PUBLIC

CT-2024-010

THE COMPETITION TRIBUNAL

IN THE MATTER OF the *Competition Act*, R.S.C. 1985, c. C-34;

AND IN THE MATTER OF certain conduct of Google Canada Corporation and Google LLC relating to the supply of online advertising technology services in Canada;

AND IN THE MATTER OF an Application by the Commissioner of Competition for one or more orders pursuant to section 79 of the *Competition Act*.

B E T W E E N:

COMMISSIONER OF COMPETITION

Applicant

and

GOOGLE CANADA CORPORATION AND GOOGLE LLC

Respondents

SUPPLEMENTARY MOTION RECORD

VOL 2 of 11

PUBLIC

August 22, 2025

DAVIES WARD PHILLIPS & VINEBERG LLP
155 Wellington Street West
Toronto ON M5V 3J7

Kent E. Thomson (LSO# 24264J)
Tel: 416.863.5566
Email: kentthomson@dwpv.com

Elisa K. Kearney (LSO# 49342T)
Tel: 416.367.7450
Email: ekearney@dwpv.com

Chantelle T.M. Cseh (LSO# 60620Q)
Tel: 416.367.7552
Email: ccseh@dwpv.com

Chanakya A. Sethi (LSO# 63492T)
Tel: 416.863.5516
Email: csethi@dwpv.com

Chenyang Li (LSO# 73249C)
Tel: 416.367.7623
Email: cli@dwpv.com

Tel: 416.863.0900
Fax: 416.863.0871

Lawyers for the Respondents, Google Canada Corporation and Google LLC

PUBLIC

TO: **ATTORNEY GENERAL OF CANADA**
Department of Justice Canada
Competition Bureau Legal Services
Place du Portage, Phase I
50 Victoria Street, 22nd Floor
Gatineau, QC K1A 0C9

Alexander M. Gay
Email: alexander.gay@cb-bc.gc.ca
Donald Houston
Email: donald.houston@cb-bc.gc.ca
John Syme
Email: john.syme@cb-bc.gc.ca
Derek Leschinsky
Email: derek.leschinsky@cb-bg.ca
Katherine Rydel
Email: katherine.rydel@cb-bc.gc.ca
Sanjay Kumbhare
Email: sanjay.kumbhare@cb-bc.gc.ca

Tel: 613.296.4470
Fax: 819.953.9267

Lawyers for the Applicant, the Commissioner of Competition

AND TO: **THE REGISTRAR OF THE COMPETITION TRIBUNAL**
Thomas D'Arcy McGee Building
90 Sparks Street, Suite 600
Ottawa, ON K1P 5B4

Email: Tribunal@ct-tc.gc.ca

PUBLIC

INDEX

Exhibit 2 to the Cross-Examination of Professor Tadelis

Federal Court Cour fédérale

Date: 20220112

Docket: T-40-18

Citation: 2021 FC 1435

Fredericton, New Brunswick, January 12, 2022

PRESENT: Madam Justice McDonald

BETWEEN:

PAID SEARCH ENGINE TOOLS, LLC

Plaintiff/
Defendant by Counterclaim

and

GOOGLE CANADA CORPORATION,
GOOGLE LLC AND ALPHABET INC.

Defendants/
Plaintiffs by Counterclaim

PUBLIC JUDGMENT AND REASONS

(Confidential Judgment and Reasons were issued December 17, 2021)

Table of Contents

I. Overview

[1] Advertising on the internet is now ubiquitous, but in 2000 it was an emerging field of opportunity for advertisers. This is a case about a patent designed to support advertisers in managing their bids for online advertising space on paid search engines.

[2] The plaintiff, Paid Search Engine Tools, LLC (PSET) owns Canadian Patent No. 2,415,167 (167 Patent) titled "Paid Search Engine Bid Management" for an invention described as "management of advertising expenses in online media". PSET alleges that the defendants have infringed certain claims of their patent and they seek damages.

[3] The defendants deny any infringement, and by counterclaim seek a declaration that the patent claims are invalid on various grounds including anticipation, obviousness, insufficiency and inutility.

[4] This trial in this matter was conducted via videoconference pursuant to the *Remote Trial Protocol* Order of March 8, 2021.

[5] For the reasons that follow, after construing the asserted claims of the 167 Patent, I have concluded that the asserted claims at 28, 44, 59 and 75 (and the dependent claims) as well as claim 9 are invalid on the grounds that they are not sufficiently disclosed. I have also concluded that the asserted claims at 1 and 17 (and the dependent claims) are invalid as they are anticipated or obvious.

[6] I have also concluded that the defendants have not infringed the 167 Patent.

[7] In light of my findings, I decline to address the issue of damages.

II. Background

A. The Parties

[8] PSET is incorporated in the State of Ohio with a registered office in Hamilton, Ontario. The 167 Patent has a Canada filing date of July 5, 2001, and a publication date of January 10, 2002. The patent was issued on March 21, 2017 and expired July 5, 2021. The 167 Patent claims priority to US provisional patent application 60/215,976 (976 Patent) filed on July 5, 2000.

[9] Google Canada Corporation is incorporated in the province of Nova Scotia. Alphabet Inc. is incorporated in the State of Delaware, and was created in 2015 through a corporate restructuring in which Google Inc. became a subsidiary of Alphabet. In 2017, Google Inc. was converted into Google LLC. Unless the context requires otherwise, I will refer to the defendants collectively as “Google”.

[10] Google owns the infrastructure that operates Google Ads and formerly AdWords. AdWords Select was launched in February 2002. AdWords Select was renamed to “AdWords” and became “Google Ads” in 2018. Unless the context requires otherwise, I will use the phrase “Google Ads” to refer to these products.

B. Patent Background

[11] In the early 2000s, with the increased use and popularity of the internet, advertisers were attempting to reach internet users by paying search engines to have their information or advertisements displayed in response to searches. In paid search engines, advertisers could pay to have their information appear in a certain position in the search engine's search result. This is described in the 167 Patent as follows:

> In a paid Internet search engine, content providers submit bids for each one or more keywords they desire to associate with their site. The paid search engine will respond to a user's request for sites with one or more keywords, by producing a list of links to those sites that have submitted bids on those keywords. The order in which links are identified is determined by the bid amounts provided by the sites – the site with the largest (cumulative) bid(s) for the keywords(s) identified by the user, appears first in the list of the sites presented to the user, followed by the site with the second largest (cumulative) bid(s) and so on.

[12] Cost-per-impression was one method by which advertisers were charged for this service, meaning advertisers paid an amount each time their advertisements appeared in response to a search query regardless of whether the searcher actually followed through on the advertisement and visited their website.

[13] In May 1999, the search engine GoTo.com (GoTo) introduced a new method for advertisers called sponsored search results. This new method was pay-per-click advertising, meaning that the advertiser was only charged for having their information appear in the search results when the searcher actually clicked on the advertiser's information (typically their URL website address). This was a more attractive model for advertisers as it meant they were only

paying for advertisements directed to those who actually showed some interest in their website. GoTo obtained a patent for its pay-per-click advertising model in July 2001.

[14] On GoTo, advertisers chose the amount of money (usually in cents) they were prepared to pay (the bid) to have their information displayed in response to the search of certain words or phrases (keywords). Advertisers who were vying for the top position in the search results had to be prepared to pay more for the chosen keywords in order to rank above competing advertisers.

[15] In February 2000, the named inventors of the 167 Patent – Juan Velez and Daren Murrer – met to discuss internet marketing. Mr. Velez was working in e-commerce and marketing. Mr. Murrer was involved in online businesses that used pay-per-click online advertising on GoTo.

[16] Mr. Murrer explained to Mr. Velez the challenges of managing as few as 10-12 keywords on GoTo, and how time consuming it was to have to manually change bids for each keyword in order to stay competitive. Mr. Murrer explained having to look at the keywords he bid on to determine his bid position and then having to change the bid to close the gap between his bid and the bids of other advertisers for that position. Mr. Velez told Mr. Murrer that he should be managing hundreds of keywords. According to the inventors, this discussion led to an all-day brainstorming session where they developed the idea that became the 167 Patent.

[17] Following this brainstorming session, Mr. Murrer and Mr. Velez contacted Mark Soper to work on computer code to implement their concept. Mr. Soper, who worked on the project over a weekend, developed code that could retrieve keyword search results from GoTo and arrange the results in a table. This table displayed the top 20 bids for a keyword, and showed the number of times the keyword was searched. This allowed advertisers to see if there was an "optimization opportunity", namely, a gap between their bid and the next high advertisers bid, which could then be collapsed (a bid collapse). For example, if Bidder A bids $0.05 and the next highest bidder, Bidder B, bids $0.08, the "optimization opportunity" is for Bidder B to reduce their bid to $0.06 while still remaining in the top position relative to other bidders in the search results.

[18] Mr. Velez and Mr. Murrer filed a provisional patent application on July 5, 2000 (referred to as the 976 Patent). They offered a commercial product known as the Keyword Bid Optimizer (KBO or KBO tool) which they operated from their website: PaidSearchEngineTool.com. The KBO tool worked by collecting information from GoTo (each night) on PSET customers' keywords and the bids on those keywords. These search results were then displayed in a report for PSET customers with columns showing: the keyword(s), the number of times it was searched (views), the top 20 bids for the keyword, and where the advertiser's bid (in red) ranked relative to others. This report shown below from the 976 Patent, is also Figure 4 of the 167 Patent:

Keyword report for www.gardens-alive.com

A yellow row indicates that www.gardens-alive.com was NOT found in the first 20 results for that keyword.

A Red $ indicates that www.gardens-alive.com was found in that position for that keyword.

Keyword	Views	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
beneficial insect	69	$0.06	$0.05	$0.05	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
cover crop	43	$0.03	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
earth friendly product	19	$0.05	$0.03	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
environment	6758	$0.65	$0.64	$0.62	$0.56	$0.55	$0.52	$0.47	$0.45	$0.43	$0.30	$0.26	$0.24	$0.22	$0.20	$0.20	$0.20	$0.16	$0.15	$0.14	$0.14
fertilizer	1518	$0.49	$0.48	$0.46	$0.45	$0.42	$0.37	$0.36	$0.35	$0.27	$0.25	$0.24	$0.23	$0.20	$0.18	$0.17	$0.15	$0.14	$0.10	$0.10	$0.07
flower gardening	473	$0.29	$0.29	$0.28	$0.27	$0.25	$0.23	$0.21	$0.17	$0.15	$0.15	$0.12	$0.08	$0.07	$0.06	$0.05	$0.05	$0.05	$0.02	$0.02	$0.02
garden	11233	$0.86	$0.85	$0.84	$0.82	$0.80	$0.77	$0.73	$0.68	$0.67	$0.66	$0.60	$0.53	$0.50	$0.39	$0.36	$0.35	$0.35	$0.35	$0.31	$0.30
garden alive	290	$0.02	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
garden pest	438	$0.17	$0.16	$0.15	$0.15	$0.11	$0.10	$0.05	$0.02	$0.01	$0.01	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
garden tip	131	$0.15	$0.14	$0.14	$0.13	$0.08	$0.05	$0.01	$0.01	$0.01	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00

[19] From this report, the advertiser could see that they had the top position for the keyword "cover crop" with their bid of $0.03. However, they would also see an optimization opportunity to reduce their bid from $0.03 to $0.02 while maintaining this top position relative to the other bidders.

[20] According to the inventors, PSET started selling the KBO tool commercially and at one time had over 500 subscribers. During the three years that the KBO tool was sold, PSET reported revenues of: $56,000 in 2001, $43,000 in 2002 and $21,000 in 2003.

[21] In May 2001, GoTo and PSET entered into a 6-month agreement in which GoTo agreed to feed search results directly to PSET in exchange for PSET providing the URLs and historical account information of their customers. This agreement was not renewed.

[22] By June 2002, GoTo, who was now operating as Overture, introduced optimization tools directly on their website. According to PSET, these tools provided the same service and information as provided through their KBO product. According to Mr. Murrer, this move by Overture effectively put PSET out of business and by 2003 PSET stopped selling the KBO tool. PSET claimed that Overture used PSET's method and techniques. In October 2003, Overture was purchased by Yahoo.

[23] Mr. Murrer confirmed that the KBO tool did not automatically adjust bids, did not show the number of clicks on an ad, and did not display information about clickthrough rates (CTR, or how often an ad was displayed and clicked). The KBO tool was not itself a search engine and it did not run an auction to determine what ads would be displayed.

[24] Mr. Murrer claims that the 167 Patent discloses more than the KBO tool. The details of the 167 Canadian patent history was introduced into evidence (Exhibit 2).

[25] Mr. Murrer had limited knowledge of the 167 Patent proceedings in Canada but confirmed that PSET had two patent applications in the US that claimed priority to the 976 Patent – the 450 Patent and the 912 Patent. He also confirmed that PSET's patent infringement claims against Google (and Yahoo and Microsoft) in the US regarding these patents were dismissed (Exhibit 15).

