with simplistic assessment, authorities may be tempted to be overly interventionist, to spend too many resources and to ignore relevant information." If evaluation is driven by external accountability (to verify whether the CA delivers its objectives) and especially if undertaken by the CA itself, the CA is prone to fall into the trap that is also identified by Chang and Harrington (2010), i.e. CAs will not seek to maximise deterrence, but focus on something that is observable/measurable (e.g. the proportion of Art.102 cases that are won, or the number of cartels detected). This can have important feedback effects, not just for evaluation, but also for success in achieving the ultimate objectives of competition policy and in extreme cases even bias decision-making. These reasons may justify, in some cases, the use of an external auditor to conduct the assessment.

4.1 Unobserved impact of competition policy

As mentioned above, impact evaluations necessarily focus on only a selected part of the total effect of competition policy and law enforcement. Some of the impacts are not observed and are therefore not measurable, or at least, very difficult to estimate. The impact estimates presented above therefore do not include the deterrent effect of competition policy, or behaviour that remain undetected, or conducts that the CA detects but decides not to investigate or intervene. Unless the evaluation acknowledges how this selection process works, any impact estimate will be inherently biased.[29]

Impact evaluations also typically ignore the possibility that there may be deterred pro-competitive cases. Baker (2003) notes that he suspects the costs from deterred pro-competitive activities not to exceed the direct costs (to the firms) of enforcement. Although he does not go into further analysis, this statement seems intuitively reasonable on the assumption that only those pro-competitive conducts are deterred that are less profitable than the firm's assessment of the expected cost of litigation, implying that if a conduct is sufficiently pro-competitive, it is more

---

[29] See section 8 below and Davies and Ormosi (2010) for more detail. Carlton (2009) also warns about this in his discussion of measuring the average impact of merger control.

likely to take place. A statement like this would of course only hold true if the CA is unbiased and free of hostility towards efficiency gains.[30]

To account for the deterrent effect of enforcement, one would need to be able to measure deterrence. While this is still an area which defies definitive quantification, there exists some work of interest, particularly where a competition policy index is constructed from institutional, legal, and social factors that are expected from economic and legal theory to deter certain types of behaviour. For example Buccirossi *et al.* (2009b) created an index of a set of institutional and enforcement features that are expected to have an impact on the level of sanctions incurred by those who are convicted, the (perceived) probability of being detected and convicted, and the (perceived) probability of being wrongly convicted or acquitted. Using these types of indices allows an ordinal measurement of the levels of general deterrence across countries.

A related issue arises regarding cases that the CA did not detect. For example empirical works suggest that not more than a fifth of all cartels are detected (see section 8.1 below) Should these forgone opportunities appear in impact evaluations (for example as an opportunity cost)?

Although deterred cases are not observed, and therefore no specific estimates of the deterrence effect are included in CAs' aggregate evaluations, the OFT has acknowledged this issue in its recent impact assessments (for example, (2011a, p.24). To provide a suggestive ball-park figure for the magnitude of deterrent effects, it uses the assumption that, for every investigated case, there are five other cases which do not occur because they are deterred, and therefore the estimated impact based on intervened cases are multiplied by five. The 5:1 multiplier is, of course, highly speculative and is based on a survey conducted by Deloitte (2007) for the OFT. In a subsequent survey by London Economics (2011), also commissioned by the OFT and using a similar methodology, the evidence points, if anything, to a significantly higher multiplier. However, whatever the value of the multiplier, the line of reasoning requires the questionable assumption that deterred cases share the same characteristics as intervened ones, which – according to economic theory – is unlikely to be the case. Taking the example of cartels, economic theory implies that more stable cartels are less likely to be deterred and detected.[31] Therefore, any estimate will ignore those cartels that are likely to be the most harmful. Similar arguments could also be used for mergers. If mergers with high price-increasing

---

[30] See Werden, Joskow et al. (1991) on the evolution of the treatment of efficiencies in antitrust.

[31] Chang and Harrington (2010) argue that less stable cartels are more likely to apply for leniency. If cartel investigations are triggered dominantly by leniency applications (in the EC 2/3 of the cartel investigations result from leniency) then the investigated cartels will be the less stable ones.

effects are more likely to be deterred,[32] then impact estimates based on investigated cases only will be negatively biased.

Another seemingly simple way to get a grasp of the level of deterrence is to look at how the number of cases changes over time. This leads to the problem that the number of detected cases is an ambiguous indicator of deterrence, as an increased number of observed cases could mean either an increase in the detection rate or a decrease in the deterrence rate of enforcement, or both. A change in the number of observed cases can only be meaningfully interpreted if, at the same time, the change in detection rate is also known. This of course may not pose an insurmountable problem in areas where the rate of detection is expected to be constant, which is likely the case in merger control regimes with compulsory pre-merger notification rules. Barros, Clougherty et al. (2010) build on this feature when they estimate the deterrent effect of merger policies. Otherwise, Ormosi (2011) proposes a way to estimate how detection rate changes over time in anti-cartel enforcement, which in turn could be used to make inferences on the rate of change in deterrence. It may indeed be this latter category (the marginal impact of competition policy on deterrence) where future work should focus, i.e. whether additional Government spending increases deterrence.

We return to sample selection, detection and deterrence in the final chapter.

---

[32] For example Barros, Clougherty et al. (2010) suggest that firms may reduce the restrictiveness of their merger in order to increase the chances of approval.

5   Ex-post Evaluation by the Competition Authorities

In the main, the evaluations described in the three previous chapters are ex-ante. However, the CAs also undertake ex-post evaluation, usually on a more ad hoc basis.

Ex-post evaluations can be loosely broken down into two main types. Qualitative studies typically survey the market post intervention in an attempt to find out whether the predictions at the time of the intervention have proven to be true or not. These studies systematically examine how market entry and expansion conditions changed in the years following intervention and assess whether the predictions at the time of the decision turned out to be accurate. The other type relies more on quantitative works by trying to establish causality between the intervention and, for example, a change in the market price. In practice it is common that if the evaluation is commissioned by the CA then a mixture of these two types of method is used. The following are some illustrative examples.

The most notable examples, published and therefore in the public domain, is a series of reviews for the UK, where the CC, the OFT and the relevant Government Department have regularly conducted (or commissioned) retrospective merger reviews, reflecting upon market developments following the merger interventions (PricewaterhouseCoopers (2005); UK Competition Commission (2008); Deloitte (2009), LEAR (2011)).

Various methodologies have been used for this purpose. For example the 2009 evaluation (see Chapter 7) used interviews with market participants, and simulations, whereas a 2011 assessment chose to rely on DiD and a survey of market actors. Follow-up questionnaires and/or interviews with the interested parties and related firms often provide invaluable insights but are inevitably prone to a number of potential limitations: low response rates, respondent bias, the parties often have short corporate memories, and in their view, interventions can often be overtaken by other subsequent and more important events. In a much earlier study, Clarke *et al.* (1998) includes a number of examples of the latter. The OFT also reviews (or commissions) two of its cases every year.[33] These assessments involve a post-event monitoring of market developments.[34] The OFT and the CC also regularly conduct an assessment of the impact of market studies.

---

[33] For example, in 2011 it evaluated the impact of the 2001 abuse of dominance case against Napp Pharmaceuticals (2011d), and the 2005 consumer enforcement case against Foxtons (2011e).

[34] An OFT example (2007) is the impact assessment of the Payment Systems Task Force review.

A different type of exercise is when the survey focuses on the perception of stakeholders regarding the CA's practices in general. The purpose of these surveys is typically to provide an objective measure of the experiences and the satisfaction of stakeholders with the CA's work. The surveyed stakeholders may include parties to the given case, third party businesses, interested businesses, professional advisors, or government bodies. Examples for this are given by the UK's Competition Commission (2009), and the EC (2010).

The EC commissions some of their evaluation work to external experts. Buccirossi *et al.* (2006) for example use a combination of an event study and a survey to examine whether the EC made the right decision in the Pirelli/BICC merger (see Chapter 7). A recent example from the new Competition Commission of Mauritius (CCM) (2011) examined the evolution of market price, the change in the market structure and a calculation of possible savings to consumers following its intervention in an abuse of dominance case (see Chapter 7).

Some studies have assessed the quality of CA decision-making by the frequency of court appeals and/or the success rates in those appeals.[35] For certain purposes this may be a valuable extra source of (presumably well-informed and objective) evaluation, but obvious limitations include the likelihood of selection bias, and the fact that court decisions will sometimes involve judgement on the correctness of legal process rather than economic substance. Although courts analyse cases ex-post, they typically rely only on information that had already been available at the time of the CAs intervention. Bergman (2008) also points out that relying too much on appeal success ratios may have an incentive effect (similar to what has been discussed above), because CAs may become more interested in pursuing simple cases or enforcing only against blatantly illegal conduct.

Certain high-profile annual reviews provide an alternative approach to ex-post assessment, based on peer review evaluation of the performance of different CAs around the world, e.g. the Global Competition Review and OECD country reviews. These enable international comparisons over time at the aggregate level, but are often based on subjective opinion rather than quantitative methodologies.

Some of the CAs have also conducted or commissioned case studies of the successfulness of their enforcement toolkit. For example the success of merger remedies has been analysed by the USFTC (1999), the EC (2005) and the UK CC (2010).[36] These evaluations look at whether past merger remedies were suitable to remedy a competition problem.

Similarly to what has already been argued in the previous chapter, there is an innate risk in the internal assessment of specific cases, which stems from the risk of

---

[35] Bergman (2008) provides a brief survey of these and discusses their limitations in more detail.

[36] See also an EC commissioned study by Davies and Lyons (2007) on merger remedies.

self-reporting bias, especially in cases still pending appeal, as the CA has a strong incentive to report high consumer savings otherwise to avoid providing evidence against their own case.[37] Given the potential risk of this bias, third-party evaluations may be preferred in some circumstances.[38]

---

[37] Bergman (2008) also emphasises this possibility.

[38] A similar recommendation is given by Buccirossi, Ciari et al. (2006).

6 Evaluation Techniques: previous literature, pros and cons

The next two chapters will focus more specifically on the alternative techniques which might be used to evaluate the impact of a particular intervention for a particular case. This chapter first defines the three main classes of methodologies: simulation, event studies and differences-in-differences. These were first introduced briefly in section 1.3, and they were referred to, in passing, in the previous chapters. This chapter now provides more extensive definitions, surveys of the literatures in which the methodologies have been used, and discusses their 'pros' and 'cons', i.e. advantages and disadvantages. The next chapter then provides an overview of six specific case studies in which they have been employed, in order to illustrate the types of data they require, and theoretical and empirical models on which they might be based.

6.1   Simulation and structural models

By simulation, we refer to evaluation based on the following three 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).[39]

(iii)   Using the model to assess how the equilibrium would change with and without an event/intervention (for example comparing the pre-merger equilibrium with the hypothetical non-intervened equilibrium[40]).

---

[39] By substitution of these parameters into the equilibrium conditions derived in stage 1, one can then recover (i.e. solve for) other, unknown, parameters, such as firms' marginal costs.

[40] In principle, one might also simulate a switch in the prevailing behaviour of firms as a consequence of the merger; for example, if it is suspected that the merger would result in a coordinated effect between merging and non-merging firms, one could compute alternative collusive post-merger equilibria. Rather more difficult are the possibilities that firms might want to reposition their products, or where the potential for new entry becomes important.

In principle, simulation may be either ex-ante or ex-post, but most applications are ex-ante in the sense that they are based on data available at the time of the intervention – even if they are actually conducted retrospectively.

6.1.1 Previous literature

The early development of simulation was firmly rooted in the academic literature and some of the seminal contributions include Farrell and Shapiro (1990), Werden and Froeb (1994), Hausman and Leonard (1997), Werden (2000) and Nevo (2000). These references all relate to merger simulation which is overwhelmingly the most common area of application. Fairly comprehensive surveys are provided by Davies and Lyons (2007), Budzinski and Ruhmer (2010) and Buccirossi *et al.* (2006). Simulation is far less frequently applied in evaluating the effects of abuse of dominance or in cartel interventions, but see Verboven and Van Dijk (2009).

Increasingly over the last 20 years, simulation has featured in specific merger investigation by the CAs and advisers to the parties. Again, these are well documented in the literature – fairly comprehensively, for example by Budzinski and Ruhmer (2010), including: Interstate Bakeries/Continental Baking (US 1995), Kimberley-Clark/Scott (US 1996), Volvo/Scania (EU 1999), Lagardere/Natexis/VUP (EU 2003), Nuon/Reliant (NMa Netherlands 2003), Oracle/PeopleSoft (EU and US 2004). These and other cases led to subsequent seminal simulation papers in the academic literature including Pinske and Slade (2004), Ivaldi and Verboven (2005), and Peters (2006).

6.1.2 Pros and cons

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. These qualifications are well documented elsewhere (see for example Buccirossi *et al.* (2006)).

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. These may be appropriate for industries which are either homogeneous products, with or without capacity constraints, or with horizontally differentiated products. In such industries an emphasis on price and quantity may be more appropriate. However, where products are vertically differentiated and competition is effected more through product quality and innovation, these models are typically incapable of describing the competitive process. Other aspects which are difficult, if not always impossible, to model completely satisfactorily are: (i) possible changes in conduct post-merger (coordinated effects), (ii) buyer power, and (iii) bidding markets (which is still in its 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. Ashenfelter and Hosken (2008, p.36) summarise on this count by suggesting that "careful evaluation of their effectiveness seems long overdue".