[26] PSET's patent application to the European Patent Office, which also relied on the 976 Patent, was refused (Exhibit 24). On January 18, 2013 and September 25, 2014, the US Patent

and Trademark Office cancelled the claims pertaining to the 450 and 912 Patents (Exhibits 26 and 27).

III. Evidence

[27] The following is a brief summary of the witness evidence. I include this summary to provide an overall factual and contextual framework. Where relevant, I also note my general observations as to the reliability of the witnesses' evidence. The particulars of some of the witness evidence will be addressed in more detail in the analysis of the issue to which it relates.

A. PSET's Fact Witnesses

[28] **Juan Velez** is a co-inventor on the 167 Patent. He has a chemical engineering degree, and an MBA. He does not have a computer science background or a background in auctions. Mr. Velez gave evidence about the meeting he had with Daren Murrer in February 2000, when they developed the concept for the patent. He testified that he and Mr. Murrer also worked with Cotty England and Mark Soper to develop the patent, and consulted his professor, Jon Keel.

[29] Mr. Velez had limited knowledge of how the patent could work in practice as his involvement was focused on the marketing of the patent.

[30] **Daren Murrer** is the co-inventor on the 167 Patent and a founder of PSET. Mr. Murrer explained that PSET was also doing business as Earth Internet Services or Earth LLC in Ohio, before becoming its own LLC. Mr. Murrer has no formal education in computer science, online

marketing or auction design. He has experience with online sales through his other businesses where he created a website to sell products.

[31] He also testified about the brainstorming session with Mr. Velez in February 2000. He explained that the code created by Mr. Soper allowed for automatic bidding, but he was not comfortable implementing that step immediately due to liability concerns. He confirmed that when they launched the KBO tool, the automatic bid feature was not included. He explained that the KBO tool was used on various search engines including GoTo, Kanoodle, and Ah-ha.

[32] Mr. Murrer provided helpful contextual background, however, he claimed to lack knowledge on details of the patent filing history in Canada and PSET's patent litigation in the US.

[33] **Mark Soper** is a self-taught software programmer with no formal education in software programming or computer science. He testified about the code he wrote for PSET. He confirmed that this code did not identify bid gaps and did not do calculations, but it was capable of doing automatic bidding. He confirmed that ad quality and ad relevance were not assessed by the code he developed. Mr. Soper was a credible and straightforward witness.

[34] **Jon Keel** taught Mr. Velez in a course on internet marketing. He did some work for PSET and became a part owner in 2000. He described the process of managing bids prior to the use of the KBO tool. He was questioned about an email he sent to Planet Ocean Communications (Planet Ocean) where he provided them with the KBO tool (Exhibit 46).

[35] Patent Agent **Shauna Paul** testified about the prosecution of the 167 Patent, including the timing of the registration of the patent in Canada. She testified that she was taking instructions from PSET's US patent attorneys.

[36] **Daniel Boberg** worked at GoTo from 1999-2009. He described GoTo as the first sponsored or paid listing search engine. Mr. Boberg explained the bidding process on GoTo.com, GoTo's Direct Traffic Centre and the impact of PSET's KBO tool on GoTo.

B. PSET's Expert Witnesses

[37] **Jessie Stricchiola** was qualified on consent as an expert in search engine marketing, search engine optimization, paid search, digital marketing and web traffic analysis. She does not have any formal education in computer science, economics, or computer coding.

[38] Ms. Stricchiola has worked in the field of search engine optimization since 1998. Her company, Alchemist Media, works in strategy, implementation and management of various digital marketing efforts by clients involving search engine optimization, paid search engine advertising, website analytics and related areas. She is the co-author of *The Art of SEO: Mastering Search Engine Optimization.* Ms. Stricchiola is a cofounder of SEMPO, the Search Engine Marketing Professional Organization, and served on the board of directors for two years.

[39] Ms. Stricchiola prepared the following reports, marked as Exhibits 31, 32, and 33:

Report on Infringement, dated January 29, 2021 (Stricchiola First Report)

Report on Validity, dated April 7, 2021 (Stricchiola Second Report)

Responding Report, dated May 7, 2021 (Stricchiola Third Report).

[40] Overall, her evidence was helpful. However, her evidence lacked particularity with respect to claims construction and infringement. I would describe it as generalized. Further, her reliance on selected Google marketing material over Google technical information impacts the objectiveness of her opinion on infringement.

[41] **Dr. Stephen Becker** is an economist with expertise in corporate financial analysis and the evaluation of economic damages, including intellectual property damages and patent infringement damages. Dr. Becker provided two reports marked as Exhibits 44 and 45 as follows:

Report on Reasonable Royalty, dated January 29, 2021

Report on Apportionment and Non-Infringing Alternatives, dated May 7, 2021.

[42] **Dr. Ernan Haruvy** was qualified on consent as an economist with expertise in auction theory, auction design, procurement, online advertising, and quantitative and qualitative methods for data analysis, including financial analysis and predictive analysis. Dr. Haruvy has a PhD in economics from the University of Texas at Austin (1999) and completed a post-doctorate fellowship at Harvard Business School (2000-2001). He is the Cleghorn Faculty Scholar (Full) Professor of Marketing at McGill University. He is also affiliated with the University of Texas at Dallas. Dr. Haruvy provided three reports marked as Exhibits 54, 56, and 58, as follows:

Report on Revenues and Profits, dated January 29, 2021 (Haruvy First Report)

Report on Validity, dated April 7, 2021 (Haruvy Second Report)

Responding Report on Remedies, dated May 7, 2021 (Haruvy Third Report).

C. Google's Fact Witnesses

[43] **Dr. Eric Veach** holds a PhD in Computer Science from Stanford University. In 2000, he joined Google and worked with the online advertising team. Dr. Veach explained the ad system used by Google in the early 2000s and described the development of Google's AdWords in July 2000. According to Dr. Veach, by November 2000, Google began to consider ways to improve AdWords. In May 2001, Google began building the program that became AdWords Select, which was launched in February 2002. AdWords Select focused on pricing and ranking, and incorporated ad quality considerations. According to Dr. Veach, Google used a second price auction. He also explained Google's Smart Ad Selection System (SmartASS) which uses machine-learning technology.

[44] **Gerald Dischler** is the VP and General Manager of Ads at Google. He previously worked with Google's ads team on the back-end technology of search advertising. Mr. Dischler now leads product management, engineering and design for the advertising team at Google. Mr. Dischler testified about Google's products and Google's ad system, as well as Google's ad quality. Mr. Dischler's evidence at trial differed from his discovery evidence on some material points. Where there were discrepancies, I prefer his discovery evidence or the documentary evidence.

[45] **Salar Kamangar** began working at Google in 1999. He worked on ads at Google in 2000, including Google's in-house ad product called Premium Sponsorships. Mr. Kamangar gave evidence about Google AdWords and AdWords Select from a high-level design standpoint.

[46] **Thomas Iljic** joined Google in 2015 and is currently the product manager for the text and shopping auction. In this capacity, Mr. Iljic deals with ads auctions and shopping ads auctions for Google.com. Mr. Iljic provided evidence on Google Ads from the perspective of an advertiser. He explained how ads are selected, ranked and priced.

[47] **Joshua Moser** leads Google's search bidding product team. Mr. Moser has been with Google since 2013. Mr. Moser explained Google Ads Help Centre and Google Ads bidding system.

[48] **Dr. Eric Schmidt** is the former Executive Chairman and CEO of Google Inc. He gave evidence on the Google advertising system model beginning in 2001 and he explained how the auction-based system evolved. Dr. Schmidt also explained the impact of Google Ads on Google's revenues. He confirmed that in the early 2000s, Google implemented financial restrictions. While PSET placed significant emphasis on Mr. Schmidt's statements made during an interview, I accord these statements little evidentiary value as I view them more as public relations statements rather than statements of the inner workings of Google.

[49] **Desmond Keane** is the senior engineering director responsible for the site reliability engineering team at YouTube, which is a part of Google. In 2004, Mr. Keane started with Google as a systems administrator. In 2007, Mr. Keane managed a team of systems administrators and software engineers responsible for internal business applications and core internal infrastructure systems. In 2015, he was the engineering director for all of the Ads site reliability engineering teams. Mr. Keane gave evidence on Google's infrastructure and confirmed that AdWords is accessible in Canada. He provided helpful detail on Exhibit 111, which provides a high-level overview of Google's infrastructure – from the boundary of things that run in Google data centres (including the Google Ads system) to the end user interacting with Google's home page from their phone or computer.

[50] **Jessie Brader** was a part owner of Planet Ocean – a company that published a digital newsletter regarding search engine information in the early 2000s. Ms. Brader gave testimony on Planet Ocean's operations as well as an email exchange with Jon Keel about the KBO tool (Exhibit 46). She testified that this email was typical of the type of communications received by Planet Ocean in the early 2000s.

[51] **Kevin Lee** started Did-It in 1996. He explained that Did-It was an organic search technology company that evolved into a company that assisted marketers in monitoring their positions in search engines. He explained pay-per-click advertising in the late 1990s through the early 2000s. Did-It used GoTo as a search engine as noted in Exhibit 112. According to Mr. Lee, there were a number of companies in the early 2000s focusing on the bid search and

management business. His evidence was that this was a busy area of development with a number of companies building similar products at the same time.

[52] **Neela Morrison** is senior corporate counsel at Google LLC. She gave evidence on the corporate structure of Google LLC, Alphabet Inc., and Google Canada.

[53] **James Maccoun** is patent counsel at Google who explained Google's patent licensing agreements.

[54] **Amrit Nandan** is Director of Finance with the Ads Business Unit at Google. He provided financial information regarding Google's product areas.

[55] **Buck Farmer** is a senior finance business intelligence analyst with Google who provided evidence on Google's financial allocation system.

D. Google's Expert Witnesses

[56] **Dr. David Parkes** is a professor of computer science at Harvard University. Dr. Parkes has a Master's Degree in Engineering and Computing Science from Oxford University and a PhD in Computer and Information Science from the University of Pennsylvania. Dr. Parkes was qualified as an expert in relation to computer science, auctions, market design, search engines, machine learning, artificial intelligence, e-commerce and internet advertising, including search engine marketing. He has provided two reports marked as Exhibits 122 and 123 as follows:

Report on Patent Claim Construction and Validity, dated January 29, 2021 (Parkes First Report)

Responding Report, dated April 7, 2021 (Parkes Second Report).

[57] On claims construction, Dr. Parkes occasionally took a literal approach to the claims language rather than a purposive approach. Likewise, his skilled person possesses significant academic credentials which is not reflective of those who would be using the online advertising products in the early 2000s. However, subject to these comments, I generally preferred Dr. Parkes' evidence and his claims construction.

[58] **Michael Grehan** was qualified on consent as an expert in relation to search engines, search engine marketing, and digital marketing. Mr. Grehan authored the book *Search Engine Marketing: The Essential Practice Guide* and was the Chair of SEMPO. He authored a report dated April 7, 2021. However, as I have concerns about the objectiveness and independence of Mr. Grehan's opinions, I do not accord his opinion much weight.

[59] **Steven Tadelis** was qualified as an expert economist with expertise in the economics of institutions, economic analysis for business decisions, including those related to digital advertising and marketing strategies, contract theory, strategic sourcing and pricing, online auctions and pricing structures. Dr. Tadelis provided one report dated April 7, 2021. I have concerns about the objectivity of his report, as it was revealed in his oral evidence that a consulting firm – Analysis Group – was heavily involved in the preparation of his report. He was also selective in the Google information he relied upon.

[60] **Christopher Bakewell** was qualified on consent as an expert on intellectual property valuation and licensing related issues, including economic issues related to the determination of a reasonable royalty. Mr. Bakewell provided a report dated April 7, 2021. I have concluded that Mr. Bakewell's approach to the royalty analysis was too narrow and relied upon unsupported assumptions. If I had assessed damages, I would not have afforded his report much weight.

[61] **Errol Soriano** was qualified on consent as an expert in assessing, analyzing, and quantifying financial damages and profits; the evaluation of business interests; and forensic accounting, including in the context of intellectual property disputes like patent infringement disputes. Mr. Soriano provided one report dated April 7, 2021.

IV. The 167 Patent

[62] The patent describes the advertisers' burden of managing keyword bids on paid search engines as paid search engines became increasingly popular. As bidding on keywords increased, bid rankings changed more frequently, and managing positions became an inefficient and time-consuming task. The patent states:

> To foster competition, paid search engines have provided facilities for bidders to monitor certain statistics, such as a daily count of "hits" on particular keywords, and reports of current bids on a given single keyword. However, paid search engines have not, to date, made such competitive information readily accessible. For example, a bidder can only view current bid positions of one keyword at a time, and has no mechanism for quickly identifying large gaps in bid amounts indicative of an opportunity for bid optimization. For a content provider managing tens or hundreds of keyword bids, the burden of evaluating each keyword individually can be substantial.