6.1.3 Back of the envelope (simple) simulation

As mentioned, full-fledged simulation is extremely demanding of data and research time. Faced with this, CAs in their own evaluation exercises often employ simplified versions of simulation models (the only viable option if the CA is to evaluate a range of mergers investigated in a given year). Typically, most of their effort is directed to those mergers where the product is homogeneous, and can be reasonably described by the Cournot model, or simple models of differentiation. In order to calibrate these, extraneous estimates of demand elasticities are required, but there is often a scarcity of good estimates. In these circumstances, Werden (2008) reports that the US typically employs a range of 1 to 1.5 for the aggregate industry demand elasticity, but is not obvious whether this is a reasonable ball-park range. Such inelastic demand will inevitably lead to high predicted price increases. As noted by Bergman (2008, p.394) "There exist amazingly few econometric studies of the price effects of mergers, considering the economic importance of mergers and given that merger effects is a topic that is well suited for this type of quantitative analysis". Below, we discuss Connor's painstaking meta-analyses of cartels, but there appears to be no equivalent for mergers. At the very least, there is an obvious gap to be filled by a similar sort of meta-analysis of the industry demand price elasticity – given its key role in most simulation.

6.2   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.

6.2.1 Previous literature

Event studies of mergers typically examine the effect of the announcements of the merger and the CA's decision (i.e. the type of intervention) on the valuation of the merging firms and their rivals.[41] For Europe, Duso *et al.* (2007) apply the methodology to a sample of 151 mergers, 1990-2002, and show that, for at least half of them, rival firms benefit after the merger is announced. According to theory, mergers will benefit rivals only if they are price raising, i.e. anticompetitive. *Duso et al.* (2011) subsequently focus on the effects of the Commission's decisions (merger remedies) and find that, on average, remedies seem to be only partially capable of reversing announcement abnormal returns – especially for Phase 2 decisions. For Australia, Diepold, Feinberg et al. (2006) apply the methodology to 50 mergers handled by the Australian Competition and Consumer Commission between 1996-2003, and find, for example, that the CA did not intervene in any of the mergers which the event study suggested to be pro-competitive, i.e. consistency between the CA and the financial markets. The methodology has also often been applied to specific case studies, for example by Simpson and Hosken (1998) for four FTC investigations of retail mergers between 1984 and 1993, and Warren-Boulton and Dalkir (2001) for the Staples – Office Depot merger.

Event studies on cartels typically examine the impact of dawn raids and the subsequent CA decisions, for example Langus and Motta (2010) for EC cartel enforcement, and Bosch and Eckard (1991) for US DoJ decisions. Lübbers (2009) studied the effect of cartelisation (a coal syndicate) in Germany, 1893-1913, where the events analysed were the foundation of the syndicate and two major modifications to the original contract. Applications of the methodology to cases of monopoly abuse are somewhat less frequent, but examples for specific cases include Bittlingmayer and Hazlett (2000) for Microsoft, and Sidak (2003) for WorldCom. In a recent novel application of the approach, Broek *et al.* (2012) study a sample of 66 Dutch antitrust investigations (including cartels). Their approach

---

[41] Eckbo (1985) is an early discussion of how the effects on rivals' valuation provides important additional information.

entails decomposing the estimated value loss to the firms into three parts: (i) the penalty itself, (ii) an estimate of future lost rents due to the assumed future removal of the overcharge, and (iii) the 'reputation loss'. The latter reflects the loss in future rents because, in the light of the prosecution, the firm's investors, customers and suppliers revise 'the terms of trade with which they do business with the firm' (Broek *et al.* (2012, p.11)). This is estimated to account for one third of the overall lost value. While this particular quantification must be very approximate, (being derived as a residual, and therefore sensitive to inevitable errors in quantifying the overcharge in component (ii)), methodologically this is an interesting idea on how to interpret the estimated magnitudes of lost value from event studies.

More generally, Carletti *et al.* (2011) analyse the impact of the introduction, or modification in competition laws in 18 jurisdictions.[42] Other related subjects where event studies have been used include private antitrust litigation (Bizjak and Coles, 1995) and the effects of entry (Whinston and Collins, 1992). Cichello and Lamdin (2006) provide a fairly comprehensive survey across all areas of policy, and present a more complete list of references.

6.2.2 Pros and cons

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 can then 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 a CA as it tackles the issue of information asymmetry between itself 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

---

[42] Other examples of measuring the effect of regulatory change, although not for competition laws, include: Becher (2009), who looks at the effect of US interstate deregulation in the banking sector; and Prager (1989), who examines the effect of the Interstate Commerce Act in the railroad industry. A major problem with using event studies to measure the impact of regulatory changes is the definition of the event window. Regulatory changes are lengthy procedures, typically starting with drafting at administrative level, before the proposal is carried on to parliamentary committees, brought to the chamber floor, and finally approved by the parliament. During this process, information about the proposal is often made public. As the content of the legislative draft may continuously evolve, it is hard to define an appropriate event window.

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 some 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.[43]

Mergers are the main area of competition policy where event studies have been used. Table 6.1 summarises how each of the three constituent events are expected to change the merging parties' market value. At each stage, there is the potential for ambiguity. For example an increase in valuation of the merging parties may reflect either pro-competitive effects (efficiency gains), or anticompetitive effects (exclusion, dominance, or collusion). Duso *et al.* (2007) argue that this ambiguity can be resolved by observing the change in the valuation of rival firms. According to most static oligopoly models, horizontal mergers will result in a higher product price unless there are offsetting efficiency gains; while the former will also benefit rivals, the latter will not. However, as the final row shows, the ambiguity does not vanish completely. For example, although an observed negative change in the rivals' valuation may reveal pro-competitive (efficiency) expectations, it may also indicate anticompetitive (exclusion) effects.

Table 6.1   Expected effects of a (horizontal) merger on merging firms' and rivals' asset values

Source of post-merger gains	Merger proposal		Announcement of investigation		Merger clearance	
Effect on:	Merging firms	Rivals	Merging firms	Rivals	Merging firms	Rivals
Dominance or collusion	+	+	–	–	+	+[44]
Efficiency	+	–	–	+	+	–[45]
Exclusion	+	–	–	+	+	–

Source: Davies and Ormosi (2010)

There is also the possibility that the merger may be interpreted as a signal that other firms in the same market will engage in merger activities in the near future,[46]

---

[43] See also Malkiel (2003) for a more general criticism of EMH.

[44] The situation would be different for vertical mergers, which can potentially have foreclosure effects, and would therefore result in negative abnormal returns.

[45] Assuming that competitors cannot free-ride on merger generated efficiencies.

[46] Cox and Portes (1998) claim that was the case in the merger of SBC Communications and Pacific Telesis Group.

which would increase the market valuation of rivals, resulting in the same sign of change as with collusive and/or dominant outcomes. Although Da Graca (2006) proposes a method for eliminating simultaneity biases, this still remains an under-explored issue. A further criticism is that the methodology does not separate out the market's anticipation of the CA's eventual decision, therefore its reaction to an 'antitrust event' may equally be explained by the market updating its beliefs about a particular antitrust decision, once the uncertainty about the merger investigation is resolved. Duso *et al.* (2010) address this problem by using observable merger characteristics to estimate the probability of a particular decision and correct the average abnormal returns accordingly.

The pros and cons are similar for assessing cartel enforcement. However, in this case (characterised by a higher level of CA secrecy, in order to ensure the success of the investigation), there is more chance that the first event (typically in the form of a dawn-raid) will actually be 'unanticipated news' to the market. Also, the theoretical expectations are less ambiguous. The effect of news of the investigation and the ultimate decision nearly unambiguously reduce the valuation of the parties. The one exception is when insufficient sanctions are announced. In this case the negative effect stemming from the elimination of cartel profit may be mitigated by a smaller than expected fine.

An obvious question in general is how well event studies predict actual outcomes. For mergers, Duso *et al.* (2010) compare the results of an event study with an ex-post analysis of balance sheet profits. They find that in some cases the abnormal returns measured in event studies are positively and significantly correlated with the ex-post measured profitability of the same mergers. This provides some affirmative evidence, but further, and possibly deeper, statistical analysis is required before any definite conclusions can be drawn on this.[47]

Finally, in spite of the general presumption that event studies are easy-to-use and data are easily accessed from financial databases, there will be 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, especially in markets where firms are small and rivals are scarce. Moreover, very often the parties are large conglomerates and/or multinationals, and the market concerned may constitute of only a small part of its aggregate activities. Where this is the case, it can often prove difficult to identify any effect on

---

[47] For example when calculating the ex-post profit effect of the merger, the counterfactual used is the median firm in the same market minus the effective rivals. The authors assume that these firms are not strongly affected by the merger, but this assumption needs more evidence especially for large mergers, which may affect the profit figures of other – not directly competing – firms in the market. The main worry however is that their results in general show that the stock market reaction is positively correlated with the ex-post development of profit typically in cases where a longer event-window is used. This raises doubts whether the stock market reaction is actually a reaction to the merger or it picks up other confounding effects as well.

the firm's valuation resulting from an event in a small market in a particular country. Some of these practical difficulties are cited by inter alia Buccirossi *et al.* (2006, p.187-8). A specific example is provided by Beverley (2008) who attempts to apply the methodology to a sample of nine UK Competition Commission merger inquiries but is ultimately frustrated by an inability to locate sufficient competitors with traded equities and merging parties for whom the market concerned accounts for a sufficiently large proportion of their activities. For these reasons, it seems likely that event studies applied across samples of mergers may suffer from an inherent sample selection bias.

6.3   Difference-in-Differences

Difference-in-differences (DiD) methods belong to a broad category of methodologies sometimes known (rather unhelpfully) as evaluation methods (see Buccirossi *et al.*, 2006), which also includes natural experiments and matching methods. The basic idea is to evaluate 'performance' before and after an event (or sometimes before, during and after) in the market concerned relative to performance in another similar (control) market, unaffected by the event. The standard DiD application is typically econometric, in which the performance measure (usually product price) is tracked over time to include the pre- and post-periods in the treatment (e.g. merger) and compared against the same in the control market, with the use of dummy variables.

DiD analysis is typically conducted ex-post. The time lag before it is undertaken reflects a trade-off between choosing a sufficiently long post-event period to gain a better grasp of long-term effects, and avoiding the choice of a time period which is so long that it would compromise the chances of finding a practicable control to emulate the counterfactual. There are also some instances of ex-ante DiD analyses, for example for predicting the effect of cartel intervention (which presumes overcharge is removed).

6.3.1 Previous literature

Examples of DiD for cartels include Symeonidis (2002), who assesses the impact of the introduction of the 1956 Restrictive Trade Practices Act in the UK. Levenstein and Suslow (2006) examined the impact of the 1982 United States' Export Trading Company Act by comparing manufacturing industries with and without export cartel exemption.

Connor *et al.* (1998) attempted to measure the impact of 112 American hospital mergers, using a database of around 3500 hospitals. Their method essentially follows a DiD approach, where the treatment markets were those with mergers, and the control was taken from the remaining set of markets.

There are few examples where DiD has been applied to abuse of dominance cases, deterrence, or other effects of public intervention. An early equivalent is Shaw and Simpson (1986), who analyse the erosion of market dominance in markets where the UK Monopolies and Mergers Commission was active (markets not investigated by the MMC were the control group). There are also some examples where DiD has been used to assess the impact of other types of regulatory intervention. For instance, Cooper, Gibbons et al. (2010) use a DiD estimator to test whether the introduction of hospital competition in the English NHS in January 2006 has led to increased efficiency.

6.3.2 Pros and cons

The difference-in-differences approach enjoys an (at least superficial) appeal that it uses observed data from the relevant product market (i.e. what actually happens) in comparison with a control, where the event does not occur. Thus the counterfactual is not dependent on untestable, maybe restrictive, theoretical assumptions.

However, the other side to the coin is that the methodology is inevitably atheoretical. Since much depends in evaluation on the nature of the counterfactual this means that a key part of the methodology – identifying an appropriate control group – is also atheoretical. As such, there is a danger that the choice of counterfactual (control group) is constrained by 'what is out there', i.e. the best of a set of alternatives, none of which is entirely appropriate. When choosing the control market the ideal is that it is characterised by the same supply and demand shocks as the treatment market, which makes it possible to filter out the net effect of the analysed event by controlling for these supply and demand shocks. But Simpson (2008) warns about the danger of this assumption, claiming that even the same supply and demand shocks may influence prices differently in the two markets.

On a practical level, the control should have a sufficient number of members and time observations to emulate the random variation which would occur in the treatment group post intervention that is unrelated to the intervention itself. Most research however just assumes that the only difference between the treatment and the control groups is the treatment.[48] Meyer (1995) identifies some of the limitations of this approach, such as omitted variables, trends in outcomes, measurement error, simultaneity, selection bias, omitted interactions, etc.

Although DiD is most typically applied as an ex-post tool for evaluating the effect of mergers, this raises a – typically unanswered – question: if the chosen control group is other non-merging rivals, then how does one allow for the possible externalities that the merger has on those rivals? Tenn and Yun (2010) for example

---

[48] See Buccirossi, Ciari et al. (2006) for a discussion of the selection bias which will occur if the set of unobservable characteristics which affect the decision to merge also affect the performance of the parties post-merger.

looked at the effect of US divestitures in the Johnson & Johnson/Pfizer acquisition. Although the authors concede that their choice of control group (similar brands in the same category) may mean that the treatment (divestitures) had an effect on the control group, they claim that this does not raise serious concerns when simply looking at whether divestitures had an effect at all on prices. We are not convinced that this is indeed the case. Ashenfelter and Hosken (2008) examined the price impact of five US mergers using a set of different control groups. In the event, their preferred choice was private label products sold in the same industry. Their rationale for this was that private labels are only weak substitutes, from the consumer's viewpoint, for the higher quality branded products affected by the merger. At least to the current author, familiar with the UK supermarket industry, this would seem to be highly contestable. There are similar doubts with Dobson and Piga (2011), who use as controls airline routes from different but close airport-pairs.

Overall, difficulties in identifying a satisfactory control must raise the worry that DiD may only be useable for a fairly small sample of markets, and a sample which may not be representative of the population as a whole.

6.3.3 Before and after

One of the most commonly used methods of evaluation is to conduct a simple before and after comparison of, say, price. This is strictly not difference-in-differences since there is no control group. In effect the implication, which is often only implicit, is that there is no need for a control, or that the default counterfactual is one of merely the status quo: absent the merger or intervention, the market would simply have performed identically before and after. Clearly, the results from such a methodology should be interpreted with caution – at the least, one would look for qualitative sources (e.g. interviews/questionnaires) to substantiate the choice of a no-change counterfactual.

A slightly more sophisticated variation on this theme is to econometrically estimate an equation including a simple dummy variable reflecting the effect of the intervention, and adding other 'control variables' in the hope that they capture, and control for, other exogenous changes in the market. However, this still requires the assumption that the counterfactual is no change since the additional exogenous explanatory can, at best, control for confounding factors which would otherwise obscure what is the equivalent to no change after the event.

6.4   Choosing between methodologies

In effect, much of our discussion so far can be interpreted as questioning the nature of the counterfactual assumptions employed by different methodologies – both in

specific cases, but also more generally – are some methodologies intrinsically more counterfactual-aware than others?

Simulation, by its nature, places the choice of counterfactual conspicuously at centre stage: a specific oligopoly model is selected, and this immediately reveals the nature of the assumed counterfactual equilibrium. Similarly, it must be calibrated transparently with key parameter estimates. A related issue was raised by Davis and Garces (2009) who pointed out that, in merger simulation, typically the pre-merger equilibrium is used as counterfactual, which ignores the possibility that the market was not in equilibrium pre-merger and the merger happened to get to equilibrium. Sidak and Teece (2009) also warn about the sensitivity of analysis to the choice of the counterfactual.[49]

In DiD, the control plays the role of the counterfactual – for example, what would have happened in the UK, had retail price margins for books not been repealed - might be captured by what actually happened over the same period in Germany, where it was not repealed (Davies and Olczak (2008)). Here, the choice of counterfactual is less theoretically driven, and the strength of the methodology rests on whether there really is a control (with adequate data) which is sufficiently similar. In practice, data expediency may sometimes distract attention from how closely this condition is met.

While it is less common to think of the counterfactual in the typical event study, implicitly it is still there – captured by whatever comparator share price index the practitioner uses to compute abnormal returns. Here there is a trade-off between using a general index, which is less likely to be sensitive to market-specific exogenous events, and more customised sector-specific indexes, which may not be truly independent of the event at issue.

An ideal approach for comparing the alternative evaluation methodologies would be to assemble an ideally very large, random sample of cases, and attempt to apply all the methodologies to all cases in the sample. This would help both in assessing their relative practicabilities and in identifying any systematic differentials in their estimates. While it is likely that some CAs and advising consultancies do sometimes conduct parallel assessments during their conduct of a particular case (experimenting simultaneously with, say, an event study and a simulation), these do not appear in the published literature of course. More generally, attempting such a task across a large sample of cases would be difficult, and, in the event, has not occurred to date.

Another possible way of finding adequate counterfactuals could use some sort of propensity score matching. For example for cartels, one could estimate the

---

[49] In one example they cite, they argue that, when using a SSNIP test in periods of economic downturn, the counterfactual for a proposed merger may be that prices that would otherwise fallen would might be stabilised.

probability of being in a cartel (using a latent dependent variable and market and firm specific independent variables) and match observations (from both treatment and non-treatment) that produced similar probabilities. The non-treatment cases that produced similar estimated probabilities could then be used as counterfactuals.

6.4.1 Comparing counterfactuals: cartels as a case study

In this sub-section, we attempt to illustrate some of the issues in choosing the 'right' counterfactual with the following example. We use an existing database on cartel overcharge already in the public domain, constructed by Connor and various co-authors. Connor's meta-analysis draws together the results from a large number of primary studies, including 800 observations, taken from nearly 400 cartel episodes across the world.

This is only a second best to an ideal dataset, because the overcharge estimate for each cartel has been computed typically using only a single methodology – for very few of the cartels do we have alternative estimates using different methodologies. This means that relatively higher power matched sample tests are impossible, but the sheer size of the sample should nevertheless reveal some reasonably meaningful insights into the estimates generated by different methodologies.

The version of the database we use here is summarised in Bolotova and Connor (2006). The main purpose of their paper was to examine the determinants of overcharge, and the methodologies used to derive the estimates were merely a minor side issue to be controlled for. For us here, this is the main focus of attention.

They classify the estimation methods used in the primary studies into eight broad groups. Of these, we discard about 30% of cases which belong to three of the groups since they do not correspond to any of our methodologies above.[50] Three of the retained categories (Table 6.2) correspond roughly to the component parts of DiD: Price before (Price after) are comparisons between the within-cartel-period price and the price immediately before (after) the cartel period, and Yardstick competition are relative to prices in yardstick 'analogous markets that were believed to be free from cartelization'. The Price war method compares prices within the lifetime of a cartel with prices under temporary cartel breakdown. Finally, the Econometric category, is very broad based, but appears to include estimates based on simulation-type methods.

Table 6.2 shows the proportions of cases in each category and our calculation of their sample mean overcharge estimates.[51] They fall within a broadly similar range

---

[50] These groups were: 'sundry/unknown', 'historical records', and 'cost calculations'.

[51] C&B do not report these means directly in their paper, but we have recovered them from their reported preliminary regression of overcharge solely against dummy variables for each category – in that case, regression coefficients coincide with the sample mean values.

(23–45 per cent), but with an intriguing ranking – price wars being the highest and price-after the lowest.

Table 6.2   Choice of counterfactual for estimating overcharge

	%	Mean overcharge
Price before	33	29
Price after	11	23
Price war	2	45
Yardstick competition	11	39
Econometric	15	31

Source: Author's calculations using data reported in Bolotova and Connor (2006)

A simple null hypothesis is that each sub-sample (i.e. for each methodology) is randomly drawn from the same population distribution of overcharges (i.e. no systematic tendency for specific methodologies to over- or under-estimate). Abstracting from systematic measurement errors,[52] if the null is rejected, this implies some systematic difference between methodologies in how they measure the counterfactual. In common with others working in this area, Connor and Bolotova refer to the counterfactual as the 'but for' price. However, this is not without ambiguity: is the 'but for' price the 'competitive' price or the market price that would have obtained under a set of identical conditions (including market structure) except for the existence of the cartel? Contrary to some discussion, the correct answer would appear to be the latter and this leaves open the possibility that a defendant might argue that damages should be only moderate because the 'but for' should be the tacitly collusive price.

To pursue this further, consider the relative magnitudes of the sample means in the above Table 6.2.[53] Suppose, merely for the sake of the argument, that the yardstick method identifies the 'competitive' outcome as the counterfactual. If so, we know that on average cartels set a price 39% higher than the competitive level. This then allows us to interpret the counterfactuals identified in the three other categories as follows:

- during price wars the price falls to 45 per cent below the cartel price, i.e. 6 per cent below the competitive level.

- the price before cartel is typically 10 per cent *higher* than the competitive level

- the price after cartel is typically 16 per cent *higher* than the competitive level

[52] Or that the researchers' choice of estimator is not independent of the actual cartel price.

[53] For this purpose the 'econometric' category is discarded because it includes a heterogeneous variety of implicit oligopoly models (counterfactuals) across papers.

Taken at face value, this implies the following typical time path for price. Before the cartel, price starts from a supra-competitive level then rises during the cartel period, before falling post-cartel. However, post-cartel, price remains at a supra-competitive level.[54] On the other hand, in the typical price war price falls substantially below the competitive level (implying harsher punishment than Nash reversion).

Of course, this is very highly speculative – it is based on simplistic interpretation of sample point estimates without attention to statistical significance. This is deliberate because of doubts about the quality of at least some of the estimates in Connor's database. Ehmer and Rosati (2009) have re-worked this database to exclude all estimates not meeting a variety of selection criteria, and this leads to a drastic pruning in the sample. Future work is anticipated which will investigate whether the estimates reported in OXERA (2009, p.90-92) are robust to such a pruning, and whether or not they differ significantly.

---

[54] Regarding post-cartel prices, Kovacic, Marshall et al. (2007) report that in the post-plea period products with two conspirators continue to be priced as if the explicit conspiracy never stopped, while products with three or four conspirators return to pre-conspiracy pricing, or lower, quite quickly. Sabbatini (2008, p.501) concludes his study of Italian milk cartels by noting that cartels may continue as 'well-established' rules post-detection. This possibility, that tacit collusion might often survive cartel busts is also implied by Fonseca and Normann (2009). Harrington (2004) raises a slightly different argument: the parties might moderate price reductions mindful of the signal that this would send to the courts regarding the magnitude of the previous overcharge. Regarding pre-cartel price, Harrington (2006) points to various instances where cartel formation is preceded by a significant price decline, and suggests that this might be accounted for by, variously, cost/demand shocks, entry and/or capacity expansion, and the breakdown of tacit collusion. However, as Harrington notes, this question has so far attracted little or no theoretical examination.

7 Example Case Studies of Evaluation Methodologies

The purpose of this section is to present six selected case studies to illustrate the different methodologies described above. For each one, we describe the case itself, the evaluation methodology and the data and results. We then point to the key issues, problems and reservations, and discuss the choice of counterfactual. The cases have been selected so as to include two examples each of differences-in-differences, event studies and simulation, we have also ensured a mix of mergers, cartels and monopoly abuse, and the authors include academics, consultants and competition authorities.

7.1 Case 1: Two UK Mergers (difference-in-differences)

This first example case study was a retrospective one – actually involving two quite separate merger cases. It was conducted by the LEAR consultancy (Buccirossi et al, 2011) and commissioned by the UK's Competition Commission (CC). The purpose was to evaluate the quality of the CC's decision-making, assumptions and analysis, with these two particular mergers chosen as illustrative examples, rather than because of any specific significance of the mergers themselves. Their focus was on the effects of a merger on price, using the difference-in-differences methodology applied within a regression framework, which also included other explanatory variables to further control for confounding factors.

GAME/Gamestation

The GAME/Gamestation merger was a merger between two specialist retailers of video games in the UK, and analysis was conducted for both 'mint' (brand new) and 'pre-owned' games because there had been controversy in the original investigation as to whether these belonged to the same market, or formed separate markets. In a first stage, LEAR conducted a simple before-after analysis of national average prices, and in a second stage a difference-in-differences analysis using the merging parties as the treatment group and their competitors as the control group. The data were for the prices of about 200 different games sold in 9 different retailers both before and after the merger. The pre-merger data were available from the CC's original investigation and the post-merger prices were collected in a survey conducted by LEAR itself.

Their main findings were that (i) there was a reduction in the general level of prices in the years following the merger, and (ii) this downward trend was more marked for the merging parties than for the market as a whole. On this basis, LEAR concluded that the CC had made the correct decision in not intervening in the merger, and that this was probably circumstantial evidence that not only did the

merger not enhance market power, but also it appears to have had an efficiency enhancing effect.

Waterstone's/Ottakars

The Waterstone's/Ottkars merger was a merger between two specialist 'High Street' book retailers in the UK. Again, the focus was on price although ideally measures of quality, such as range of titles and quality of service would also have been examined. As a first step, designed to establish the geographic dimension of the market, LEAR conducted an analysis of the variability in price across the parties' stores, but this was inconclusive. As a second step, it assembled two different datasets – one at the individual store level for Waterstone's (200 titles across 60 different stores) and the other at the national level, designed to establish the effect of concentration both locally and nationally. Here, there was both a before- and after-merger dimension. At the local level, the control was achieved by comparing stores where there was an overlap between the parties pre-merger, against those where there was not. At the national level, the control was the parties' main rivals. In both cases, the methodology suggested that the merger had no price raising consequences. However, this was possibly due to an exogenous toughening in competition at the time due to the growth of the internet sellers (notably Amazon) and sales of books by the giant UK grocery supermarkets.

These two cases help illustrate some more generic features of the difference-in-differences methodology. First, the demands on data (and therefore time and financially) are often intense – many of the most successful previous applications have involved retail markets with a local as well as national dimension which adds an additional potential source of variability. It also helps considerably if the pre-data were already collected and available from the original investigation – as they were here. However, even in these two data-rich examples, it proved impossible for a very able team who were not overly time/cost constrained to undertake any tests of the key service-quality dimension.

The choice and availability of alternative counterfactuals is also a key issue. In this case, as in many other local retail cases, the existence of some local markets where only one of the parties was present pre-merger is extremely valuable – an attempt to establish what would happen if two competitors become only one is facilitated if we can already observe some markets where this is only one. On the other hand, the use of competitors' prices as a control is potentially problematic. In most oligopoly models, the equilibrium response by the outsiders to a merger is rarely zero, and in general it will be optimal for rivals to raise their prices post-merger. If so, interpreting rivals' prices as indicative of what would have happened had the merger not occurred is questionable.