[63] The patent sought to address this inefficiency. First, the patent describes a process for accumulating customers' account and keyword information (FIG. 1) and generating a keyword report (FIG. 2):

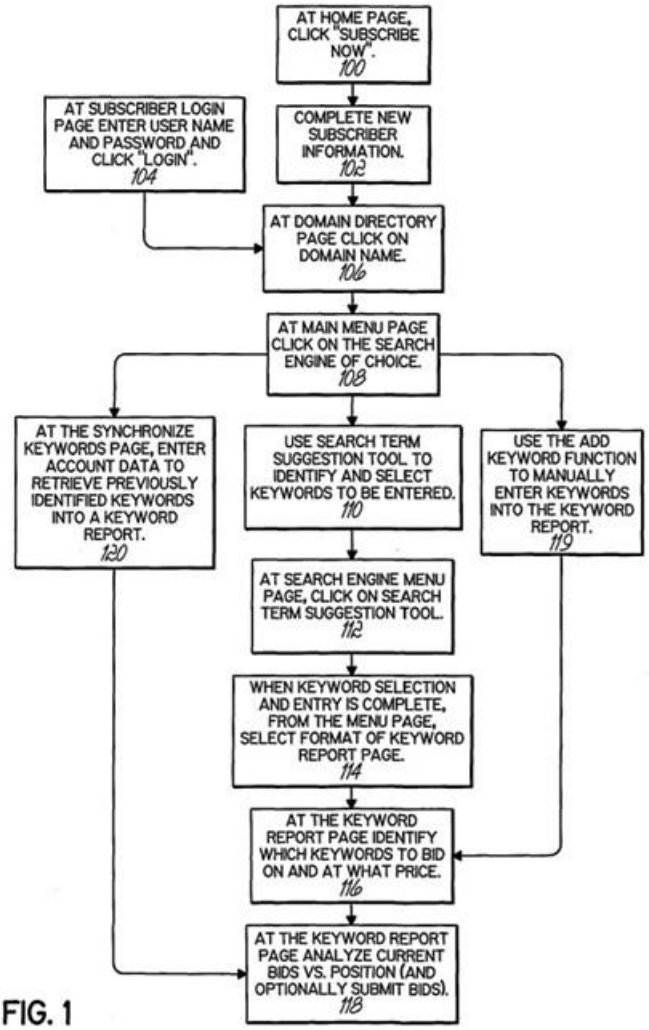


FIG. 1

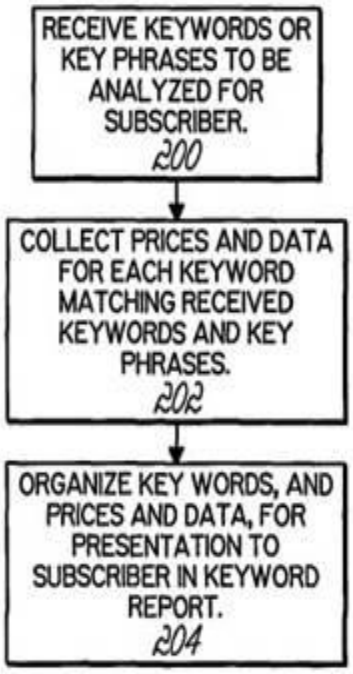


FIG. 2

[64] The patent also describes a method for identifying opportunities for customers to optimize their keyword bids:

> A method and apparatus for improving efficiencies in the current paid search engine keyword bidding market and optimizing use of use of such engines. The system accumulates bid amounts for a plurality of target keywords at one or more paid search engines and presents bid amounts to a user enabling the user to evaluate

and optimize bids on those keywords. Bid amounts of keywords of interest are highlighted (302). Differential bids can be identified to optimize bids. Keyword bid changes are monitored to identify changes of interest to a potential bidder (306).

[65] This method of keyword bid monitoring is illustrated below in FIG. 3:

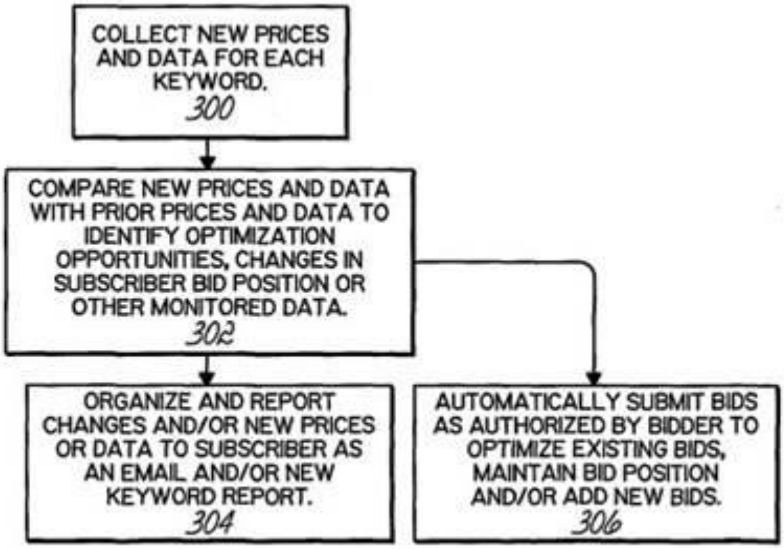


FIG. 3

V. Claims Construction

A. Legal Principles

[66] The Supreme Court in *Whirlpool Corp v Camco Inc*, 2000 SCC 67 [*Whirlpool*] states at paragraph 45 that "[t]he key to purposive construction is therefore the identification by the court, with the assistance of the skilled reader, of the particular words or phrases in the claims that describe what the inventor considered to be the 'essential' elements of his invention."

[67] The Supreme Court also stated in *Free World Trust v Électro Santé Inc*, 2000 SCC 66 [*Free World Trust*] at paragraph 44 that "[t]he courts have traditionally protected a patentee from the effects of excessive literalism."

[68] The relevant date for the purposes of claims construction is the publication date (*Whirlpool* at para 55). Here the publication date is January 10, 2002.

[69] The principles of claim construction are well summarized by Justice Fothergill in *dTechs EPM Ltd v British Columbia Hydro and Power Authority*, 2021 FC 190 at para 113 [*dTechs*] as folllows:

> The canons of claim construction are found in the Supreme Court of Canada's decisions in *Consolboard Inc v MacMillan Bloedel (Saskatchewan) Limited*, [1981] 1 SCR 504 at 520, *Whirlpool* at paragraphs 49 to 55, and *Free World Trust v Électro Santé Inc*, 2000 SCC 66 [*Free World Trust*] at paragraphs 44 to 54. They are the following:
>
> (a) the words of the claims must be read in an informed and purposive way with a mind willing to understand, viewed through the eyes of the person of ordinary skill in the art [PSA] as of the date of publication having regard to the PSA's common general knowledge;
>
> (b) the *Patent Act* promotes adherence to the language of the claims. This allows the claims to be read in the manner the inventor is presumed to have intended, and in a way that is sympathetic to accomplishing the inventor's purpose, which promotes both fairness and predictability;
>
> (c) the whole of the patent's specification should be considered to ascertain the nature of the invention, and the claim construction must be neither benevolent nor harsh, but should instead be reasonable and fair to both the patentee and the public. The focus of the analysis is on the claims; specifications will be relevant only where there is ambiguity in the claims (*AstraZeneca Canada Inc v Apotex Inc*, 2017 SCC 36 at para 31); and

(d) claim construction must be the same for the purpose of validity and for the purpose of infringement.

B. Person of Skill in the Art (PSA)

[70] As noted in *Whirlpool* at paragraph 53:

> … the patent specification is not addressed to grammarians, etymologists or to the public generally, but to skilled individuals sufficiently versed in the art to which the patent relates to enable them on a technical level to appreciate the nature and description of the invention [citations omitted].

[71] The PSA has "common knowledge in the trade" (*Whirlpool* at para 70) and "is thought to be reasonably diligent in keeping up with advances in the field to which the patent relates" (*Whirlpool* at para 74).

[72] The parties disagree on the knowledge and skills of the PSA.

[73] Google's expert Dr. Parkes says that the skilled person has "a university degree in computer science or a degree in a related field with approximately 2-5 years of experience in internet technologies generally", as well as "at least 2 years of experience with online marketing, and knowledge of search engine marketing". His skilled person would also know market design, including auction design, through a university degree or 3-6 years working on online market systems (Parkes First Report at para 24).

[74] Although both PSET's experts Ms. Stricchiola and Dr. Haruvy claim to rely upon Dr. Parkes' skilled person, they nonetheless provided their own opinions on the attributes of the skilled person which differ from Dr. Parkes on key aspects.

[75] Ms. Stricchiola says the skilled person requires "a general understanding of the structure and function of the internet as a whole" and would be familiar with "web browsers, websites, webpages, website URLs, domain names, clicks, users, and hyperlinks, and/or links". She says the skilled person would be familiar with "paid search advertising within search engines", and would be familiar with "target keywords, bidding, CPC or PPC (cost per click and pay per click, respectively), clickthrough rates (CTR), impressions, landing pages, campaign management and optimization, budget, maximum bids, rankings, results, positions, competitor keywords, user demographics, search and query volume, keyword research, and related topics" (Stricchiola First Report at para 49-52).

[76] In her opinion, the skilled person would need to have access to a web developer and/or a computer programmer/coder.

[77] Dr. Haruvy says that the skilled person would have experience and familiarity with the internet search industry and would have experience or familiarity with basic concepts in bidding. He agrees that the skilled person would include or have access to a coder. Dr. Haruvy says that the skilled person would be "a manager or executive at a portal, ISP/IAP, search engine, or advertiser tasked with making strategic pricing decisions in respect to paid search engine advertising". Further "[t]he familiarity required with (1) internet search concepts, (2)

bidding concepts, and (3) access to a coder, means that the manager described here is more likely to come from a high tech business background which would often involve college education with related exposure" (Haruvy Second Report at para 61-64).

[78] Dr. Haruvy states that there is nothing in the 167 Patent that requires much familiarity with specialized auction formats (Haruvy Second Report at para 65).

[79] I agree with Ms. Stricchiola that the skilled person would have a general understanding of the structure and function of the internet as a whole and an understanding of search engines and general search engine functionality, including from both a user and advertiser perspective. I agree that this understanding would include how to perform a search, what search results are, the difference between paid and non-paid search results, and how and where advertisers can appear within search results.

[80] I do not agree with Dr. Parkes that a degree in computer science would be necessary. I note, although not determinative, that the named inventors did not have such degrees. Furthermore, Mr. Soper – the coder for the 167 Patent – does not have a formal education in software programming.

[81] The issue that arose in cross-examination was whether the PSA personally had computer coding skills or had access to someone with those skills. Although this is an area where the experts appear to diverge, they all agree that some coding is necessary to put the 167 Patent into practice. Accordingly, coding knowledge or access to that knowledge was necessary for the

PSA. However, as noted by Dr. Haruvy, the relevant computer coding skills in the early 2000s were comparatively basic. Therefore, the skilled person would be someone with at least 2 years of coding experience or access to someone with at least 2 years of coding experience. The skilled person may therefore be a team.

[82] In my view, Dr. Parkes overstates the skilled person attributes. I do not agree that the skilled person requires a university degree as that was not reflective of the skill set of those working with paid search engines in the early 2000s. I do not reach this conclusion based upon the profile of the co-inventors, but rather based upon the state of the relevant technological development at the relevant time.

C. Common General Knowledge (CGK)

[83] In *Apotex Inc v Sanofi-Synthelabo Canada Inc,* 2008 SCC 61 [*Sanofi*] the Court noted that "[c]ommon general knowledge means knowledge generally known by persons skilled in the relevant art at the relevant time" (at para 37). The relevant time for assessing CGK is the publication date (*Eli Lilly Canada Inc v Mylan Pharmaceuticals ULC,* 2020 FC 816 at para 187 [*Eli Lilly*]).

[84] The publication date of the 167 Patent is January 10, 2002.

[85] Although the experts are not in complete agreement on what would make up the CGK, I would characterize their disagreement as more in relation to the depth of knowledge the PSA would possess, rather than the areas of knowledge.

[86] Based upon the above, in my view, the PSA would have knowledge of the following concepts:

a) Advertising on the internet, including banner ads on webpages.

b) How paid search engines like GoTo functioned.

c) Online pricing models for advertising including: cost-per-mille (CPM – a cost per thousand impressions); cost-per-click (CPC – a cost per clickthrough); and cost-per-action (CPA – where the advertiser is charged each time a user takes a predefined action).

d) Web browsers such as Netscape's Navigator and Microsoft's Internet Explorer.

e) Web scraping by fetching webpages and parsing out particular information from that web page.

f) The major search engines and web directories, including Google, Yahoo!, AltaVista, Lycos and AskJeeves.

g) Web crawling, which involved the parsing of Internet webpages to produce an index of webpages such that each webpage could be associated with search terms.

h) How to use keywords (individual words, word combinations, or short phrases) with a search engine, and understanding the association of keywords with webpage meta tags.

i) Search engine optimization (SEO), being the process of optimizing a webpage to achieve better results (higher rankings) on search engines.

j) Online auctions, including various types of auctions (like eBay), types of bidding, and auction participant behaviors including proxy bidding and the use of automatic bidding agents.