7.2   Case 2: Cheese in Mauritius (back-of-envelope difference-in-differences)

This case has been chosen as a very recent example of an in-house ex-post impact evaluation by a young competition authority - the Competition Commission of Mauritius (CCM) (2011) and its intervention in 2010 against an abuse of monopoly. The abuse was the practice of offering retroactive rebates on Kraft branded block processed cheddar cheese in exchange for supermarket premium shelf space, not only for Kraft's cheese but also for other Kraft-branded products, including chocolates, biscuits and powdered juice. The evaluation was conducted unusually soon, only one year, after the CCM's intervention. Typically, studies only tend to be conducted a number of years after in order that more post-intervention data are available. However, as pointed out by CCM, this was two years after the investigation-launch and they judged that the impact was beginning to be felt even whilst the investigation was underway. The assessment examined price and the change in market structure. The data used were collected from various governmental institutions as well as the market players themselves.

It found that, post-intervention, two new brands entered, driving the HHI index down from 8,200 to less than 5,000. In itself, this was considered to be a positive impact since CCM viewed the offending practice to be exclusionary and it saw as its objective the reduction of barriers to entry. It also found that the average price of the two incumbents (Kraft and Chesdale) increased by 6 per cent in the post-intervention year compared to an annual average increase for the three years pre-intervention of 16.5 per cent. In addition, following the entry of the two new brands, the average price of the two incumbents actually fell by 4.5 per cent; and there was a significant drop in the price of block processed cheddar cheese in almost all the supermarkets across Mauritius.

This is an example of what we call a very simplified back-of-envelope difference-in-differences, in which the CCM assumed that, had it not intervened, there would have been no entry, and price would have continued to rise at the same rate as it had prior to the intervention. Clearly, such simple assumptions are really only guesses and open to question. On the other hand, they render the required analysis much less data and time-demanding. Other authorities quite often employ simple before-after comparisons, for example the OFT in its ex-post impact evaluation of the NAPP abuse case (OFT, 2011f).

7.3   Case 3: Pirelli/BICC merger (event study)

The Pirelli/BICC merger study was commissioned by the European Commission (EC) and undertaken by LEAR Consulting (Buccirossi et al, 2006). It used, as an example, the acquisition by Pirelli of six manufacturing plants in Italy and the UK in the power cables industry. The merger had originally been investigated on the basis of fears of collective dominance (coordinated effects) in two specific product

markets (in which Alcatel was the main rival) and single dominance (unilateral effects) in another.

The data employed were taken from the Datastream database on the stock prices of the rivals and customers of the merging parties at three sub-events: when the merger was first proposed, then at EC's announcement that it was going to instigate a Phase 2 investigation, and finally at the clearance decision. On the basis of their results, LEAR concluded that the merger was pro-competitive, that there was a price reduction as a result of the merger, and that EC's clearance decision was correct.

In many ways, this particular study is a classic example of an event study. LEAR itself concluded (2006, p.102) that "Overall this event study has been a successful one, because it has provided some clear results on the effect of the merger. If compared to many of the results generally observed in the literature, the effects that we have estimated are particularly significant." This is certainly true. Indeed, this particular industry should provide fertile ground for the use of an event study since many of the relevant firms are large and quoted on stock exchanges – the latter being a necessary condition for an event study. Having said this, even in such a large scale important industry such as this, only 5 of the 11 rivals were publicly quoted and only 17 of the 29 customers. So at the least this raising concerns about selection bias. Moreover, many of the 22 quoted sample firms were multinational and diversified, raising doubts whether their stock valuations would be sensitive to a merger between two of their rivals in just two countries in only one of their lines of business.

In fact, inspection of the detailed results of this event study reveals that the merger had no impact on rivals (the expectation is that a coordinated effects merger would raise the rivals' profitability) and a mixed impact on customers (the expectation is that an anti-competitive merger would lead the financial market to reduce its assessment of customers' profitability). Indeed, more customers than not appear to have benefited according to the stock market, and to cite extreme examples, Energia's stock increased by 16 per cent, Verbund by 9 per cent and National Grid by 6 per cent at the time of the merger announcement. However, on reflection, these magnitudes appear to be implausibly high – by how much would the price of one input in one market have to fall for the profitability of the using firm to rise by as much as 16 per cent? Moreover, as just noted, there was an absence of significant impacts on rivals – while it is indeed true that this result is consistent with the merger having no adverse competitive impact, it is also consistent with a more cynical interpretation that the merger (whatever its competitive impact) would have no noticeable effect on the future profitability of such large multinational diversified firms. If the latter interpretation is correct, this suggests that the event study may often be impotent in even the most superficially promising cases.

7.4   Case 4: Antitrust fines in South African bread and flour cartels (event study)

Similar to the previous case, this case also employs an event study but applied to estimate the impact of fines on the stock market valuations of infringing firms involved in bread and flour cartels in South Africa (Darj *et al.*, 2011). It has been selected deliberately as a second example of evaluation from a relatively young competition authority outside of Europe.[55] It is also far less technical, and was presumably much less time-consuming than the previous case.

Interestingly, the authors of the study comment: "In South Africa, there are not many listed companies that have been found guilty of contravening the Act and as such there are only a limited number of companies that we can focus on (p. 6)". However in this case three key firms involved in the cartels are quoted on the Johannesburg Stock Exchange, and the authors were able to conduct a straightforward but competent event study analysis. Contrary to event study merger applications, in which the focus is on rival and perhaps customer firms, the objective here was much simpler – to identify the impact of a fine on just the profitability of the firm itself. The authors' chosen event dates were the day on which a fine was imposed and then the day of any subsequent settlement with the Commission. The resulting estimates of impact on future profitability were generally plausible. On average, the reduction in company value over the event window was about 7.5 per cent - fairly typical of studies such as this.

This case is a good example of how an event study can be used successfully in a straight forward to quantify the effects of intervention on the firm concerned – as is possible with fines. Importantly, the impact of a fine may exceed the simple cost of the fine itself because of other potential indirect effects on the firm's future profitability. In principle, these might include a deterrent effect – the firm will be less likely to be involved in future cartels – and the possibility that its customers might be more cost-conscious when dealing with the firm in future. There is no real complication regarding the counterfactual which is fairly obvious in the case of fines.

7.5   Case 5: Volvo-Scandia (simulation)

The Volvo-Scandia case is selected as an example of an academic merger simulation – in fact, a full-fledged technically quite advanced simulation of a proposed merger between Volvo and Scandia in the European truck (lorry) market. The authors, Ivaldi and Verboven, were involved in the original European Commission investigation of the case, but here we cite their account as subsequently described in their article in an academic journal (Ivaldi and Verboven, 2005). That article

---

[55] Other better known cartel cases are already referenced in section 6.

emphasises the methodology more than the results of the evaluation *per se*, and it has since become a widely cited standard reference in Industrial Organisation texts used to illustrate state of the art simulation techniques.

At the heart of their methodology is the econometric estimation of the demand system of equations for trucks. This uses panel data for 16 different Member States over two years. Prices are the list prices of a base model and sales are total sales for the model range. The econometric model used is nested logit – a special case of the random utility function – and it is necessary to include variables reflecting the different characteristics of different models of trucks. However, and this is a potential weakness of their estimation, only one truly technical characteristic was employed – the truck's horsepower. In practice, of course, technical specification (and therefore consumers' willingness to pay) is likely to depend on many other technical features of a particular truck. Perhaps such data were unavailable. Estimation of the demand system provided the authors with estimates of the relevant price elasticities.

The effects of the merger were then simulated by introducing their chosen non-cooperative oligopoly model (i.e. unilateral effects) and analytically solving for (i) the equilibrium prices assuming first no merger (i.e. before) and then (ii) with the merger (i.e. after). In this case the merger was actually prohibited so the difference between the before and after price equilibria generates an estimate of the price rise avoided by the EC's prohibition. It appears, on the basis of this simulation, that the EC made broadly the correct decision.

This case illustrates well many of the standard strengths and weaknesses of merger simulation. It forces the analyst to specify quite explicitly what is the chosen counterfactual, but it is also possible to simulate alternative counterfactuals (oligopoly games). In itself, the rigour of the simulation model also requires the analyst to form a detailed understanding of the nature of competition and demand in the market concerned. On the other hand, results can be very sensitive to alternative assumptions, and data demands can be prohibitive – often requiring approximations (e.g. here the paucity of measures of product characteristics). Finally, when conducted at the level of technical complexity employed here, simulation can be time- and resource-consuming and require highly skilled econometricians.

7.6   Case 6: UK merger review (back-of-envelope simulation)

This final case is chosen as a contrast to the previous one in that the simulation was deliberately much more simplistic but also far less time consuming. It was conducted in a matter of days, rather than months or years.

The chosen study is a broad-based evaluation of eight different mergers, conducted by the consultants Deloitte's for a consortium of UK authorities (2007).[56] Their purpose was to conduct a review of merger decisions under the relatively new Enterprise Act of 2002, in particular, reviewing the decisions and the quality of decision making on which they were based, in the light of subsequent developments in the market. The assessment was based on two approaches. The first was a qualitative survey using interviews with market participants, supplemented by questionnaires and publicly available data. The second was to conduct ex-ante simulation on as many of the cases as possible, simulating what the likely consequences of the mergers would be, but using only information that was available to the authorities at the times of their decisions. The simulations, which were undertaken by the present author, were confined to back-of-envelope simple methods which could be effected quickly and with minimal data requirements.

In the event, given data limitations, simulation was possible for only four of the eight cases. In each of these, the product was effectively homogenous and capacity constraints were typically an issue. For this reason, the model used was a capacity-constrained Cournot model. Given the simplicity of the model and the limited nature of the available data, the purpose of the simulation was not to provide precise quantified estimates of by how much each of the mergers would have raised price, but more to suggest broadly the sort of magnitudes which might have been involved. As it turned out, these simulations were supportive of the authorities' decision, in that they suggested that there would have been very high increases in the case where the merger was prohibited, but only very moderate for the other three, all of which were cleared. In one of those three cases however, the market conditions suggested that, perhaps, the merger might have coordinated effects, and these could have led to much higher price increases. In that case, the Competition Commission judged that coordinated effects were, in fact, unlikely. Here, the simulation confirmed that the clearance decision was correct, but conditional on the assumption that coordinated effects would not occur. Whether or not that assumption could, in principle, be evaluated by an ex-post evaluation, but that was not the purpose of this particular project.

This study illustrates quite well what can, and what can not, be achieved by back-of-envelope simulation. Its strength is in speed and simplicity, and thus the capability of application to a number of different cases in a fairly short period of research time. Its weakness is that it can often only be applied to simple markets (here, in which competition could be described by simple Cournot models, albeit capacity constrained), or in which simple diversion ratios can be used to model product differentiation. Moreover, a danger with simulation models is the spurious sense of precision and accuracy it might encourage. In this particular project, the authors correctly stressed that such simulations should be viewed only as devices

---

[56] The consortium included the Competition Commission, the OFT and the Department of Business, Enterprise and Regulatory Reform.

for showing how different assumptions about the market and the nature of competition might translate into different evaluations of the likely impact of different decisions. In other words, their role is to provide only one of a number of different aids to the decision-making.

The other noteworthy feature of the study was the combination of the quantitative simulations with the qualitative survey. The latter provides a number of additional insights which could not have been identified from a quantitative analysis alone – (i) in three cases there was significant entry which had not been anticipated by the CC; (ii) apparent inconsistency across cases in the threshold number of firms in local markets which the CAs used to action interventions, (iii) a question mark on how best to handle self-supply when defining vertically related markets.

8   Some Important Under-developed Issues

Throughout the report, we have flagged where there are gaps, uncertainties and unresolved issues in the previous literatures and practices. In this penultimate section, we discuss three of the most important in greater depth. These reflect the author's own views on the areas in which further work is particularly required.

8.1   Selection bias and deterrence

Inevitably, evaluation studies focus on documented cases in the public domain, and amongst these, usually the ones where an intervention has actually occurred or been seriously contemplated. As noted frequently in the literature, this raises a variety of doubts about whether such a sample can provide an unbiased estimate of the benefits of competition policy. It also leaves open the problematic issue of how to treat the deterrent effect. This section draws together various strands in a fairly disparate existing literature on selection bias and deterrence into a coherent framework which highlights the most pressing questions for future research.

8.1.1   Some potential sources of selection bias

As stressed throughout this report, one source of bias comes from the fact that different evaluation methodologies are less practicable for some types of markets and interventions (e.g. event studies for markets populated by unquoted or conglomerate firms). In addition, previous literature contains numerous other instances of why evaluation is susceptible to selection bias problems; the following examples illustrate some of the issues.

In the empirical literature on **cartels** (e.g. duration and overcharge) it is widely understood that the samples analysed may be intrinsically biased because they are drawn exclusively from detected investigated cases, which may not be representative of the unknown population of undetected cases.

When the rigour of **merger policy** is evaluated using a sample of *unchallenged* mergers, Carlton (2009) points to an easily overlooked bias. He poses the question: 'suppose we observe that the mean price increase in a sample of unchallenged mergers is negative, can we deduce that the CA is sufficiently (or even over-) strict?' The answer is no because, even with a lax CA (inclined to Type I errors), such a sample will include 'good' mergers alongside any incorrectly permitted 'bad' (price increasing) mergers. He suggests that any such evaluation should more properly compare outcomes with the CA's predictions at the time of its decisions to estimate the systemic bias in enforcement.

It is widely acknowledged that the beneficial **deterrent** effects of competition enforcement are likely to be considerable, probably far outweighing the measurable benefits of the actual caseloads of CAs. It follows that any evaluation of the benefits of policy based only on investigated cases may be a serious underestimate, probably by an order of magnitude; however, we know remarkably little about the magnitudes of this global underestimate.

The remainder of this section introduces a classification scheme designed to structure the various dimensions of potential selection bias and highlight directions for future research.

8.1.2 How much of the iceberg lies below the waterline?

Figure 8.1 suggests a stylised classification scheme to describe the full distribution of all potential competition cases in the population.[57] Some of these cases are deterred and never occur; amongst the undeterred cases, some will be detected but some not; and then, within the detected set, some are investigated and some are not.[58]

This classification is used to illustrate some of the selection problems which would need to be resolved if one is to extrapolate from what we know (the investigated cases) to what we do not (the deterred, undetected and un-investigated parts of the population). We proceed in two steps, considering first the relative sizes (frequencies of cases) of the classes, and then the potential heterogeneity between the classes (measured by, say, mean expected harm), which reflects any potential selection bias.

---

[57] This notion of a well-defined population of potential cases is not unproblematic, given that it is to include all deterred anti-competitive practices. Applying such a classification, analogously, to the law criminalising murder would require one to quantify the number of murders that would be committed were the practice not illegal.

[58] In reality the picture is even more complex. For example investigated cases can be broken down into cases settled by the parties and cases that reach the CA's final decision. Within this latter group there are cases where the CA intervenes and ones without intervention. Impact evaluations are typically based in this latter sub-sample of cases. This means selection issues arise on at least two extra levels.

Figure 8.1   A general classification of potential competition cases

Deterred ω	Undeterred (1-ω)			
	Undetected (1-ϕ)	Detected ϕ		
			Uninvestigated (1-σ)	Investigated σ

Source: Davies and Ormosi (2010)

Relative frequencies

Denoting the conditional probabilities by: deterrence rate (ω), detection rate (ϕ), and investigation rate (σ), it follows that the sample of investigated cases:

- **Property 1**: represents only a (perhaps very) small proportion, (1-ω)*(ϕ)*(σ), of the population of all potentially relevant cases.

- **Property 2**: fails to capture any beneficial deterrent effect.

- **Property 3**: fails to capture the 'missed opportunities' represented by harmful cases that are either wrongly un-investigated or undetected.[59]

The first two are trivial, but the third is often overlooked in evaluation studies.

A back-of-envelope quantification

It remains for future work to establish how rigorously these three properties might be quantified, but to illustrate how information already in the public domain might be employed, we draw on a rare qualitative study by Deloitte (2007), commissioned by the OFT.[60] This involved interviews/telephone surveys of lawyers, economists and companies. The key findings (2007, pp. 7-12) on deterrence are shown in Table 8.1. From the survey of legal advisers, they suggest that, for each merger blocked or modified by the CA, there were at least another five proposed mergers that were abandoned or modified on competition grounds. Their 'multiplier' was slightly smaller for potential abuse of dominance cases, but higher for commercial

---

[59] We abstract for the moment from possible offsetting effects, e.g. deterrence of welfare-enhancing mergers, and incorrect Type II intervention decisions.

[60] The Dutch competition authority (NMa) also conducted a survey (2005) to estimate the scope of merger deterrence. See also OFT (2011c) for a recent updating of Deloitte's (2007) study.

agreements (Article 101). According to their survey of the companies themselves, the reported multipliers were all considerably higher.

Table 8.1   Deterrence multipliers

	Legal survey	Company survey
Mergers	5:1	–
Cartels	5:1	16:1
Commercial agreements	7:1	29:1
Abuses	4:1	10:1

Source: derived from Deloitte (2007)

Turning to undetected cases, Deloitte report that the number of 'under the radar' (i.e. undetected by the OFT) mergers was at least as high as the number which are blocked or modified following intervention by the UK competition authorities.[61]

These estimates should be interpreted with considerable caution (see the qualifications stressed in the report itself), but taken at face value, and arbitrarily assuming a value of 0.5 for σ the investigation rate (in principle, this could be computed from the CA's own records), we can back out estimates of the probabilities as in Table 8.2, and then the population frequencies as in Table 8.3.

Table 8.2   Deterrence and detection probabilities

	ω[62]	φ
Cartels	0.56	0.17[63]
Mergers	0.56	0.5

Source: Author's calculations using the data from Table 8.1.

The three above properties can then be quantified as: (i) the investigated sample is only about 10 per cent of the full population of potential mergers and less than 4

---

[61] Strictly speaking the survey reports that the number of undetected problematic mergers was at least as high as the number of investigated ones, in which case the total detection rate may be much lower than 50%.

[62] This uses the Deloitte finding of 5 deterred cases for each investigated case, $\omega = 5(1-\omega)\phi\sigma$, from which ω can be calculated using the estimations for ϕ, and the arbitrarily chosen 0.5 for σ.

[63] Based on Ormosi (2011) discussed below.

per cent of potential cartels,[64] (ii) there are five (fifteen) times as many deterred as investigated merger (cartels). On the other hand, (iii) there are three (ten) times as many 'missed opportunities' (undetected or un-investigated) as investigated mergers (cartels) cases.

Table 8.3   Calculated frequencies for categories

	Cartels	Mergers
Deterred	0.556	0.556
Undetected	0.368	0.222
Un-investigated	0.0377	0.111
Investigated	0.0377	0.111

Source: Author's calculations using the data from Table 8.1.

Needless to say, these estimates are presented for merely illustrative purposes – they show the implications of the Deloitte study in particular, but we reserve judgement on whether or not they might be credible.

8.1.3 Heterogeneity between case classes (selection bias)

It should also be stressed that the above calculations relate only to the *frequencies* of each class – they do not quantify the relative *amounts* of deterred harm or missed opportunities. This would only be true if the expected harm of cases in all categories were identical. However, the possibility of selection bias raises suspicion that expected harm will differ systematically between the classes of case in Figure 1.[65]

First, it is important to note that the distinction between categories is not clear cut, and that the frequencies and make-up of the classes are endogenous to the policy decisions of the CA. Consider mergers as an example. In jurisdictions where there is a compulsory pre-merger notification regime (most of the world) a regulatory threshold demarcates the cases that do not have to be notified. Above the threshold there are three types of cases: (1) those where there is only a trivial preliminary screening, (2) phase 1 cases; (3) phase 2 cases. Although all cases above the

---

[64] Even if we assume that all known cases are investigated (i.e. $\sigma = 1$) these ratios will still only be 22% and 10% respectively.

[65] The OFT (2010c) p.22. speculates about applying a multiplier to account for the deterrent effect of its enforcement activity, where the value of the multiplier is derived from the Deloitte survey. This method implicitly assumes that the distribution of case types in the investigated and deterred classes are similar.

threshold theoretically have to be investigated (i.e. the CA has no discretion), it is arguable whether a simplistic screening of the merger (or even a phase I review, as discussed below) should qualify as an investigation. So the question for mergers is: where to draw the line below which we consider cases as un-investigated.[66]

The possibility that CA discretion may have an important distortional impact on evaluation is not confined to just mergers, and the choices made by the CA are an integral part of its policy and should therefore also form part of any evaluation. The key idea here is that a CA is confronted with a set of potential cases where it could assemble sufficient information to undertake a full investigation (and subsequently possibly intervene) but, in the event, it chooses not to pursue them all, for example because there is a low success rate if the case is appealed.[67]

This possibility cannot be excluded as one possible explanation for the previously highlighted scarcity of evaluations of abuse of dominance cases – perhaps CAs choose to intervene less frequently in such cases. Similarly for cartels, Chang and Harrington (2010) argue that the CA's decision to investigate a case is a matter of CA choice, and success is negatively correlated with the CA's workload.

Of course, given finite resources, coupled with a need to substantiate its impact, it is rational for any CA to pursue the 'easy options'; i.e. easier cases at the expense of more difficult cases (for which the probability of 'success' is lower or where there is greater uncertainty). This may have serious implications for selection bias – both for policy in general and in specific parts (e.g. abuse of dominance). We believe that this issue merits further research.

A related danger is highlighted by Neven and Zengler (2008, p.477), who suggest that the evaluation programme itself may introduce an additional motive for distortional discretion in CA conduct: "Faced with simplistic assessment, authorities may be tempted to be overly interventionist, to spend too many resources and to ignore relevant information". This returns us to the question of what is the purpose of evaluation, and who undertakes it. If evaluation is driven by external accountability (to verify whether the CA delivers its objectives) and especially if undertaken by the CA itself, the CA is prone to fall into the trap identified by Chang and Harrington (2010), that it will not seek to maximise deterrence, but focus on something that is observable/measurable (e.g. the proportion of Article 102 cases that are won, or the number of cartels detected). This could easily mean that enhanced deterrence is not in the interest of the CA if that were to reduce its caseload – and consequently the number of cases won. Although this might seem a questionable assumption, there is some intuition to support it:

---

[66] In non-mandatory pre-merger notification regimes (such as the UK) the situation is different as the CA has more discretion whether to investigate mergers.

[67] Neven (2006) reports a 98% success rate for Abuse of dominance cases (as opposed to a 75% success rate for Article 101 and 58% for merger cases).

deterrence is difficult to measure and therefore include in any performance measure of the CA. If the CA's budget is based on previous years' performance, which excludes deterrence, then the assumption becomes more plausible.

8.1.4 Existing academic literature on detection and deterrence

The task of calibrating the three parameters and quantifying likely magnitudes of associated selection biases is formidable. In our opinion the investigation rate is potentially the easiest to estimate, and could be acquired from the CA's own records. Detection and especially deterrence are more problematic but there have been attempts to estimate the probability of both of these categories. These are summarised in Table 8.4.

The detection rate is not directly observable of course, but a body of emerging research offers some promise that indirect estimation may be possible using relatively atheoretical statistical methods. This research also has implications for the rate of deterrence, but may involve additional theoretical modelling of the creation of mergers, cartels and other business practices. More work is clearly needed – for example, on whether observed CA activity and latent deterrence are complementary or substitutes.[68] Similarly, little work has been done on determining the key drivers of antitrust deterrence. As described earlier, the results from Broek, Kemp et al. (2012) suggest that deterrence may be driven by reputational factors as well as the legal penalty per se. More generally, there may be fruitful avenues of research drawing on the wider legal deterrence literature. For instance, the critique by Donohue and Wolfers (2006) - of the common mistakes made in the statistical analysis of the deterrent effect of the death penalty - might be helpful.

Finally, it should be acknowledged that over-stringent policy can sometimes deter welfare *enhancing* practices or proposals. This would include mergers which might lead to strong efficiency effects; vertical practices (perhaps RPM) where the beneficial efficiency gains outweigh any anti-competitive foreclosure effect; or even cases where firms may refrain from fierce price competition because they are deterred by predatory pricing policy. This is rarely even discussed in CA evaluations, although admittedly quantification would be difficult. An interesting example is given by Eckbo (1992), who finds no evidence of Canadian mergers having anticompetitive effects. If so, it follows that any deterred merger must be either a welfare loss, or at least no gain.

---

[68] This might be approached statistically by comparative analysis of high and low deterrence authorities.

Table 8.4   Previous attempts to estimate investigation, detection and deterrence rates

	Probabilities of interest	Source of data	Literature examples
Observed	Investigation rate	CA's records	–
Unobserved	Detection rate	Atheoretical empirical methods	Bryant and Eckard (1991) Ormosi (2010)
		Theory-based empirical methods	Miller (2009) Harrington et al. (2009) Chang et al. (2010)
		Surveys	OFT (2007) Beckstein and Landis-Gabel (1982)
Unobserved	Deterrence rate	CA's records	Barros et al. (2010)
		Theory-based empirical methods	Block et al. (1981) Harrington et al. (2009) Cheng et al. (2010)
		Surveys	NMa (2005)

Source: Author's compilation

8.2   The need for longer term studies

In general it is difficult to assess the long-run effect of any public policy, given the exponential rise of possible compounding effects with time, which makes it difficult to distinguish effects that were caused by the intervention from effects triggered by exogenous factors. One potential way of doing this is through case studies regarding the evolution of markets following regulatory intervention.[69] This could highlight concerns about the time dimension which applies to any one-off evaluation (no matter what the methodology is) – that it runs the risk of closing the story prematurely. The wider IO literature (both theoretical and empirical) suggests various possibilities for how a specific event might trigger a sequence or chain of subsequent events – each of which might be evaluated independently, but which are in reality clearly path-dependent. For example, the literature on endogenous mergers alerts us to the possibility that, if merger A is cleared, this makes a subsequent merger B more or less likely. Similarly, in failing firm merger cases, the consequences of intervention may include subsequent alternative merger proposals

---

[69] Another way would be to examine how competition policy shaped firms' long-run decisions that are formed immediately following the policy intervention. This way the scope of possible compounding effects is reduced to minimum and we get some sort of idea about longer-run impacts. To define long-run decisions for example Symeonidis (2002) refers to those decisions that determine the set of conditions that firms must take as given when making short-run choices. One example would be firms investing in long-run R&D projects following competition intervention.

by other parties as happened in the case of Airtours.[70] Airtours had to change its brand name (My Travel group) following the blocking of its takeover attempt on First Choice. Eventually My Travel was acquired by Thomas Cook Travel.

There is case study evidence which suggests that sometimes when some practice is prohibited, it is replaced by others. Clarke *et al.* (1998) cite examples from the UK where firms responded to the prohibition of one form of vertical restraint by introducing an alternative form, or where a prohibited restraint was replaced by full-fledged integration. In other cases, such as following intervention in exclusionary abuse of dominance the market may trigger market entry in the longer-run. For example following the US intervention in the bundling of Microsoft Internet Explorer web-browser and the Windows operating system the market of web-browsers now accommodates three larger firms (Internet Explorer, Firefox, and Google Chrome) and a fringe with a number of other browsers.