[87] On the issue of auctions, I find that Dr. Parkes' CGK is broader in scope than the field of the invention. The advanced technical aspects of auction theory and market design fields that Dr. Parkes describes would not be part of the relevant common knowledge of the skilled person at the relevant time.

[88] Dr. Haruvy also disagrees with Dr. Parkes' CGK regarding auctions. Dr. Haruvy states:

> I agree that the Parkes Skilled Person would know, as part of their common general knowledge, that bidding agents on eBay and similar sites had the potential for "bid-snipers" […]. The Parkes Skilled Person, however, would also know that unlike the single-item auctions on eBay, "winning" in the context of a paid search engine is fundamentally different from winning on eBay. Whereas "winning" on eBay or Yahoo! Auctions means taking home the prized item, winning on a paid search engine using CPC pricing is measured by a different metric - a user clickthrough. Being listed first in search engine results page is not a "win" if users do not actually click through on the displayed advertisement. Similarly, not obtaining the first position (which would clearly be a "loss" in an eBay auction), may still be a "win" in a paid search engine if it nevertheless results in a user click through (Haruvy Second Report at para 82).

[89] I conclude that a detailed understanding of auction theory is not required to understand the technical area of the 167 Patent. The PSA would have the following general knowledge of auction concepts:

- The difference between an “open” auction where all participants (bidders) know the bids of other parties, and a “closed” or sealed auction where bidders do not know the bids of others;

- First price auctions (like GoTo), where the winner pays their bid amount. In a first price closed auction, this means the winner is the highest bidder; and,

- Second price auctions (sometimes called Vickrey auctions) are auctions where the winner pays the amount of the second-highest bid. A second price auction allows bidders to bid their true value knowing that they will not be paying more than necessary.

D. Claim Terms Needing Construction

[90] In their statement of claim, PSET is asserting 6 independent claims (in bold) and the following dependent claims:

a. **Claim 1** and dependent claims 2-5, 7-9
b. **Claim 17** and dependent claims 18, 19, 22, 24, 25, 27
c. **Claim 28** and dependent claims 29-34, 36, 37, 39-43
d. **Claim 44** and dependent claims 45, 46, 55-58
e. **Claim 59** and dependent claims 60-65, 67, 70, 72-74
f. **Claim 75** and dependent claims 76-81, 83, 86, 88 and 89.

[91] Dr. Parkes construed all the claims and provided a claims chart.

[92] Ms. Stricchiola and Dr. Haruvy did not construe each asserted claim in their reports, stating that it was not necessary to do so where the claims would be understood by the skilled person according to their ordinary meaning. Ms. Stricchiola and Dr. Haruvy did not construe the following claims: 4, 5, 7, 18, 22, 24, 25, 27, 31, 33, 34, 36, 37, 39, 40, 41, 42, 43, 45, 46, 55, 56, 57, 58, 60, 61, 62, 63, 64, 65, 67, 70, 72, 73, 74, 76, 77, 78, 79, 80, 81, 83, 86, 88, and 89.

[93] Overall, I prefer the claims construction of Dr. Parkes to that of Ms. Stricchiola and Dr. Haruvy. Dr. Parkes provided a fulsome and objective approach to construction and he individually construed each claim, even when he acknowledged that is was difficult to do so.

[94] In addition to not addressing each of the asserted claims, I would describe Ms. Stricchiola's claims construction as not fully comprehensive and often lacking in detail and substance. Her approach was more aspirational rather than practical. Dr. Haruvy's construction was less comprehensive, and was more of a critique of Dr. Parkes' construction.

Claim 1 (paragraph lettering added)

1. A method of managing a bidder's bid to pay for the use of a bidder-supplied Internet link, the bid identifying a target keyword and an amount the bidder is willing to pay to a search engine when the search engine presents the bidder-supplied link as one of a plurality of search results in response to a search using the target keyword and the link is then used by a recipient of the search results, comprising:

   (a) retrieving a web page used to supply and gather information on the bidder's bids made to the search engine, the web page presenting one or more keywords and information on the bidder's bids relating to those keywords;

   (b) entering into the web page a target keyword and the amount the bidder is willing to pay when the bidder-supplied Internet link is presented in a list with other search results in response to a search using the target keyword and the bidder-supplied Internet link is used by a recipient of the search results;

(c) authorizing an automatic reduction of the entered bid amount to a lower amount that will not change the position of the bidder-supplied Internet link relative to the other search results presented in the list of search results; and

(d) retrieving a web page reporting at least one previously authorized automatic reduction in the amount bid and paid for a use of the bidder-supplied Internet link.

[95] Claim 1 is a method claim that outlines various steps.

[96] All experts agree that **bidder** is a reference to the advertiser or their representative who provided the advertisement and who will be charged by the paid search engine when users click on its advertisement.

[97] The **bidder's bid** is the amount of money that the advertiser is willing to pay for a click on their hyperlink and/or to maintain a specific ranking or position for that hyperlink.

[98] I construe the step in 1(a) **retrieving a web page used to supply and gather information on the bidder's bids made to the search engine** to mean that the bidder obtains a webpage that presents the information on the bidder's bids and allows the bidder to input information. The phrase **the web page presenting one or more keywords and information on the bidder's bids relating to those keywords** refers to the webpage displaying the bidder's keywords, the current bids on those keywords and information, such as the URL.

[99] For step 1(b), I construe **entering into the web page a target keyword** as the step when the bidder inputs a keyword that is the target of its advertising and the **amount the bidder is willing to pay** is the entered bid amount. The **bidder-supplied internet link is used**

by a recipient of the search results is the advertiser's (bidder)'s ad containing a hyperlink to their website.

[100] The experts disagree on the construction of 1(c) which states **authorizing an automatic reduction of the entered bid amount to a lower amount that will not change the position of the bidder-supplied Internet link relative to the other search results presented in the list of search results.**

[101] Dr. Haruvy states that 1(c) would be understood to incorporate: "(1) a generalized second price mechanism; (2) automation of a possible bid reduction; and (3) optimization" (Haruvy Second Report at para 111).

[102] He goes on to define a generalized second price auction (GSP auction):

> A generalized second-price mechanism ("GSP") is not explicitly referred to by that name in the patent but is a mechanism nevertheless clearly described in the patent. It refers to an auction mechanism for multiple items, where in the present case the multiple items are the different ranks/slots in the search results. Each bidder places a bid. The highest bidder gets the first slot, the second-highest, the second slot and so on. The highest bidder pays for a click through at the price bid by the second-highest bidder plus some minimum bid increment (i.e., $0.01), the second-highest pays for a click through the price bid by the third-highest plus the minimum bid increment, and so on (Haruvy Second Report at para 112).

[103] Dr. Haruvy also highlighted that second price auctions and GSP auctions "have little in common in terms of auction properties" (Haruvy Second Report at para 94). Ms. Stricchiola also construes claim 1 to incorporate a GSP auction.

[104] I agree that 1(c) makes reference to the automation of a bid reduction. However, I do not accept a construction of claim 1(c) that introduces an auction mechanism as it does not disclose or support a construction that it is a stand-alone or self-supporting method for an auction. Claim 1(c) states that it is designed to function within a paid search engine.

[105] I adopt Dr. Parkes' construction of 1(c) to mean a bid collapse. The words **will not change the position of the bidder-supplied Internet link relative to the other search results presented in the list of search results** means that the amount submitted to the paid search engine will not change the position or rank of the bidders link in the search results list relative to the other search results even though the bidder's bid amount has been lowered.

[106] For step 1(d) Ms. Stricchiola construes **amount bid and paid** as being different amounts. She says the amount bid is the bidder's bid and the amount paid is the final amount charged by the search engine, which may be lower than the amount bid as a result of the authorized automatic reduction (Stricchiola First Report at para 61). Dr. Haruvy agrees that the bid amount is not necessarily the amount paid (Haruvy Second Report at para 123).

[107] Dr. Parkes says that the **amount bid and paid** refers to one amount only. He says that it means the bid submitted to the paid search engine is the amount paid for a click on the link (Parkes First Report, pg 58). He claims this is how the mechanism of a first price auction as used on GoTo.com and other paid search engines worked.

[108] I adopt Dr. Parkes' construction of 1(d) being a reference to one amount. This is consistent with the words chosen to describe the claim and is consistent with the intention that the patent is designed to function on then existing search engines. While bidding and paying may happen in two steps, the structure of the words used only supports an interpretation of it being a reference to one amount of money. It does not say "amount bid and amount paid" which would be suggestive of different amounts. Nor does it refer to "amounts". Ms. Stricchiola and Dr. Haruvy's construction of 1(d) would invite the possibility of multiple changes in the "amount". However, the inventors chose the singular "amount" and not "amounts". As well, there is no reference to the "final amount" if the phrase is intended to reference changes made throughout or during the bidding process.

Claim 2

2. The method of claim 1 further comprising retrieving one or more web pages that present statistic data on activity relating to the target keyword at the search engine.

[109] I construe **statistic data** to be a reference to statistics of the target keyword including the number of times a particular keyword was searched or used at the search engine.

Claim 3

3. The method of claim 2 wherein the statistic data comprises one or more of:
   the rate of use of a target keyword by users of the search engine;
   demographics of users of the target keyword;
   demographics or commercial information regarding bidders on the target keyword;
   identification of additional keywords used in conjunction with the target keywords by searches of a Internet search engine; and
   identification of additional keywords bid upon by bidders on the target keyword.

[110] This is a method that relates to the statistic data referred to in claim 2. The information referred to here is gathered from the search engine(s) and is made up of search (query) volume of a target keyword or topics related to the target keyword.

[111] I do not agree with Ms. Stricchiola and Dr. Haruvy who state that CTRs would be a rate of use. CTRs were part of the CGK and if the inventors intended to include that function here it would have been specified.

[112] Dr. Parkes states the **rate of use of a target keyword by users of the search engine** means the frequency of searches of the target keyword at the paid search engine, and points to the "view" column in the sample report at Figure 4 of the 167 Patent as an example (Parkes First Report, pg 59).

[113] Neither Ms. Stricchiola nor Dr. Haruvy construed the remaining terms of claim 3.

[114] The **demographics of users of the target keyword** means the information on users of the paid search engine who have searched using that keyword and would include gender, age, and location (Parkes First Report, pg 59).

[115] The **demographics or commercial information regarding bidders on the target keyword** would mean the information about the entities that are bidding on the target keyword on the paid search engine and such information might include the business sector or size of a business.

[116] I construe **identification of additional keywords used in conjunction with the target keywords by searchers of a Internet search engine** to mean the words used in the search string alongside the target keywords by searchers of an Internet search engine. I construe **identification of additional keywords bid upon by bidders on the target keyword** to mean other keywords bid on by the bidders who also bid on the target keyword (Parkes First Report, pg 60).

Claim 4

4. The method of claim 1 wherein the authorized automatic reduction is constrained by a minimum currency amount.

[117] This limits claim 1 such that the automatic reduction does not occur when the difference between the "entered amount" and the lower amount is smaller than a minimum currency amount (Parkes First Report, pg 60).

Claim 5

5. The method of claim 4 wherein the minimum currency amount is one cent.

[118] This limits claim 4 and claim 1 to a minimum bid of one cent.

Claim 7

7. The method of claim 1 wherein the authorized automatic reduction is a result of a change in a bid of another party.

[119] I construe this to mean that any authorized reduction because of a change in a bid at the paid search engine of another bidder on the target keyword.

Claim 8

8. The method of claim 1 wherein the authorized automatic reduction is performed in response to bids of other parties and other data.

[120] I agree with Dr. Parkes' construction that **other data** is a reference to non-bid data such as changes in keyword bid position, or an increase or decrease in the use of the keyword by users of the paid search engine. This is supported by Figure 4, which provides for reporting of keyword views data to a bidder (Parkes First Report, pg 61).

Claim 9

9. The method of claim 8 wherein the authorized automatic reduction is based on demographic information pertaining to searchers of the target keyword.

[121] As I have construed claim 8 in relation to keywords, I adopt Dr. Parkes' construction that **demographic information pertaining to searchers of the target keyword** to mean the demographics of the searchers of the target keyword(s) on the paid search engine and would include information like the age, gender, or location of a user.

Claim 17 (paragraph lettering added)

17. A method of managing an offeror's offer for a keyword made to a search engine, said offer identifying an amount said offeror will pay upon a searcher's use of an offeror-supplied reference located upon the keyword within said search engine, comprising

(a) receiving an authorization from said offeror,

(b) after receipt of said authorization, monitoring keyword offers at one or more Internet search engines to identify a change in said offeror's offer of interest to said offeror, and

(c) implementing said change in said offeror's offer on behalf of said offeror based upon the previously received authorization without further intervention of said offeror.

[122] Claim 17 is a method claim with various steps.

[123] Both Ms. Stricchiola and Dr. Haruvy construe the phrase **offeror's offer** as being broader than a "bidder's bid" (Stricchiola First Report at para 66; Haruvy Second Report at para 132).

[124] Dr. Parkes construes **offeror's offer** to means a bidder's bid and **identifying an amount said offeror will pay upon a searcher's use of an offeror-supplied reference** as meaning the money amount the bidder will pay the paid search engine when a searcher clicks their URL or hyperlink (Parkes First Report, pg 62).

[125] I adopt the construction that the offeror's offer is the bidder's bid.

[126] The step at 17(a) refers to the receiving the authorization for the offer.

[127] Ms. Stricchiola construes 17(b) **change in said offeror's offer of interest to said offeror** to be "contemplating that the offeror may be willing to offer different amounts for a clickthrough based on various criteria 'of interest' to them. This element does not specifically limit the types of information or criteria that may be of interest to an offeror, other than that they relate to offers at one or more (paid) search engines" (Stricchiola First Report at para 66). I find this construction is vague and does not provide guidance to the PSA on how to put this element into practice.

[128] I adopt Dr. Parkes' construction of 17(b) **monitoring keyword offers at one of more Internet search engines**, to mean repeatedly checking keyword bids at one or more paid search

engines, where monitoring can be done either automatically or manually (Parkes First Report, pg 63). He construes **identify a change in said offeror's offer of interest to said offeror** to mean "using the values of the monitored keyword bids to identify changes of interest" and either increasing a bid to recapture the bidder position or decreasing a bid where there is a bid gap (Parkes First Report, pg 63).

[129] The phrase at 17(c) – **implementing said change in said offeror's offer on behalf of said offeror based upon the previously received authorization without further intervention of said offeror** is construed by Ms. Stricchiola as referring to the automated function of changing the offer as a result of prior authorization by the advertiser (Stricchiola First Report at para 66). Dr. Haruvy adopts a similar construction (Haruvy Second Report at para 138).

[130] I adopt a construction of 17(c) that means making the change of interest to the offeror's offer and submitting the changed offer to the paid search engine.

[131] I do not agree that an automated function is specifically called out in claim 17(c) in relation to implementing the authorized change. The change is referenced in relation to the authority to do so, rather than an automatic function. Therefore the "change" called out in claim 17(c) may be done manually. This claim speaks to the authority to make the change – not the process by which the change is implemented.

Claim 18

18. The method of claim 17 wherein the identified change creates a differential in offers meeting certain criteria.

[132] I agree with Dr. Parkes that the skilled person would understand this to mean that "the identified change of interest creates a differential in the bids on a keyword meeting a certain criteria, for example, creating a differential of more than $0.01, or some other amount" (Parkes First Report, pg 64).

Claim 19

19. The method of claim 18 wherein said criteria identify differentials in offers characteristic of optimization opportunities.

[133] This claim relates back to claim 18 and speaks to the change being in relation to optimization opportunities (meaning closing a bid gap).

Claim 22

22. The method of claim 17 wherein the change comprises increasing an offer to obtain a rank position.

[134] I construe this to mean an increase in a bid to capture or recapture a desired position.

Claim 24

24. The method of claim 17 wherein the change is identified as a result of an increase or decrease in the use of a keyword by searchers of the search engine.

[135] This means, in relation to claim 17, a change that is promoted by the frequency of which a keyword is searched on a paid search engine.

Claim 25

25. The method of claim 17 wherein the change is identified as a result of a change in an offer of another party.

[136] I construe this to mean a change that is prompted by a change made by another user.

Claim 27

27. The method of claim 17 wherein said change is generated in response to offered prices and other data.

[137] I construe this to mean a change that is made in response to monetary and keyword information.

Claim 28 (paragraph lettering added)

28. A method that manages pay-per-click advertising, by determining an amount to be charged in response to a click of a hyperlink associated with a target keyword, comprising

(a) accessing with a computer processor, a particular amount a first advertiser is willing to be charged in response to a click of a hyperlink associated with said first advertiser;

(b) accessing with a computer processor, a statistic of relevance to said first advertiser and a statistic of relevance to a second advertiser, said statistics being related to one or more of rate of use by users, number of times a hyperlink was viewed, data that relates to an increase or decrease in the use of a keyword by users, demographics of users associated with a keyword, or demographics of advertisers associated with a keyword;

(c) without human intervention, determining with a computer processor, an amount to be charged to said second advertiser in response to a click of a hyperlink associated with said target keyword and said second advertiser; and

(d) wherein said amount to be charged to said second advertiser is determined using said particular amount, and is also determined using said statistic of relevance to said first advertiser and said statistic of relevance to said second advertiser.

[138] Ms. Stricchiola construes claim 28 as a new method that is different from claim 1, and is "directed to determining the actual amount to be charged (the price) for a click to an advertiser's ad, by accessing various other information beyond the bidder's bid […]." She construes this as being accomplished with "steps for accessing statistics related to the measurement of an advertiser's ad's 'relevance' to a searcher's (user's) search for a target keyword" (Stricchiola First Report at paras 67-69).

[139] According to Ms. Stricchiola's construction, the phrase at 28(b) **statistic of relevance** would include "relevance-related measurements, including but not limited to a 'statistic on a rate of use of a keyword' (such as query volume), clickthrough rate (CTR), and other advertiser-based statistics" (Stricchiola First Report at para 69).

[140] Ms. Stricchiola construes 28(d) – **determined using said particular amount, and is also determined using said statistic of relevance to first advertiser and said statistic of relevance to second advertiser** – to "broadly describ[e] a general process by which the search engine determines the actual amount to charge an advertiser for a click to their ad whereby a statistic of relevance for two or more competing ads is used to determine the cost of a click to an ad" (Stricchiola First Report at para 69).

[141] Dr. Parkes construes claim 28 as calling out a method to manage pay-per-click advertising with steps to determine the amount to charge for pay-per-click advertising in an auction (Parkes First Report, pg 65).

[142] Dr. Parkes construes **an amount to be charged in response to a click of a hyperlink associated with a target keyword** to be the amount that is charged to the advertiser when the hyperlink is clicked (i.e. the cost or price per click). He states "a 'target keyword' would be understood to mean a particular word, combination of words or short phrase that a bidder is bidding on in relation to pay-per-click advertising, such as pay-per-click advertising at a paid search engine" (Parkes First Report, pg 66).

[143] Dr. Parkes also states that the skilled person would understand 28(a) **accessing with a computer processor, a particular amount a first advertiser is willing to be charged in response to a click of a hyperlink associated with said first advertiser** to mean "using a computer that is determining the price, and reading from somewhere a bid amount of a 'first advertiser'. The Skilled Person would understand this bid amount to be a bid on the target keyword" (Parkes First Report, pg 66).

[144] Claim 28(b) **rate of use by users** is construed by Dr. Parkes to mean "the rate of use of a target keyword by users of a paid search engine" (Parkes First Report, pg 66). Ms. Stricchiola and Dr. Haruvy did not construe this term.

[145] Dr. Parkes states that the skilled person would understand 28(b) **accessing with a computer processor, a statistic of relevance to said first advertiser and a statistic of relevance to a second advertiser, said statistics being related to one or more of** to mean "using a computer and reading from somewhere a 'statistic of relevance' to the 'first advertiser' and a 'statistic of relevance' to a 'second advertiser'" where these are two distinct advertisers (Parkes First Report, pg 66).

[146] Dr. Parkes construes 28(c) **without human intervention, determining with a computer processor, an amount to be charged to said second advertiser in response to a click of a hyperlink associated with said target keyword and said second advertiser** to mean a computer automatically determining the pay-per-click amount to the second advertiser when their is a click of their hyperlink associated with the target keyword (Parkes First Report, pg 67).

[147] Contrary to Ms. Stricchiola's construction that **statistic of relevance** relates to the advertisement itself, Dr. Parkes interprets **statistic of relevance** to relate to the advertiser.

[148] I agree with Dr. Parkes and construe **statistic of relevance** to be a reference to the statistics relevant to the advertiser in relation to keywords. Nowhere does the patent nor the specifications speak to the assessment of the advertisement itself. The consistent theme throughout the patent is the activity of other advertisers relative to keywords. This is consistent with the language of the claims and the overall invention of the patent – that is, to improve the efficiencies in online marketing and address competition among advertisers.

[149] For claim 28, I adopt Dr. Parkes' construction.

Claim 29

29. The method of claim 28 wherein at least one said statistic comprises a statistic on the rate of use of a keyword.

[150] This would be understood to mean that the rate of use of a keyword by users is a statistic of relevance.

Claim 30

30. The method of claim 28 wherein at least one said statistic is related to a number of times a link to an advertiser's web site is produced by a paid search engine over a period of time.

[151] The skilled person would understand this to be a reference to the number of times an advertiser's hyperlink was viewed over a period of time at a paid search engine.

Claim 31

31. The method of claim 28 further comprising accessing additional keywords related to the target keyword.

[152] This would be understood to be a step of obtaining other keywords that have been used in conjunction with the target keyword (Parkes First Report, pg 68).

Claim 32

32. The method of claim 28 further comprising providing at least one said statistic for display on a single display screen to a user.

[153] I adopt Dr. Parkes' construction that '**a user**' means a searcher on the paid search engine and that this claim refers to a display of at least one of the statistics accessed (Parkes First Report, pg 68).

Claim 33

33. The method of claim 32 further comprising providing said amount for display on said single display screen.

[154] This relates back to claim 32, and would be understood to be the amount to be charged for the advertiser's hyperlink on the display screen.

Claim 34

34. The method of claim 28 further comprising providing said amount for display on a single display screen to a user.

[155] Dr. Parkes states that the skilled person would understand this to be "adding an additional step to claim 28 of providing for display on a single display screen to 'a user' the 'amount to be charged to said second advertiser', where 'a user' means a searcher on the paid search engine, i.e., displaying the cost per click amount of an advertiser's hyperlink in the paid search results" (Parkes First Report, pg 69).

[156] I adopt Dr. Parkes' construction.

Claim 36

36. The method of claim 28 wherein said determining step is performed by a computer responsible for making decisions on said second advertiser's use of pay-per-click marketing.

[157] I adopt Dr. Parkes' construction that this claim is "limiting claim 28 such that the computer determining 'an amount to be charged to said second advertiser' (i.e. determining the cost per click) is the same computer that is 'responsible for making decisions on said second advertiser's use of pay-per-click marketing'". Dr. Parkes also construes **computer responsible for making decisions on said second advertiser's use of pay-per-click marketing** to mean "the computer managing the use of the pay-per-click marketing for the advertiser, i.e., the computer responsible for managing the bid of the second advertiser" (Parkes First Report, pg 69).

Claim 37

37. The method of claim 36 further comprising sending said amount from said computer responsible for making decisions.

[158] I agree with Dr. Parkes' construction that the skilled person would understand this claim to be "adding an additional step to claim 36 in which the computer that determines an amount to be charged sends the amount to be charged somewhere. For example, the computer determines and then emails the amount to the second advertiser, or sends it to another computer" (Parkes First Report, pg 69).

Claim 39

39. The method of claim 28 wherein said amount meets pre-identified requirements of said second advertiser

[159] I construe this to mean that the amount referenced in claim 28 is determined by requirements as set by the second advertiser.

Claim 40

40. The method of claim 28 wherein determining the amount includes the use of a minimum currency amount.

[160] This limits the method of claim 28 to a minimum amount of money.

Claim 41

41. The method of claim 28 further comprising identifying keywords that have been used in conjunction with the keyword.

[161] This phrase would be understood to be a reference to the identification of other keywords that have been used in conjunction with the submitted keyword.

Claim 42

42. The method of claim 28 wherein one of said statistics comprises a statistic on the rate of use of a keyword and the other of said statistics is related to a number of times a link to an advertiser's web site is produced by the paid search engine over a period of time.

[162] I adopt Dr. Parkes' construction that the skilled person would understand this claim to be "limiting claim 28 to when the two statistics of relevance are 'a statistic on the rate of use of

a keyword' and a statistic 'related to a number of times a link to an advertiser's web site is produced by the paid search engine over a period of time'" (Parkes First Report, pg 70).

Claim 43

43. The method of claim 28 wherein said hyperlink associated with said first advertiser is ranked below said hyperlink associated with said second advertiser.

[163] This would be understood to be a reference to the position of first and second advertisers' hyperlinks in a paid search engine's results page.

Claim 44

44. A method that manages pay-per-click advertising, by determining an amount to be charged in response to a click of a hyperlink associated with a target keyword, comprising

    accessing with a computer processor, a particular amount a first advertiser is willing to be charged in response to a click of a hyperlink associated with said first advertiser,

    accessing with a computer processor, first and second different statistics related to one or more of a rate of use by users, number of times a hyperlink was viewed, data that relates to an increase or decrease in the use of a keyword by users, demographics of users associated with a keyword or demographics of advertisers associated with a keyword, and

    without human intervention, determining with a computer processor, an amount to be charged to a second advertiser in response to a click of a hyperlink associated with said target keyword and said second advertiser,

    wherein said amount to be charged to said second advertiser is determined using said particular amount, and is also determined using said first and second statistics.

[164] The experts agree that claim 44 and claim 28 are identical except that claim 44 refers to "statistics" whereas claim 28 refers to "statistics of relevance".