It has long been recognised that horizontal mergers may sometimes be an alternative to cartelisation. Symeonidis (2002) shows that cartel legislation in the UK in the 1950s provoked a subsequent merger wave through the 1960s. Similarly, following the 2002 detection of a cartel, which included Mittal Steel and Arcelor (the first and second-largest steel producers in the world), in 2006 the two firms merged.[71]

Sidak and Teece (2009) provide an insightful discussion on why competition analysis has been so one-sidedly fixated on static effects. They draw upon recent developments in evolutionary economics, the behavioural theory of the firm, and strategic management to propose a more dynamic and robust support for competition economics. Although their discussion focuses mostly on enforcement issues, much it has direct relevance for ex-post impact evaluation as well.

Finally, liberalisation and privatisation (and potentially the advocacy activities of CAs that may trigger liberalisation) will also have long as well as short-term impacts. Sector inquiries by the EC are prime examples of evaluation that looks back to market interventions such as liberalisation and examines how the market evolved in the medium and longer run.[72] For example the EC report on leased telecommunication lines found that the 1998 full liberalisation of the EU telecommunications infrastructure and services brought about greater choice for

---

[70] The Airtours/First Choice merger was blocked in 2000 by the EC although the decision was overturned by the ECJ in 2002 – although the deal by then had been long dead.

[71] Although Kumar, Marshall et al. (2011) claim that ceteris paribus, firms choose cartels over mergers, suggesting that they perceive the payoffs to a cartel to be higher than the payoffs to the merger of the same firms.

[72] The EC has so far published sector inquiries in the following markets: pharmaceuticals, financial services, energy, local loop, leased lines, roaming, and media (3G). For more information see: http://ec.europa.eu/competition/antitrust/sector_inquiries.html

users, lower prices and better quality of services across the EU.[73] At the other extreme, the inquiry into the European energy sector, found that years after the liberalisation of the internal energy market, barriers to free competition remain.[74] OECD Country Reviews also frequently provide a longer perspective analysis of liberalisation and privatisation activities in the past.

To summarise, it matters what happens after the intervention, and 'after' should sometimes be interpreted as long-term and not too narrowly. The importance of looking at longer effects is unquestionable. As Werden (2008) pointed out citing Easterbrook (1992): "An antitrust policy that reduced prices by 5% today at the expense of reducing by 1% the annual rate at which innovation lowers the cost of production would be a calamity. In the long run a continuous rate of change, compounded, swamps static losses." Some evaluation studies have acknowledged this, albeit indirectly, by considering sequences of cases, for example Sabbatini (2008) for Italian baby-milk; Pinske and Slade (2004) for a sequence of mergers in UK beer and Nevo (2000) who includes various different mergers in his simulations, but these are the exception rather than the rule.

8.3   The broader impact of competition policy

Most of the literature, and therefore discussion, in this report has focussed rather specifically on the impact of enforcement on consumer welfare. However, we should also be concerned with the wider impact on the economy as a whole and the macro aggregates of growth, investment, employment as well as productivity and innovation.

There are empirical studies which explore these areas, but in general the literature is thin and of variable quality. This section provides a brief overview, distinguishing between two strands of literature: works that examine the impact of *competition* on productivity, growth, etc, and research that attempts to measure the impact of *competition enforcement* on the same factors.

8.3.1 The effects of competition

There are grounds for expecting that competition can help fuel economic growth through at least three different channels: (i) exerting pressure on firm management to reduce inefficiencies (Meyer and Vickers (1997), Schmidt (1997); McKinsey

---

[73] http://ec.europa.eu/competition/sectors/telecommunications/archive/inquiries/leased_lines/index.html

[74] Inquiry pursuant to Article 17 of Regulation (EC) No 1/2003 into the European gas and electricity sectors.

Global Institute (2010b, p.28)); (ii) by fostering innovation,[75, 76] and (iii) by guiding a better allocation of resources to more efficient firms (Arnold *et al.*,2008). It can also help create a more favourable macroeconomic environment in various other ways, such as through reducing inflation rates,[77] or improving international competitiveness.[78]

The main challenge in this literature is to find an adequate measure of competition.[79] Simple measures such as market concentration do not always provide a reliable stand-in for competition.[80] Price-cost margins (PCM) have been a popular measure although not fully supported by economic theory.[81] Boone (2008) offers an alternative to PCM; a measure that has a more robust theoretical backing and has the same data requirements as for PCM. One frequently used way of measuring competition follows the seminal paper by Nickell (1996), who estimates the impact of competition on firm performance. To proxy for competition, Nickell used a mix of measures including market share, concentration, import penetration, a survey-based measure of competition, and a measure of average rents.[82] Another way of measuring product market competition is given by Przybyla and Roma (2005) who use a different mix of measures, which includes mark-up measured as the inverse of the labour income share in the economy, profit margin as the ratio of operating surplus to output, profit rate as net unadjusted operating surplus/capital stock, a World Economic Forum (2002) index number, the level of regulation, and market openness. Finally, Haffner *et al.* (2000) use simple measures such as similarities and convergence over time of price structures, differences in price levels and estimates of the levels and trends of profit margins. The rationale behind this is the intuition that if competition was intensive then cross-border competitive pressure would result in prices converging at a European level. Another frequently used 'proxy' for competition is trade liberalisation, which is used in time-series or

---

[75] Gilbert (2008) surveys economic theory and empirical studies on the relation between market structure innovation.

[76] In a survey of more than 26,000 manufacturing establishments across 71 countries Dutz et al. (2011) found empirical evidence that innovation drives employment growth and that the underlying innovations are fostered by a pro-competitive business environment. .

[77] Przybyla and Roma (2005).

[78] Mitschke (2008).

[79] Aghion and Griffith (2005) give a detailed overview of the development of econometric techniques aimed at measuring competition, productivity, and innovation.

[80] See for example Boone (2001).

[81] See Boone (2008) for a brief discussion on this matter.

[82] Nickell finds that competition is associated with higher rates of total factor productivity growth.

longitudinal structural analyses. Most of these studies find robust evidence for the positive effect of competition on firm performance, growth and productivity.[83]

8.3.2 The effects of competition enforcement

One example of a study of this type is the OFT (2011b) who examine the impact of competition interventions on various elements of economic growth, such as natural resources, capital, innovation, and management. A key challenge in this line of literature is to find appropriate measures of competition enforcement or competition policy.[84] Buccirossi *et al.* (2009a) warn of this and claim that the lack of evidence of a positive effect of competition policy in some of the studies is a result of inadequate measures. In this respect, competition policy indices, such as the ones in World Economic Forum (2002), Buccirossi *et al.* (2009a), and Hüschelrath and Leheyda (2010) may be better suited to capture those characteristics that are likely to have an impact on the effectiveness of policy. The composition of a competition policy index typically relies on institutional and enforcement characteristics such as the degree of formal independence of the CA, the separation of adjudicatory and prosecutor functions within the CA, the scope of investigative powers of the CA, the level of enforcement (size of sanctions, size of budgets and resources, etc)., and the seriousness of sanctions that the CA can impose.[85] The related econometric problems are familiar in any international comparisons based on production functions or related concepts, e.g. identification, simultaneity and the requirement that the underlying functional forms are stable across countries.[86]

Competition policy, when interpreted more widely than just the enforcement of competition laws, has also contributed to economic growth. This includes government policies on facilitating market entry, for example by liberalising or privatising markets. European-wide market liberalisation for example has led to greater competition and consequently reduced prices across Europe.[87] More specifically there is evidence that market liberalisation has contributed to

---

[83] See OECD (2011). For a review of studies on non OECD member countries see: Tussie and Aggio (2006) and Parikh and Stirbu (2004).

[84] Some of these aggregate studies are discussed and well summarised by Bergman (2008), who provides some of the key references.

[85] Which reflects the understanding that the effectiveness of competition law enforcement largely depends on the strength of legal systems, the quality of the judiciary system, and other indicators of legal origin (La Porta, de Silanes et al. (2008)).

[86] One area where more work would be welcome is a comparison of the impact of competition enforcement on socio-economic factors depending on whether the CA's decisions are rooted in static or dynamic economic analyses. Following the arguments from Sidak and Teece (2008) one would expect that the latter should show a more positive impact.

[87] McKinsey Global Institute (2010a)

innovation and productivity, for example in the UK,[88] India,[89] or China.[90] Reducing barriers to entry and trade, and eliminating price controls also increases employment rate through increased demand in markets.[91]

The impact of competition policy in the geographically broader sense is a yet unexplored area, where more work is long overdue. Competition policy impacts conducts that are turning increasingly global in their scope. Therefore the enforcement of competition law and the design of competition policy in one country will inevitably have an effect on other economies, although effects of competition policy in one country may be difficult to disentangle from the effects of competition policy from another country. The challenges of measuring this type of impact are similar to some of the already discussed issues.

---

[88] Aghion, Blundell et al. (2009).

[89] Sivadasan (2009)

[90] Zheng and Ward (2011)

[91] Griffith, Harrison et al. (2006)

9 Conclusions

By way of a conclusion, I turn to three questions on which I have been asked to advise as part of my research brief: (i) why should a CA carry out evaluations of its own work? (ii) how should it do it? (iii) what are the challenges that CAs face in the field? In each case, my answers necessarily express personal opinions, and while they are based on a fairly extensive experience in the field, they are the inevitably opinions of someone who is obviously committed to the evaluation project.

9.1 Why should CAs evaluate their own work?

As explained in the opening section of this report, evaluation has been carried out for a variety of reasons, some of which are clearly in the self-interest of Competition Authorities. Here, I will highlight three.

First, on the assumption that competitive markets are an essential ingredient of an efficient, innovate and flexible economy, there is clearly a role for active competition law and policy. In turn, given that governments face a multitude of potential demands on the public purse, it is important that the enforcement of competition policy (broadly defined) should attract an appropriate share of the state's budget. The evidence from the competition authorities who have been most prominent to date in evaluating competition policy (the UK, US and EC), is that the enforcement of competition law does indeed provide exceptional 'value for money.' For example, the OFT currently estimates that every £1 it spends yields approximately £10 of benefits to consumers. While one can always argue about and perhaps question the details involved in deriving such a figure, the sheer magnitude speaks for itself – especially when it is remembered that conservative assumptions are used throughout and, crucially, that no allowance has been made for the probably substantial additional benefits deriving from deterrence.

In my opinion, it is important that the value of competition policy is appreciated beyond the walls of the competition authorities, and that any competition authority should be able to confidently and justifiably claim that the benefits deriving from its activities far exceed the costs involved.

Second, at a more micro level, regular and fairly comprehensive evaluation, should provide an essential ingredient in internal resource allocation decisions within the authority. When allocating its budget between its different activities (merger control, detection of cartels, monitoring of dominant firms), it is important to know what is the likely marginal product of spending an additional euro in each area. More generally, allocational decisions – say between detection and deterrence - should be informed by some notion of the potential pay-offs.

Third, in any busy competition authority, there is an inherent danger that once a decision or intervention has been made, it is consigned to a file which is then never again opened. Ex-post evaluation provides probably the best means for ensuring that the authority becomes aware of 'what happens next' in a market. Of particular importance is the question of how markets react to, and perform, after interventions. Moreover, evaluation provides a valuable check of whether the assumptions, analysis and decision making taken at the time of the decision vindicated by what happens in subsequent years.

However, there are some caveats, two of which deserve special emphasis. The first is that we should not become slavishly attached to the precise estimates derived in any evaluation. By its nature, evaluation will always be an imprecise subject, based inevitably on assumptions which are typically approximate and sometimes (regarding the counterfactual) never verifiable on a purely empirical basis. Evaluation is there to guide rather than dictate. For this reason, it may always be prudent to avoid giving a spurious impression of precision by presenting estimates of impact in the form of a range rather than a precise single figure.

Second, and relatedly, there is an inherent danger that over-reliance on impact estimates can distort allocational decisions within an authority. Faced with a corporate objective of maximising the benefits to cost ratio, the danger is that certain activities may go relatively neglected because their impact is more difficult to measure. This is most obviously true for policies designed to enhance deterrence – almost by definition, an anti-competitive act which is deterred will not show up in any output measure. But, beyond this, certain policy areas are more difficult to evaluate than others, and there is a danger that these may tend to be neglected if the CA attaches over-importance to being able to measure impacts. This is certainly true for advocacy and education, and perhaps also Article 102 cases. Hopefully this position may be at least partly redressed if the evaluation of advocacy begins to attract more attention in the future.

9.2 How should CAs conduct evaluation?

There have been two broad themes running through this report – deriving global estimates of the aggregate impact of an authority's activity, and then more specifically the techniques which can be used to evaluate individual cases. It is my view that the competition authority should do both – regularly (perhaps annually) evaluating its overall impact and also devoting some resources to in-depth ex-post evaluations of at least some specific cases. This is very much the model currently used by the OFT.

There are a number of issues and decisions which would need to be resolved in deriving aggregate impact evaluations – which rules of thumb to use for cartel overcharge and duration etc., which areas of activity to evaluate? Here, the existing

practices of the OFT, DGCOMP, FTC and DoJ might be adopted, but, in the absence of any internationally agreed best practice,[92] this need not be essential.