[165] Ms. Stricchiola construes "statistics" in claim 44 to mean something broader than (but including) the "statistics of relevance" from claim 28. She states that the difference is to whom the statistics apply (Stricchiola First Report at para 70). In my view, this construction is vague and does not promote an understanding of the language used in the claim.

[166] Dr. Parkes construes differences between claim 44 and claim 28 as follows:

> (i) In claim 28, the two statistics used to determine the amount to be charged to the second advertiser are "of relevance" to the first and second advertisers, respectively. In claim 44, there is no requirement that the statistics be relevant to the first and second advertisers.
>
> (ii) In claim 28, the two statistics used to determine the amount to be charged to the second advertiser can be the same statistic, but of relevance to each of the first and second advertisers. For example, in claim 28, both statistics could be the "rate of use by users". In contrast, in claim 44, the two statistics used to determine the amount to be charged to the second advertiser must be "different" statistics within the list of identified statistics (Parkes First Report at para 144).

[167] As I adopted Dr. Parkes' claims 28 construction, I also adopt his construction of claim 44.

[168] Further, the claims construction of the claims dependent on claim 44 will refer back to the corresponding dependent claims of claim 28.

Claim 45

45. The method of claim 44 wherein the statistic on a rate of use comprises a statistic on rate of use of a keyword.

[169] This is construed the same as claim 29 above.

Claim 46

46. The method of claim 44 wherein the statistic on a rate of use is related to a number of times a link to an advertiser's web site is produced by the paid search engine over a period of time.

[170] This is construed the same as claim 30 above.

Claim 55

55. The method of claim 44 further comprising determining a position of a hyperlink associated with an advertiser-specified Internet address relative to other hyperlinks associated with the keyword.

[171] I construe this to mean the advertiser understanding the position of their hyperlink relative to other hyperlinks in response to a keyword search.

Claim 56

56. The method of claim 44 wherein determining the amount includes the use of a minimum currency amount.

[172] Like claim 40 above, this is a reference to a minimum amount of money.

Claim 57

57. The method of claim 44 further comprising identifying keywords related to the target keyword.

[173] Like claims 31 and 41, this would be understood to be a reference to the identification of other keywords that have been used in conjunction with the submitted keyword.

Claim 58

58. The method of claim 44 wherein said hyperlink associated with said first advertiser is ranked below said hyperlink associated with said second advertiser identifying keywords related to the target keyword.

[174] Dr. Parkes construes the first part of claim 58 consistently with the construction of the same words in claim 43. However, he says that the additional phrase **identifying keywords related to the target keyword** is a typographical error (Parkes First Report, pg 74).

[175] I agree with and adopt Dr. Parkes' construction.

Claim 59

59. A computer system that manages pay-per-click advertising, by determining an amount to be charged in response to a click of a hyperlink associated with a target keyword, comprising

a memory, and

processing hardware configured to:

access from memory a particular amount a first advertiser is willing to be charged in response to a click of a hyperlink associated with said target keyword and associated with said first advertiser,

access from memory a statistic of relevance to said first advertiser and a statistic of relevance to a second advertiser, said statistics being related to one or more of rate of user by users, number of times a hyperlink was viewed, data that related to an increase or decrease in the use of a keyword by users, demographics of users associated with a keyword, or demographics of advertisers associated with a keyword, and

> without human intervention, determine an amount charged to a second advertiser in response to a click of a hyperlink associated with said target keyword and said second advertiser,
>
> wherein said amount to be charged to said second advertiser is determined using said particular amount, and is also determined using said statistic of relevance to said first advertiser and said statistic of relevance to said second advertiser.

[176] Claim 59 includes the same elements as claim 28 except that it refers to a computer system that manages pay-per-click advertising.

[177] Ms. Stricchiola states, "[c]laim 59 is conceptually similar to claim 28 with the exception that the claim is directed to a computer system rather than a method" (Stricchiola First Report at para 72). Dr. Haruvy agrees (Haruvy Second Report at para 146).

[178] Dr. Parkes agrees that claim 59 is very similar to claim 28, and the skilled person would have the same technical understanding of claim 59 as claim 28. For claim 59, "the Skilled Person would know that they require a computer (or 'a computer system') to have a memory and processing hardware configured to access that memory" (Parkes First Report at para 149).

[179] I agree that claim 59 is similar to claim 28, and adopt Dr. Parkes' construction of the essential elements as described in claim 28.

Claims 60, 61, 62, 63, 64, 65, 67, 70, 72, 73, 74

[180] I construe these claims in the same manner as the similarly worded claims 29, 30, 31, 32, 33, 34, 36, 39, 41, 42 and 43.

Claim 75

75. A computer system that manages pay-per-click advertising, by determining an amount to be charged in response to a click of a hyperlink associated with a target keyword, comprising

a memory, and

processing hardware configured to:

access from memory a particular amount a first advertiser is willing to be charged in response to a click of a hyperlink associated with said first advertiser,

access from memory first and second different statistics related to one or more of a rate of use by users, number of times a hyperlink was viewed, data that relates to an increase or decrease in the use of a keyword by users, demographics of users associated with a keyword or demographics of advertisers associated with a keyword, and

without human intervention, determine an amount, to be charged to a second advertiser in response to a click of a hyperlink associated with said target keyword and said second advertiser,

wherein said amount to be charged to said second advertiser is determined using said particular amount, and is also determined using said first and second statistics.

[181] I adopt Dr. Parkes' construction of claim 75 as being similar to claim 44 except that it refers to a computer system and not a method. Otherwise, in all other respects the claims are identical (Parkes First Report, pg 76).

Claims 76, 77, 78, 79, 80, 81, 83, 86, 88, 89

[182] I construe these claims in the same manner as the similarly worded claims 45, 46, 39, 32, 33, 34, 36, 55, 57, and 43.

VI. Validity

[183] Google raises a number of validity issues with the 167 Patent. Once a patent is issued, it is presumed to be valid (*Patent Act*, RSC 1985, c P-4 [*Patent Act*], s. 43(2)). Accordingly, Google bears the burden of proof.

[184] Below I will address the validity issues that are determinative of this action.

A. Sufficiency

[185] Google alleges that claims 8, 9, 24, 27, 28, 44, 59 and 75 are not valid as they are not sufficiently disclosed.

Legal Principles

[186] Subsection 27(3) of the *Patent Act* requires that the specification of a patent correctly and fully describe the invention and the operation or use of the invention as contemplated by the inventor.

[187] The test for sufficiency was set out in *Teva Canada Ltd v Pfizer Canada Inc*, 2012 SCC 60 [*Teva*] at paragraph 90: "the relevant question is whether the disclosure was sufficient as of the date of filing." The Court asks:

1) What is the invention;
2) How does it work; and,
3) Having only the specification, can a POSITA successfully produce the invention using only the instructions contained in the disclosure (*Teva* at paras 70-71).

[188] Sufficiency has two requirements: disclosure and enablement. As described by Justice Stratas, "the details published by the inventor in the disclosure must permit a person skilled in the art to recreate the invention as claimed" (*Hershkovitz v Tyco Safety Products Canada Ltd*, 2010 FCA 190 at para 14).

[189] A mere conception is not an invention unless the idea is also set into a practical shape (*Apotex Inc v Wellcome Foundation Ltd*, 2000 CanLII 16270 (FCA) at para 31).

[190] Further, as noted by Justice Binnie in *Free World Trust* at para 32:

> The claims cannot be stretched to allow the patentee to monopolize anything that achieves the desirable result. It is not legitimate, for example, to obtain a patent for a particular method that grows hair on bald men and thereafter claim that *anything* that grows hair on bald men infringes [emphasis in original].

Analysis

[191] The invention of the 167 Patent is a method and apparatus to ease the burden for advertisers managing keyword bids on paid search engines. The 167 Patent does not disclose the operation of a paid search engine and it does not disclose an auction. It is a method designed to work on the pricing formats (auctions) used on the existing paid search engines.

[192] Ms. Stricchiola's evidence – both in her testimony and in her expert report – was that the 167 Patent is a "concept" that requires a "further step" to put the invention into practice (Stricchiola Second Report at para 118).

[193] Both Ms. Stricchiola and Dr. Haruvy claim that the patent discloses a GSP auction, but state that the patent is not bound by a specific logic or algorithm. Dr. Haruvy states that a skilled person would be able to use non-bid data by virtue of their knowledge of multi-dimensional auctions. This, however, is in contrast to his opinion that the PSA does not require in-depth knowledge of auction design.

[194] As noted in my construction of claim 1 of the patent, I do not construe the claims language to disclose a GSP auction. The patent claims, the specifications, and the embodiments do not disclose the workings of a GSP auction or detail steps or methods to achieve that process.

B. Sufficiency of Claims 8, 9, 24 and 27

[195] Google's expert, Dr. Parkes, states that dependent claims 8, 9, 24 and 27 are not sufficiently disclosed.

[196] Dr. Parkes says that claims 8 and 9 lack sufficient disclosure or direction on how use the information from "other data" (claim 8) and/or "demographic information" (claim 9) to modify a bid amount (Parkes First Report at paras 193-197).

[197] Dr. Parkes also says that claim 24 does not provide any direction on how to modify a bid in a meaningful way based upon the increase or decrease in the use of a keyword by searchers (Parkes First Report at para 200).

[198] On claim 27, Dr. Parkes says that the reference to "offered prices" and "other data" is missing any direction or instruction on how put this information to use in practicing the patent.

[199] On claim 9, I agree that there is no disclosure of how the PSA is to authorize an automatic reduction of a bid based on the demographic information of searchers. For example, accepting that "demographic information" would include information like the age, gender, or location of a user, the patent does not enable the PSA to understand how to determine whether to reduce their bid if the searchers of their target keyword (i.e. "Christmas cards") are determined to be mainly comprised of 25-35 year-olds in Ontario.

[200] Regarding claims 8, 24, and 27, I am satisfied that the PSA with the CGK would be able to put these claims into use when these claims are considered in the context of the full patent, including the specifications and the embodiments.

[201] I therefore conclude that dependent claims 8, 24 and 27 are sufficiently disclosed.

C. Sufficiency of Claims 28, 44, 59 and 75

[202] The experts agree that independent claims 28, 44, 59 and 75 of the 167 Patent teach a different method and process from those disclosed in claims 1-27. Specifically, claims 28 and higher claim methods and systems for the management of pay-per-click advertising with a computer system (claim 59 and 75) and a computer processor (claim 28 and 44) with reference to statistics of relevance (claims 28 and 59) and statistics (claim 44 and 75).

[203] The patent language, the specifications and the embodiments do not provide any further direction on the function or operation of the pay-per-click advertising process itself. The 167 Patent is silent on technical direction on how to put the pay-per-click advertising process into practice on a computer system or a computer processor. There is no code or algorithm on how to put the claims into practice. The specifications and the embodiments also fail to disclose a methodology or code to be used in order to put the pay-per-click advertising functions into practice on a search engine.

[204] The patent also fails to explain or disclose how to make use of the statistical information referenced in these claims. The claims, specifications, and embodiments provide no direction

or information on how the statistics are compiled, or how to use this information once compiled within the computer system or computer processor. Even accepting that the relevant statistical information will differ among advertisers, the PSA must nonetheless be able to take the patent and put the patent into use based upon the claims language, specifications and embodiments. There is an absence of information in the specifications on how to compile the statistical information and how to put that information to use for the purpose of bidding.

[205] Despite the absence of information in the specifications on how or what the PSA is to do with the statistics, I recognize that the Court must endeavor to give a purposive construction to the claims. However, the Court cannot rewrite the claims and must assess the claims based on the language used. The PSA would possess knowledge of advertising on the internet and some coding ability, however, I am not satisfied that the PSA would possess the technical knowledge necessary to put these claims into practice even with the use of the broad applicable CGK.

[206] I therefore conclude that claims 28, 44, 59 and 75 (and the dependent claims) which are largely stand-alone claims within the 167 Patent, are not sufficiently disclosed and cannot be enabled. Therefore they are not valid.

[207] Having concluded that claims 9, 28, 44, 59 and 75 are not valid they will not be considered under the anticipation or obviousness analysis.

D. Priority Date

[208] The applicable filing date is relevant for the anticipation and obviousness analysis.

[209] The application for the 167 Patent was filed on July 5, 2001, and claims priority to the 976 Patent that was filed on July 5, 2000.

[210] The legal test governing priority claims to an earlier filed application is set out in subsection 28.1(1) of the *Patent Act*, which requires that the priority patent disclose the subject-matter defined by the asserted claims.

[211] In *Hospira Healthcare Corporation v Kennedy Trust for Rheumatology Research*, 2020 FCA 30 [*Hospira*] at para 63, the Court confirmed that the question of priority affects which prior art may be relevant for the purposes of the attack on the validity of the patent.

[212] In the 976 Patent application, the invention is titled "Paid Search Engine Bid Management" and the inventors are Jon Keel, Juan Velez, and Daren Murrer. The field of invention states: "[t]he present invention relates to the management of advertising expenses in online media." The 976 Patent discloses the following under the heading 'benefit'":

> Function 1 = get knowledge (Vs information) to make informed decisions about marketing using Paid Search Engines
> Element 11 = Compile the information
> Element 12 = Organize the information
> Function 2 = Monitor changes and opportunities
> Element 21 = Compile the information
> Element 22 = Compare the information
> Element 23 = Get opportunities and changes to the people or computer who make the decision about how to use the Paid Search Engines

[213] The parties agree that priority date considerations apply only to claims 1-27 of the 167 Patent, as claims at 28 and higher are different. For the priority date to apply, the 976 Patent must disclose claims 1 to 27 of the 167 Patent.

[214] Google's expert Dr. Parkes argues that the 976 Patent is limited to the KBO bid monitoring tool, and since claims 1-27 of the 167 Patent relate to a bid management tool, the 976 Patent does not disclose these claims (Parkes First Report at para 157). There is no dispute that the KBO bid monitoring tool disclosed in the 976 Patent (Exhibit 4) is the same as Figure 4 of the 167 Patent.

[215] I would characterize the main area of difference among the experts in comparing the 976 Patent and the 167 Patent, is the issues of the automatic bidding function.

[216] PSET's expert, Ms. Stricchiola, states that claims 1-27 of the 167 Patent are reasonably inferable from language of the 976 Patent. In support, she relies upon Element 23 of the 976 Patent (above) which states: "Get opportunities and changes to the people or computer who make the decision about how to use the Paid Search Engines". She argues that this is a reference to automatic bidding as outlined in the 167 Patent. Similarly, Dr. Haruvy states that in Element 23, a computer making decisions would be understood to be doing so automatically (Haruvy Second Report at para 185).

[217] Similarly, both Ms. Stricchiola and Dr. Haruvy state that "Function 2 = Monitor changes and opportunities" involves active bid management. Ms. Stricchiola argues that an automatic function inference can also be drawn from this language.

[218] There is no direct reference to an automated process in the 976 Patent, nor is there a reference to an automated process in this phrase, and I do not agree that this can reasonably be interpreted to refer to an automated step. The automation step in the 167 Patent – which is one of the core inventive steps – would need to be more clearly and obviously disclosed in the 976 Patent in order to support the priority date claim. While the word "automatic" need not necessarily appear, the PSA must ultimately be able to read the 976 Patent and logically conclude that the language speaks to automation. Here, I do not see how the PSA would be led to that conclusion.

[219] Furthermore, the use of non-bid data in the automatic reduction step is not referenced in the 976 Patent. Nor does it speak to monitoring and implementing changes as referenced in claim 17 of the 167 Patent. The 976 Patent at Function 2 only references monitoring changes and opportunities.

[220] PSET relies upon *Astrazeneca AB v Apotex Inc*, 2007 FC 688, to argue that the claim to priority is met if an "inference" can be drawn that the 976 Patent discloses the 167 Patent. However, in my view, PSET misstates the finding of the Court. It is clear that the Court was looking for an actual disclosure where it states at para 63:

> The specific issue is whether the first Swedish priority application **disclosed** the use of omeprazole and clarithromycin as a

> combination therapy. If it did not, then Logan is clearly citable art [emphasis added].

[221] The word "disclose" appears in various sections of the *Patent Act*. I do not accept that a different meaning should apply to disclosure as it is used in subsection 28.1(1) such that disclosure is achieved merely if an inference can be drawn. The priority application must disclose the patent. The 976 Patent only speaks to bid monitoring. There is no reference to management tools to assist with the bidding process, nor does it reference any automatic bidding features. Therefore, the 976 Patent does not disclose the 167 Patent.

[222] As a result, I conclude that the relevant date for assessing whether claims 1-27 were obvious or anticipated is the filing date of July 5, 2001.

E. Anticipation

[223] Google argues that the 167 Patent is invalid for anticipation, or lack of novelty. Google bears the burden of proof on a balance of probabilities standard (*Bell Helicopter Textron Canada Limitée v Eurocopter, société par actions simplifiée*, 2013 FCA 219 at para 105).

Legal Principles

[224] A patent is invalid if it is not new and the invention had been previously disclosed (*Patent Act*, s. 28.2(1)). Pursuant to paragraph 28.2(1)(b), the subject matter defined by the claim must not have been disclosed before the claim date in such a manner that it became available to the public.

[225] As described by Justice Rothstein in *Sanofi* at para 26, there are two requirements to prove anticipation: disclosure and enablement.

[226] The disclosure requirement means that "the prior art, as of the claim date, must disclose subject-matter, which if performed, would necessarily result in an infringement of the patent" (*Eli Lilly* at para 241; *Sanofi* at para 25).

[227] With respect to enablement, the Court asks "whether a [person of skill in the art] would have been able to perform the invention" (*Eli Lilly* at para 241; *Sanofi* at para 26).

[228] It is the essential elements of the claimed invention that must be enabled (*Hospira* at para 74).

[229] Enablement must come from a disclosed single prior art reference (*Eli Lilly* at para 241) with a clear direction such that "a skilled person reading and following it would in every case and without possibility of error be led to the claimed invention" (*Beloit Canada Ltd v Valmet Oy*, [1986] FCJ No 87 at para 29 [*Beloit*]; *Seedlings Life Science Ventures, LLC v Pfizer Canada ULC*, 2020 FC 1 at para 109).

[230] Justice Rothstein outlined a non-exhaustive list of factors to be considered with respect to enablement (*Sanofi* at para 37):

> 1. Enablement is to be assessed having regard to the prior patent as a whole including the specification and the claims. There is no reason to limit what the skilled person may consider in the prior patent in order to discover how to perform or make

> the invention of the subsequent patent. The entire prior patent constitutes prior art.
>
> 2. The skilled person may use his or her common general knowledge to supplement information contained in the prior patent. Common general knowledge means knowledge generally known by persons skilled in the relevant art at the relevant time.
>
> 3. The prior patent must provide enough information to allow the subsequently claimed invention to be performed without undue burden. When considering whether there is undue burden, the nature of the invention must be taken into account. For example, if the invention takes place in a field of technology in which trials and experiments are generally carried out, the threshold for undue burden will tend to be higher than in circumstances in which less effort is normal. If inventive steps are required, the prior art will not be considered as enabling. However, routine trials are acceptable and would not be considered undue burden. But experiments or trials and errors are not to be prolonged even in fields of technology in which trials and experiments are generally carried out. No time limits on exercises of energy can be laid down; however, prolonged or arduous trial and error would not be considered routine.
>
> 4. Obvious errors or omissions in the prior patent will not prevent enablement if reasonable skill and knowledge in the art could readily correct the error or find what was omitted.

Analysis

[231] PSET's experts, Dr. Haruvy and Ms. Stricchiola, agree with Dr. Parkes that claims 17, 18, 19, 22, and 25 are anticipated if the claim date is July 5, 2001 (Haruvy Second Report at paras 273, 281; Stricchiola Second Report at paras 172, 174-175, 178).

[232] The remaining claims on which the experts disagree are claim 1 (and dependent claims 2-5, 7, 8), claim 24, and claim 27.

976 Patent

[233] The first prior art to be considered is the 976 Patent, which discloses the KBO bid-monitoring tool that operated by gathering bid and keyword data from GoTo.com, and providing reports on this data. The reports displayed the bid amounts of other bidders on keywords, and notified users of opportunities to raise or lower bids. Based upon these features, the 967 Patent is very similar to the 167 Patent. However, as noted in my analysis regarding the priority date, the KBO tool did not, and the 976 Patent does not, disclose the function of automatic bid reductions as disclosed in claim 1 of the 167 Patent.

[234] On May 26, 2000, Mr. Keel sent an email to Jessie Brader at Planet Ocean (Exhibit 46), where he states he is providing "information on the new paid search engine tool". This email explains the features of the KBO tool and provides a username and password to Planet Ocean to allow them to evaluate the tool. In the email, Mr. Keel explains how the KBO tool works, and describes its features as follows as:

> *Shows the first 20 goto.com bid positions for keywords including current bid amounts
>
> *Analyzes the top 20 goto.com positions for an unlimited number of keywords
>
> *Shows your present top 20 keyword bid positions in red – your competitors' bid positions and bid amounts are shown in yellow
>
> *Automatically updates each night the bid positions and bid amounts for all keywords you've entered
>
> *Conveniently formats for easy reading and printing right from your browser
>
> *Shows you the previous month's goto.com search counts for each keyword you've entered

> *Allows you easy access to your account via a user name and password
>
> *Arranges your keywords alphabetically
>
> *Lets you easily add new keywords to track using the real time "add a keyword" feature
>
> *Daily notifies you by email of changes in your bid position and opportunities you may have to raise or lower your bid amounts
>
> *Gives you the option of managing keywords and key phrases for multiple domains.

[235] The issue is whether this email was a public disclosure of the KBO tool and therefore anticipates the 167 Patent. Although Mr. Keel claimed that this communication was confidential, there is no reference in the email to the information being provided in confidence. This was also contradicted by Ms. Brader who testified that this type of communication was common at that time. She noted that a username and password were provided, suggesting that there was no expectation that the information would be treated confidentially. This public disclosure is only relevant to this analysis if the 976 Patent fully anticipates the 167 Patent.

[236] As noted in the sufficiency analysis above, the 976 Patent does not include an automatic reduction feature, which is an essential element of claim 1 of the 167 Patent.

[237] Given this, I do not conclude that the PSA could have performed or enabled the teachings of the 167 Patent via the 976 Patent, even with the benefit of the CKG and trial and error. In my view, there are additional steps that would be necessary beyond trial and error to get to an automated process.

[238] Therefore, I am not satisfied that the 976 Patent fully anticipates the 167 Patent.

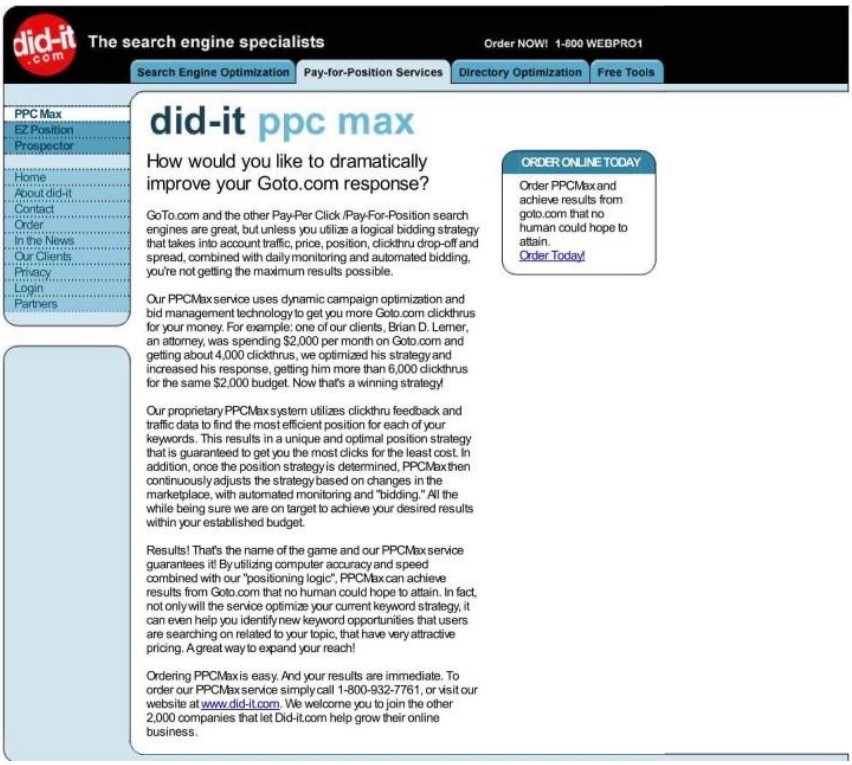
Other Prior Art

[239] Ms. Stricchiola agrees that as of the July 2001 claim date, the following prior art was available: ClickPatrol, ManageBid, Did-It, and goClick. Although Google's expert, Dr. Parkes, refers to additional prior art in his report, for this analysis it is sufficient to consider the prior art that both Ms. Stricchiola and Dr. Parkes agree was available as of July 2001.

[240] Did-It was a commercial bid management service that submitted, monitored, and maintained bids. Web archives show the service was advertised as being able to "optimize your current keyword strategy" and "help you identify new keyword opportunities." Did-It provided three services: PPC Max, EZ Position, and Prospector. PPC Max was an automated bid management tool, which monitored ongoing auctions at GoTo.com and collapsed bid gaps on behalf of users using "positioning logic". Prospector was a bid reporting service, which allowed users to check keywords and prices at various search engines, while EZ position was a bid monitoring service, which notified users of opportunities to lower their bids while maintaining their position.