Turning to the specific techniques of evaluation, Chapters 6 and 7 have shown that there is no such thing as an ideal technique. Simulation, event studies, difference-in-differences, before-after and more qualitative surveys all have strengths and weaknesses, and each is more suited to some cases than others. In an ideal world, it would be best practice to evaluate any given case using more than one different technique to seek to establish robust findings. However, cost-, data- and time-constraints will typically preclude this option. Perhaps the strongest message to draw from this report is that sufficient time and attention should be devoted to formulating an appropriate counterfactual (or alternative counterfactuals). Sometimes, the nature of the counterfactual will dictate which evaluation technique should be used.

A third broad recommendation is that the authority should adopt what might be called a portfolio approach to its evaluation activities. Thus, there is much to be said for employing a mix of ex-ante and ex-post evaluation - as do the other authorities most frequently discussed in this report. Similarly, if only because of cost-constraints, some of the evaluation should employ 'back-of-envelope' techniques, while others might be more thorough and technically advanced. Finally, there are advantages in commissioning outsiders (consultancies and academics) to do some, but not all, of the research. The obvious problem with in-house evaluation is the suspicion that it will be seen as self-justifying and without objectivity. On the other hand, it is important that the authority should develop its own technical capabilities – not only is this good for the morale of the CA's staff, but also evaluation can generate important externalities (in developing human capital and experience) for the CA's core enforcement activities.

9.2   What are the challenges?

In Chapter 8, I have already identified three areas in which I think further work is required in general. First, very little is known about the scale of undetected cases or the strength of deterrence. This means that we have only a partial understanding of the CA's full impact or, for that matter, of the scale of its missed opportunities. Second, most evaluation to date has focused on the short-term, with very few analyses of the longer-term impact, taking into account how intervention will often provoke responses by the firm in the markets concerned. Third, evaluation is often confined to just price effects, with the impact on quality, choice, innovation, productivity far less frequently assessed. This emphasis on price is understandable – partly on practical grounds and partly because competition law itself emphasises

---

[92] The author is aware that the OECD has recently embarked on its own evaluation research programme, and the possibility of establishing international best practices will be considered within this.

consumer surplus. However, potentially it seriously underestimates the true impact of competition policy and this makes it sometimes difficult to respond to demands, say from government, for quantifications of how much competition policy contributes to growth and productivity.

More specifically, and confined to the activities that *are* currently assessed and the techniques already used, the evaluation project is still very much 'work in progress'. Evaluation in some areas of policy is still under-developed, e.g. advocacy, education and Article 102. Little is known, in a comparative sense about the predictive performance of the different techniques. Here, the challenge is to widen the scope of evaluation without unduly increasing the cost. As is true for many policy areas, there is an important role for international cooperation and diffusion of experience and skills.

10   References

Aghion, Philippe, Richard Blundell, Rachel Griffith, Peter Howitt and Susanne Prantl. 2009. The Effects of Entry on Incumbent Innovation and Productivity - *Review of Economics and Statistics*, 91, 1, 20-32.

Aghion, Philippe and Rachel Griffith. 2005. *Competition and Growth: Reconciling Theory and Evidence*, The MIT Press.

Arisi, E. and F. M. Esposito. 2007. *'L'attivita di advocacy dell'Autorita Garante della Concorrenza e del Mercato: un bilancio dei primi 15 anni (Advocacy by the Italian Competition Authority: an overview over the first fifteen years)'*, unpublished paper referenced in Sabbatini (2008).

Arnold, Jens, Giuseppe Nicoletti and Stefano Scarpetta. 2008. 'Regulation, Allocative Efficiency and Productivity in OECD Countries: Industry and Firm-Level Evidence', *OECD Economics Department Working Papers*, No. 616.

Ashenfelter, Orley and Daniel Hosken. 2008. 'The effect of mergers on consumer prices: evidence from five selected case studies', *NBER Working Paper*, Nº 13859.

Baker, Jonathan B. 2003. 'The Case for Antitrust Enforcement', *Journal of Economic Perspectives*, 17, 4, 27-50.

Barros, Pedro P., Joseph A. Clougherty and Jo Seldeslachts. 2010. 'How to Measure the Deterrence Effects of Merger Policy: Frequency or Composition?', *International Journal of the Economics of Business*, 17, 1, 1-8

Beckstein, A., and H. Landis-Gabel. 1982. 'Antitrust Compliance: Results of a Survey of Legal Opinion'. *Antitrust Law Journal*, 52, 459-468.

Bergman, Mats A. 2008. 'Quis Custodiet Ipsos Custodes? or Measuring and Evaluating the Effectiveness of Competition Enforcement', *De Economist*, 156, 4, 387-409.

Becher, D. A. 2009. 'Bidder returns and merger anticipation: Evidence from banking deregulation', *Journal of Corporate Finance*, 15, 85-98.

Beverley, L. 2008. 'Stock Market Event Studies and Competition Commission Inquiries', *CCP Working Paper No. 08-16*

Bittlingmayer, G. and T.W. Hazlett. 2000. (DOS *Kapital*: Has antitrust action against Microsoft created value in the computer industry?', *Journal of Financial Economics*, 55, 329-359.

Bizjak, J. M and J. L. Coles. 1995. 'The Effect of Private Antitrust Litigation on the Stock-Market Valuation of the Firm', *American Economic Review*, 85, 3, 436-461.

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.

Boone, Jan. 2001. 'Intensity of competition and the incentive to innovate', *International Journal of Industrial Organization*, 19, 5, 705-726.

Boone, Jan. 2008. 'A new way to measure competition', *The Economic Journal*, 188, 1245-1261.

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

Bosch, Jean-Claude and E. Woodrow Eckard. 1991. 'The Profitability of Price Fixing: Evidence from Stock Market Reaction to Federal Indictments', *Review of Economics and Statistics*, 73, 309–317.

Broek, S. v. d., R. Kemp, et al. 2012. 'Reputational Penalties to Firms in Antitrust Investigations', *Journal of Competition Law & Economics*, Forthcoming.

Bryant, Peter G. and E. Woodrow Eckard. 1991. 'Price Fixing: The Probability of Getting Caught', *The Review of Economics and Statistics*, 73, 3, 531-536.

Buccirossi, Paolo, Lorenzo Ciari, Tomaso Duso, Sven-Olof Fridolfsson, Giancarlo Spagnolo and Cristiana Vitale. 2006. 'Ex-post review of merger control decisions', *European Commission*.

Buccirossi, Paolo, Lorenzo Ciari, Tomaso Duso, Sven-Olof Fridolfsson, Giancarlo Spagnolo and Cristiana Vitale. 2008. 'A Short Overview of a Methodology for the Ex-Post Review of Merger Control Decisions ', *De Economist*, 156, 4, 453-475.

Buccirossi, Paolo, Lorenzo Ciari, Tomaso Duso, Giancarlo Spagnolo and Cristiana Vitale. 2009a. 'Competition Policy and Productivity Growth: An Empirical Assessment', *CEPR Discussion Paper*, No. 7470.

Buccirossi, Paolo, Lorenzo Ciari, Tomaso Duso, Giancarlo Spagnolo and Cristiana Vitale. 2009b. 'Measuring the Deterrence Properties of Competition Policy: the Competition Policy Indexes', *Journal of Competition Law & Economics*, 7, 1, 165–204.

Budzinski, Oliver & Ruhmer, Isabel. 2010. Merger Simulation in Competition Policy: A Survey, in: *Journal of Competition Law & Economics*, Vol. 6 (2), pp. 277-320.

Carletti, Elena, Philipp Hartmann and Steven Ongena. 2011. 'The Economic Impact of Merger Control Legislation', *CEPR Discussion Paper* No. DP8447.

Carlton, Denis. 2009. 'Why we need to measure the effect of merger policy and how to do it', *Competition Policy International*, 5, 1 (Spring), 76-90.

Chang, M. H. and J. E. Harrington. 2010. 'The Impact of a Corporate Leniency Program on Antitrust Enforcement and Cartelization', available at: http://www.econ.jhu.edu/pdf/papers/WP548.pdf.

Cichello, C. and D.J. Lamdin. 2006. 'Event Studies and the Analysis of Antitrust', *International Journal of Economics and Business*, 13:2, 229-245.

Clarke, Roger, Stephen W. Davies and Nigel Driffield. 1998. *Monopoly policy in the UK: Assessing the evidence.* Cheltenham, Edward Elgar.

Colley, Liam and Marsden Philip. 2010, 'Use of the Counterfactual in Antitrust', *Competition Law Forum*, at: http://www.biicl.org/files/5106_the_use_of_the_counterfactual_in_antitrustv2.pdf

Competition Commission of Mauritius. 2011. Evaluation of CCM case

Connor, Robert A. , Roger D. Feldman and Bryan E. Dowd. 1998. 'The effects of market concentration and horizontal mergers on hospital costs and prices', *International Journal of Economics and Business*, 5, 2, 159-180.

Cooper, Z., S. Gibbons, et al. 2010. 'Does Hospital Competition Improve Efficiency? An Analysis of the Recent Market-Based Reforms to the English NHS', *CEP Discussion Paper* No 988.

Cox, A. J. and J. Portes. 1998. Mergers in Regulated Industries: the Uses and Abuses of Event Studies, *Journal of Regulatory Economics*, 14, 3, 281-304.

Da Graca, T. 2006. 'Simultaneous Mergers Probability Event Study '.

Darj R, Grimbeek S, and Muzata G. 2011. *The impact of antitrust fines on firm valuation in South Africa: the case of Pioneer Foods, Tiger Brands and Sasol Chemical Industries*, Competition Commission of South Africa

Davies, Stephen and Peter L. Ormosi. 2012. 'A Comparative Assessment of Methodologies used to Evaluate Competition Policy, *Journal of Competition Law and Economics* (forthcoming)

Davies, Stephen and Peter L. Ormosi. 2010. 'Assessing Competition Policy: Methodologies, Gaps and Agenda for Future Research', *CCP Working Paper*, 10, 19.

Davies, Stephen W. 2010. 'A Review of OFT's Impact Estimation methods', *Office of Fair Trading*, OFT1164.

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. and Matthew Olczak. 2008. 'An evaluation of the impact upon productivity of ending resale price maintenance on books', *Office of Fair Trading*, OFT981.

Davis, Peter and Eliana Garces. 2009. 'Quantitative Techniques for Competition and Antitrust Analysis', *Princeton University Press*.

Deloitte. 2007. '*The Deterrent Effect of Competition Enforcement by the OFT'*, OFT962.

Deloitte 2009. *Review of merger decisions under the Enterprise Act 2002*, A review prepared for the Competition Commission, the Office of Fair Trading, and the Department for Business, Enterprise and Regulatory Reform.

Diepold, Brian, Robert Feinberg, David Round and Jeremy Tustin. 2008. 'Merger Impacts on Investor Expectations: An Event Study for Australia', *International Journal of the Economics of Business* 15, 1, 45-62.

Dobson, P. W. and C. A. Piga. 2011. The Impact of Mergers on Fares Structure: Evidence from European Low-Cost Airlines. *Economic Inquiry* (forthcoming)

Donohue, J. J. and J. Wolfers. 2006. 'Uses and abuses of empirical evidence in the death penalty debate', *Stanford Law Review*, 58, 791.

Duso, Tomaso, Klaus Gugler and Burcin Yurtoglu. 2010. 'Is the Event Study Methodology Useful for Merger Analysis: A Comparison of Stock Market and Accounting Data', *International Review of Law and Economics*, 30, 186-192.

Duso, Tomaso, Klaus Gugler and Burcin B. Yurtoglu. 2011. 'How Effective is European Merger Control?', *European Economic Review*, 55, 7, 980-1006.

Duso, Tomaso, Damien J. Neven and Lars-Hendrik Roller. 2007. 'The Political Economy of. European Merges Control: Evidence using Stock Market Data', *Journal of Law and Economics*, 50, 455-489.

Dutz, Mark A., Ioannis Kessides, Stephen O'Connell and Robert D. Willig. 2011. 'Competition and Innovation-Driven Inclusive Growth', *The World Bank Policy Research Working Paper*, WPS5852.

Easterbrook, F. 1992. Ignorance and Antitrust. in: Jorde, T. and D. Teece, *Antitrust, Innovation, and Competitiveness*. Oxford University Press, New York/Oxford: 119-136.

Eckbo, B. E. 1985. 'Mergers and the Market Concentration Doctrine: Evidence from the Capital Market', *The Journal of Business*, 58, 3, 325-349.

Eckbo, B. E. 1992. 'Mergers and the Value of Antitrust Deterrence', *Journal of Finance*, 47, 1005-1029.

Ehmer, C. and F. Rosati. 2009. 'Science, Myth, and Fines: Do Cartels Typically Raise Prices by 25%?', *Concurrences* No 4, 1-5.

European Commission. 2011. *'Report on Competition Policy 2010'*, DGCOMP, {SEC(2011) 690 final}

European Commission DGCOMP. 2010. '*Annual Activity Report 2010'*

European Commission DGCOMP. 2010. '*DG Competition Stakeholder Study, Aggregate Report'* at
http://ec.europa.eu/competition/publications/reports/aggregate_report_en.pdf

European Commission DGCOMP. 2005. '*Merger Remedies Study*'

Evenett, S. J. 2006. 'Competition Advocacy: Time for a rethink', *Available at: www.alexandria.unisg.ch/EXPORT/DL/Simon_Evenett/28424.pdf*.

Farrell, Joseph and Carl Shapiro. 1990. 'Horizontal Mergers: An Equilibrium Analysis', *The American Economic Review*, 80, 1, 107-126.

Farrell, Joseph and Carl Shapiro. 2010. 'Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition', *The B.E. Journal of Theoretical Economics, Policies and Perspectives*, 10, 1, Article 9.

Fonseca, M. A. and H.-T. Normann. 2009. 'Mergers, Asymmetries and Collusion: Experimental Evidence', *The Economic Journal*, 118, 527, 387-400.

Gilbert, Robert J. 2008. Competition and Innovation. in: Collins, Wayne D., *Issues in Competition Law and Policy*. ABA Section of Antitrust Law: Chapter 26.

Griffith, Rachel, Rupert Harrison and Gareth Macartney. 2006. 'Product Market Reforms, Labour Market Institutions and Unemployment', *Institute for Fiscal Studies*, WP06/06.

Haffner, Robert C.G., Stephen Nickell, Giuseppe Nicoletti, Stefano Scarpetta and Gylfi Zoega. 2000. 'European integration, liberalisation and labour market performance', *Report for the fondazione Rudolfo DeBenedetti*.

Harrington, J. E. 2004. 'Post-Cartel Pricing During Litigation', *Journal of Industrial Economics*, 52, 517–533.

Harrington, J. 2006. *The Anatomy of a Cartel Price Path: Theory Meets Practice.* Keynote Speech.

Harrington, J. E. 2008. "Optimal Corporate Leniency Programs," *Journal of Industrial Economics*

Harrington, J. E. and M.H. Chang. 2009. 'Modelling the Birth and Death of Cartels with an Application to Evaluating Antitrust Policy', *Journal of the European Economic Association*, 7, 6, 1400-1435.

Hausman, J. and G. Leonard 1997 'Economic Analysis of Differentiated Products Mergers using real world data', *George Mason Law Review*, 5, 3, 321-346.

Henk, Don, Kemp Ron and Sinderen Jarig. 2008. 'Measuring the Economic Effects of Competition Law Enforcement', *De Economist*, 156, 4, 341-348.

Hurley, S. 2010. 'The use of surveys in merger and competition analysis', *Journal of Competition Law & Economics*, 7, 1, 45–68.

Hüschelrath, Kai and Nina Leheyda. 2010. 'A Methodology for the Evaluation of Competition Policy', *ZEW Discussion Paper* No.10-081.

International Competition Network. 2002. *Advocacy and competition policy*, report prepared by the Advocacy Working Group, ICN's Conference, Naples, Italy

Ivaldi, M. and F. Verboven. 2005. 'Quantifying the Effects from Horizontal Mergers in European Competition Policy', *International Journal of Industrial Organization*669-691.

Kemp, Ron and Jarig van Sinderen. 2008. 'Measuring the economic effects of competition law enforcement', *De Economist*, 156, 4, 365-385.

Kovacic, W., R. C. Marshall, et al. 2007. Lessons for Competition Policy from the Vitamins Cartel. in: V. Ghosal and J. Stennek, *The Political Economy of Antitrust,*. Volume 282: 149-176.

Kumar, Vikram, Robert C. Marshall, Leslie M. Marx and Lily Samkharadze. 2011. *Cartel versus Merger*, http://capcp.psu.edu/papers/2011/cartelvsmerger.pdf.

Kwoka J and Lawrence J. White. 2004. *Antitrust Revolution: Economics, Competition, and Policy,* Oxford University Press.

La Porta, Rafael, Florencio Lopez de Silanes and Andrei Shleifer. 2008 'The Economic Consequences of Legal Origins', *Journal of Economic Literature*, 46, 285-332.

Langus, Gregor, Massimo Motta and Luca Aguzzoni. 2010 'The effect of EU antitrust investigations and fines on a firm's valuation', *http://ftp.zew.de/pub/zew-docs/veranstaltungen/Macci/SummerInstitute/Papers/SI_Motta.pdf*.

Lear 2011. *The ex-post evaluation of two merger decisions,* A Report prepared by Lear for the UK Competition Commission.

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

London Economics 2011 *The Impact of Competition Interventions on Compliance and Deterrence,* Office of Fair Trading, OFT1391.

Lübbers, T. 2009. 'Is Cartelisation Profi table? A Case Study of the Rhenish Westphalian Coal Syndicate, 1893-1913', *MPI Collective Goods Preprint, No. 2009/9*.

Lyons, Bruce 2009. '*Cases in European Competition Policy: The Economic Analysis*', Cambridge, Cambridge University Press.

Majoras, D. P. 2005. '*A Dose of our Own Medicine: Applying a Cost/Benefit Analysis to the FTC's Advocacy Program*', Available at: www.ftc.gov/speeches/majoras/050208currebttopics.pdf.

Malkiel, B. G. 2003. '*A random walk down Wall Street : the time-tested strategy for successful investing*', New York, London, W. W. Norton.

McKinsey Global Institute 2010a. '*Beyond austerity: A path to economic growth and renewal in Europe*'.

McKinsey Global Institute 2010b. '*From austerity to prosperity: Seven priorities for the long term in the United Kingdom*'.

Meyer, D. B. 1995. Natural and Quasi-Experiments in Economics. *Journal of Business and Economic Statistics, 13*, 151-161.

Meyer, Margaret A. and John Vickers. 1997 'Performance Comparisons and Dynamic Incentives', *The Journal of Political Economy*, 105, 3, 547-581.

Miller Nathan H 2009. 'Strategic Leniency and Cartel Enforcement.' *American Economic Review,* 99(3): 750–68

Mitschke, Andreas 2008. '*The influence of national competition policy on the international competitiveness of nations: a contribution to the debate on international competition rules*', Heidelberg, Physica Verlag.

Neven, D. J. 2006. 'Competition economics and antitrust in Europe', *Economic Policy*, 48, 741 - 791.

Neven, D. and H. Zengler 2008. 'Ex-post Evaluation of Enforcement: a Principal-Agent Perspective', *De Economist*, 156, 477-490.

Nevo, A. 2000. 'Mergers with Differentiated Products: the Case of the Ready-to-eat Cereal Industry', *RAND Journal of Economics*, 31, 3, 395-421.

Nickell, Stephen J. 1996. 'Competition and Corporate Performance', *Journal of Political Economy*, 104, 4, 724-746.

Nickell, Stephen J., Daphne Nicolitsas and Neil Dryden. 1997. 'What makes firms perform well?', *European Economic Review*, 41, 783-796.

NMa. 2005. '*Research into the Anticipation of merger Control*' at: http://www.nmanet.nl/Images/Rapport%20Research%20into%20the%20Anticipation%20of%20merger%20control_tcm16-86578.pdf).

OECD 2011 'The Impact of Trade Liberalisation on Jobs and Growth: Technical Note', *OECD Trade Policy Working Papers*, 107.

OECD 2002, *Roundtable on communication by competition authorities*, at: http://www.oecd.org/dataoecd/20/40/2492536.pdf

OFT 2011a. "Positive Impact 10/11 - *Consumer benefits from the OFT's work*.

OFT 2011b. *Competition and Growth*, Office of Fair Trading, OFT1390.

OFT 2011c. "*The Impact of Competition Interventions on Compliance and Deterrence*", OFT 1391

OFT 2011d. "*Commentary on retail mergers*", OFT 1305/CC2

OFT 2011e. *Evaluating the impact of the OFT's consumer enforcement case against Foxtons in respect of commission terms in letting agreements*, OFT1346eval

OFT 2011f. *Evaluating the impact of the OFT's 2001 abuse of dominance case against Napp Pharmaceuticals*, OFT1332

OFT 2010a. *Drivers of Compliance and Non Compliance with Competition Law'*, Office of Fair Trading, OFT1227.

OFT 2010b. '*Evaluation of OFT Competition Advocacy'*, Office of Fair Trading, OFT866.

OFT 2010c. '*Positive Impact 09/10 Consumer benefits from the OFT's work'*, Office of Fair Trading, OFT1251.

OFT 2007. *Final report of the Payment Systems Task Force*, OFT901

Ormosi, P. L. 2010. 'Merger Remedies versus Efficiency Defence: An analysis of merging parties' litigation strategy in EC merger cases', *CCP Working Paper Series*, 10-1.

Ormosi, Peter 2011. 'How Big is a Tip of the Iceberg? A Parsimonious Way to Estimate Cartel Detection Rate', *CCP Working Paper*, 11-6.

OXERA 2009. '*Quantifying antitrust damages Towards non-binding guidance for courts'*, Study prepared for the European Commission, http://www.oxera.com/cmsDocuments/Quantifying%20antitrust%20damages.pdf

Parikh, Ashok and Corneliu Stirbu. 2004 'Relationship between Trade Liberalisation, Economic Growth and Trade Balance: An Econometric Investigation', *HWWA Discussion Paper*, 282.

Peters, Craig. 2006. 'Evaluating the Performance of Merger Simulation: Evidence From the U.S. Airline Industry', *Journal of Law & Economics*, 49, 627–649.

Pinske, J. and M. E. Slade 2004. 'Mergers, Brand Competition, and the Price of a Pint', *European Economic Review*, 48, 617-643.

Prager, R. A. 1989. 'Using stock price data to measure the effects of regulation: the Interstate Commerce Act and the railroad industry', *RAND Journal of Economics*, 20, 2, 280-290.

PricewaterhouseCoopers 2005. '*Ex-post evaluation of mergers'*, A report prepared for the Office of Fair Trading, Department of Trade and Industry and the Competition Commission.

Przybyla, Marcin and Moreno Roma 2005. '*Does Product Market Competition Reduce Inflation? Evidence from EU Countries and Sectors'*, European Central Bank Working Paper Series, No.453.

Sabbatini, P. 2008. 'Assessing the impact of antitrust intervention by the Italian Competition Authority', *De Economist*, 156, 491-505.

Schmidt, Klaus M. 1997. 'Managerial Incentives and Product Market Competition', *Review of Economic Studies*, 64, 191-214.

Shaw, R. W. and P. Simpson 1986. 'The Persistence of Monopoly: An Investigation of the Effectiveness of the United KingdomMonopolies Commission', *The Journal of Industrial Economics*, 34, 4, 355-372.

Sidak, Gregory J. and David Teece 2009. 'Dynamic Competition in Antitrust Law', *Journal of Competition Law & Economics*, 5, 4, 581–631.

Sidak, S.G. 2003. 'The Failure of Good Intentions: The Worldcom Fraud and the Collapse of American Telecommunications After Deregulation', *Yale Journal on Regulation*, 20, 207-265.

Simpson, J. D. and D. Hosken. 1998. '*Are Retailing Mergers Anticompetitive? An Event Study Analysis'*, Federal Trade Commission, WP-216.

Simpson, J. S. 2008. 'Difference-in-Differences Analysis in Antitrust: A Cautionary Note', *Antitrust Law Journal*, 75, 623.

Sivadasan, Jagadeesh 2009. 'Barriers to Competition and Productivity: Evidence from India', *The B.E. Journal of Economic Analysis & Policy*, 9, 1.

Slade, Margaret 2008. The effects of vertical restraints: an evidence based approach. *The pros and cons of vertical restraints*. Swedish Competition Authority, Stockholm.

Symeonidis, George 2002. '*The effects of competition: cartel policy and the evolution of strategy and structure in British industry*', Cambridge, Mass., *MIT Press*.

Tenn, S. and J. M. Yun 2010. 'The success of divestitures in merger enforcement: Evidence from the J&J–Pfizer transaction', *International Journal of Industrial Organization*, Forthcoming.

Tussie, Diana and Carlos Aggio 2006. '*Economic and social impacts of trade liberalization'*, UNCTAD.

UK Competition Commission 2009. '*Stakeholder Research 2009: A research report prepared for The Competition Commission*' at: http://webarchive.nationalarchives.gov.uk/+/http://www.competition-commission.org.uk/our_role/analysis/evaluation_reports.htm

UK Competition Commission 2010. '*Understanding past merger remedies: report on case study research'*

USDoJ. 2012. *Antitrust Division's Performance Budget Congressional Submission for 2008-2011*, at: http://www.justice.gov/jmd/2012justification/pdf/fy12-atr-justification.pdfand

UK Competition Commission 2008. '*Evaluation of the Competition Commission's past cases*'.

USFTC 1999. '*A Study of the Commission's Divestiture Process'*

USFTC 2011. '*Performance and Accountability Report 2011 and earlier years*, at:

http://www.ftc.gov/opp/gpra/2011parreport.pdf

Verboven, F. and T. Van Dijk (2009) 'Cartel damages claims and the passing-on defence', *Journal of Industrial Economics*, 57, 3, 457-491(435).

Warren-Boulton, Frederick R. and Serdar Dalkir 2001. 'Staples and Office Depot: An Event-Probability Case Study', *Review of Industrial Organization*, 19, 4, 467.

Weinberg, Matthew 2007. '*The Price Effects of Horizontal Mergers: A Survey'*, CEPS Working Paper, No. 140.

Werden, G. J. 2000. 'Expert Report in United States v. Interstate Bakeries Corp. and Continental Baking Co.', *International Journal of the Economics of Business*, 7, 2, 139-148.

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.

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

Werden, Gregory J. and Luke M Froeb 1994. 'The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy', *Journal of Law, Economics and Organization*, 10, 2, 407-426.

Whinston, Michael and Scott Collins 1992. 'Entry and Competitive Structure in Deregulated Airline Markets: An Event Study Analysis of People Express', *The Rand Journal of Economics*, 23, 445-462.

World Economic Forum 2002. *'The Global Competitiveness Report 2001-2002'*, Oxford University Press.

Zheng, Shilin and Michael R. Ward 2011. 'The Effects of Market Liberalization and Privatization on Chinese Telecommunications', *China Economic Review*, (forthcoming).