[241] The Court heard direct evidence from Mr. Lee who worked at Did-It. He testified that in 1999, Did-It started building automation tools internally to manage bids for pay-per-click advertising, and eventually released a suite of tools to the public. These tools worked on GoTo, as well as other search engines like FindWhat, Kanoodle, and Sprinx. Mr. Lee explained the various tools offered by Did-It, including PPC Max, EZ Position, and Prospector, and stated

they were all released as publicly available tools by 2001. The Did-It PPC Max functions are described in the following (Exhibit 115):



[242] The description of the Did-It PPC Max service covers the essential elements of claim 1 and claim 17 of the 167 Patent by: retrieving a webpage, entering a target keyword and the amount including the maximum amount, authorizing an automatic reduction, and generating reports

[243] In addition to Did-It, the other prior art that was available is summarized as follows:

a. **ClickPatrol** was a free, online bid management tool that submitted, monitored, and maintained bids on paid search engines like GoTo.com. Internet Archives show that its purpose was to “buy and manage ads” on a pay-per-click search engine. The interface

allowed users to view and change their bids, increase balances, and add new keywords, URLs and bid amounts for a paid search engine. ClickPatrol also had a "MoneySaver" feature, which would automatically monitor for bid gaps, such that if there was a gap between a user's bid and the next ranked bid, the service would drop the user's bid to save money while still maintaining their position.

b. **ManageBid** was a commercial bid management tool that submitted, monitored, and maintained bids at paid search engines. Users would select the search engines they wished to place bids in, and input keywords, the ranking they wished to achieve, and the maximum price they wished to pay for a keyword. ManageBid would generate a list of keywords, and return a matrix of the bids a user needed to make to achieve the desired ranking for each keyword. ManageBid could also be used to "generate lists of keywords/phrases similar to those you already know" and modify bids. Users could also set "auto-pilots" to maintain their ranking while keeping their bids optimized.

c. **goClick** was a pay-per-click search engine which used a tool called "BidMaster" to monitor and adjust bids and positions automatically. As described on the website, the "Auto-BidMaster tool adjusts your bid amounts for you once a day automatically, ensuring that you stay on top of the search results and eliminating the need to keep logging in to adjust your bids."

[244] This prior art discloses the following functionality:

- bidding for position was practiced;
- bid collapsing was practiced;
- the process of monitoring and tracking bid positions was known and practiced; and,

- there were automated bidding products available.

[245] Based on my conclusions, individually Did-it, ClickPatrol, ManageBid or goClick anticipate the essential elements of claims 1 and 17 (and the dependent claims) of the 167 Patent.

F. Obviousness

Legal Principles

[246] A patent is not valid if it is not inventive or if the invention it claims would have been obvious to the skilled person (*Patent Act*, s. 28.3). Obviousness and anticipation can be distinguished as follows:

> […] obviousness is an attack on a patent based on its lack of inventiveness. The attacker says, in effect, "Any fool could have done that." Anticipation, or lack of novelty, on the other hand, in effect assumes that there has been an invention but asserts that it has been disclosed to the public, prior to the application for the patent. The charge is: "Your invention, though clever, was already known" (*Beloit* at para 12).

[247] A patent claim will be invalid if, based upon the information that was available to the public before the claim date, the subject matter would have been obvious to the PSA.

[248] An obviousness inquiry follows a 4 step approach – the *Windsurfing* approach – as restated by the Supreme Court of Canada in *Sanofi*:

> (1) (a) Identify the notional "person skilled in the art";

> (b) Identify the relevant common general knowledge of that person;
>
> (2) Identify the inventive concept of the claim in question or if that cannot readily be done, construe it;
>
> (3) Identify what, if any, differences exist between the matter cited as forming part of the "state of the art" and the inventive concept of the claim or the claim as construed;
>
> (4) Viewed without any knowledge of the alleged invention as claimed, do those differences constitute steps which would have been obvious to the person skilled in the art or do they require any degree of invention (at para 67).

[249] If an "obvious to try" test is warranted, additional factors should be considered at the fourth step (*Sanofi* at para 69). These factors include:

> 1. Is it more or less self-evident that what is being tried ought to work? Are there a finite number of identified predictable solutions known to persons skilled in the art?
>
> 2. What is the extent, nature and amount of effort required to achieve the invention? Are routine trials carried out or is the experimentation prolonged and arduous, such that the trials would not be considered routine?
>
> 3. Is there a motive provided in the prior art to find the solution the patent addresses?

[250] The proper approach to considering the *Sanofi* inventive concept factor has been addressed in a number of cases. Recently, in *Western Oilfield Equipment Rentals Ltd v M-I LLC*, 2021 FCA 24, the Court said at para 109 that *Sanofi* is one way to assess obviousness but it is not mandatory. The Court also warned against an overly rigid approach to the obviousness inquiry.

Analysis

[251] The PSA and the CGK addressed above apply to this analysis. The inventive concept is derived from the claim construction exercise, and the Court may have regard to the patent specification where it is not possible to determine the inventive concept from the claims alone (*Apotex Inc v Shire LLC*, 2021 FCA 52 at para 67).

[252] In my view, regardless of the approach taken to assessing the inventiveness concept, it is easily arrived at here. Claims 1 and 17 (and their dependents) is a system for bidding on multiple keywords that tracks, collects and displays information for advertisers to identify opportunities to adjust bids on paid search engines. This would enable advertisers to stay competitive with other advertisers and avoid overpaying for a keyword relative to what other advertisers were paying for that same keyword. The impetus behind this invention was that paid search engines were not making this competitive information readily accessible.

[253] The 167 Patent also discloses the potential of automatic bidding. While Mr. Soper confirmed that the code he wrote for the patent enabled automatic bidding, this feature was never enabled for customers of PSET. Nonetheless, this function is called out in the patent.

[254] As noted noted, the prior art (reviewed above) demonstrates there were other products on the market at the relevant time (July 2001) that were offering the same features and functions for assistance with online marketing, including ClickPatrol, ManageBid, Did-It, and goClick.

[255] PSET raised objections to some of the prior art references on the basis that it was located by legal counsel for Google and not by Google's expert. They argued that the PSA would not have found the art after a reasonable search. However, the Court in *Hospira* states that "it is an error to exclude from consideration prior art that was available to the public at the relevant date simply because it would not have been located in a reasonably diligent search" (at para 86).

[256] Further, considering the particular field of invention – online advertising – I am satisfied that the PSA with familiarity with online research would have been able to locate these prior art references without difficulty.

[257] Considering this, there is no discernable difference between ClickPatrol, ManageBid, goClick and Did-It and the functions of claims 1 and 17 of the 167 Patent. For example, Did-It's PPC Max tool that was available by July 2001 offers the same services as those detailed in the 167 Patent.

[258] A further relevant consideration is whether "the inventor and his or her team reached the invention quickly, easily, directly and relatively inexpensively, in light of the prior art and common general knowledge" (*Sanofi* at para 71). This statement is particularly relevant in this case where the inventors testified that they reached the claimed invention after a short brainstorming session.

[259] In applying the *Sanofi* test, I conclude that even if there were any differences between claims 1 and 17 of the 167 Patent and the prior art, they are differences that would be overcome by the PSA in light of the CGK.

[260] This is not a case where the application of the obviousness analysis strains credulity. The evidence is clear that even on a rigid application of the obviousness considerations there were multiple other commercial products that performed the same essential features, and in some cases more, as those described in claims 1 and 17 of the 167 Patent.

[261] As evidenced by the prior art, a number of other companies were offering the same products as PSET. As described by Mr. Lee, in the early 2000s there were a lot of companies focusing on the bid and search management business and there was rapid development.

[262] I conclude that claims 1 and 17 (and their dependents) are invalid for obviousness.

G. Other Invalidity Grounds

[263] Google also argues that the 167 Patent is invalid on the grounds of overbreadth, utility, and non-patentable subject matter. However, given my findings on sufficiency, anticipation and obviousness, it is not necessary to address these grounds.

VII. Infringement

[264] Although I have concluded that all of the asserted claims are invalid, I will nonetheless address the infringement allegations.

[265] In their Further Amended Statement of Claim, PSET claims that Google infringes the 167 Patent in using, offering for sale and selling search engine advertising inventory to customers in Canada through its AdWords method and computer system. It is PSET's position that the 167 Patent covers core elements of the Google Ads system – including Google Ads' use of a second price auction – as well as its use of ad quality and user demographics data in the ranking of ads and in the pricing of clicks to those ads.

[266] The evidence is that Google Ads has evolved into a complicated machine driven system, however, the question on infringement is if the foundation of the Google Ads system infringes the 167 Patent. I would further note that the focus of this inquiry starts back in the early 2000s when Google was an emerging technology company and not the tech giant it is today.

A. Legal Principles

[267] The applicable legal principles are described in *dTechs* as follows:

> Section 42 of the *Patent Act* grants the patent holder the exclusive right, privilege and liberty of making, constructing and using the invention and selling it to others to be used. A patent is infringed by any act that interferes with the patentee's full enjoyment of the monopoly granted […].
>
> Pursuant to s 55(1) of the *Patent Act*, any person who infringes a patent is liable for all damages sustained by the patentee after the grant of the patent by reason of infringement. The burden of

> proving infringement rests with the party that alleges it […] (at paras 162-163).

[268] There is no infringement if an essential element is different or omitted (*Free World Trust* at para 31). If an independent claim is not infringed, then no infringement of the dependent claims is established.

[269] The six independent claims of the 167 Patent for consideration are: 1, 17, 28, 44, 59, and 75. Claims 1, 17, 28, and 44 are method claims whereas claims 59 and 75 are systems claims. To infringe a method claim, "an alleged infringer must carry out the claimed method" (*dTechs* at para 177).

[270] The burden to establish infringement rests on PSET. The question for the Court is whether Google practices the essential elements of the asserted claims. The parties take the position that all elements of all claims are essential. The analysis on infringement flows from the construction of the 167 Patent.

Analysis

[271] In support of its position on infringement, PSET relies upon the opinion of Ms. Stricchiola who states as follows in her First Report at paras 13 and 14:

> Google Ad and AdSense for search infringe the asserted claims of the 167 Patent.
>
> The 167 Patent broadly covers core elements of the Google Ads system, including Google Ads' use of a second price auction, as

> well as it use of ad quality and user demographics data in the ranking of ads and in the pricing of clicks to those ads.

[272] The foundation of Ms. Stricchiola's infringement opinion is that the 167 Patent discloses a GSP auction. This is her construction opinion and the opinion of Dr. Haruvy. I do not construe the 167 Patent as disclosing an auction format. The 167 Patent does not mention or discuss auction formats. The only reference to an auction is the reference to the GoTo search engine which was known to be operating an open first price auction. On this ground alone, Ms. Stricchiola's infringement opinion is flawed as it is premised upon a foundational assumption that I do not accept. Although this is in my view determinative of the infringement claim, I will nonetheless consider the independent claims.

[273] I will first provide an overview of the alleged infringing product.

B. Google's Alleged Infringing Products

[274] Google sells advertising on its various internet platforms. Dr. Veach, who has been with Google since July 2000, explained that the advertising offered by Google at that time was called Adwords and it was a self-service system for advertisers. He explained that there were issues with ad quality and monetizing this system.

[275] He worked on a system to optimize advertising and in February 2002, AdWords Select was launched. It differed from the early version of AdWords because it used an auction model, and advertisers were only charged when a user clicked on an ad (cost-per-click) rather than

being charged each time an ad was displayed. This system was dynamic in that the ad that was displayed to a user and the price the advertiser had to pay was determined at the time of the search itself, on a search-by-search basis.

[276] According to Dr. Veach, AdWords Select works in a closed auction system, meaning the bids of advertisers are sealed and not known to each other. The cost an advertiser would have to pay per click was determined by calculating the bid the advertiser would have to submit to obtain the same score as the runner up bidder, and rounding up. There was also an option to have advertiser's bids adjusted if they had opted to use automated bidding strategies.

[277] In AdWords Select, ads were ranked based upon two criteria: the bid and the predicted clickthrough rate (pCTR). Dr. Veach described CTRs as being a measure of ad quality. As the process was explained by Dr. Veach, if Advertiser A was bidding $1, and Advertiser B was bidding $0.25, but Advertiser B received a higher pCTR, they could be ranked above Advertiser A, despite Advertiser A bidding a higher amount.

[278] Dr. Veach described Google Ads as a complicated and continually evolving system that now uses machine learning technology.

[279] Google employees, Mr. Moser and Mr. Iljic testified that currently Google Ads runs [redacted] on every search results page generated by each user query. Auction positions are based on LTV (long term value) scores which incorporates machine learning prediction models, the context of the specific search,

[redacted], and the bid. Google also offers automated bidding strategies, which allow advertisers to interact with the ads system through automated bidding agents – which are separate from the auction.

[280] The evidence of Mr. Iljic was that Google Ads is not an open auction and not a first price auction. The auction results are not ordered by bid amount and there is no ongoing auction. He also explained that at the time of the development of AdWords Select, the focus was on the end user – being the searcher rather than the advertiser. This is consistent with Google's corporate focus on "search".

[281] Mr. Iljic was Google's witness on the auction process used by Google. He was not cross-examined about Google's auctions by PSET's lawyers. Accordingly, his direct evidence on the operation and function of Google Ads is accepted as largely uncontradicted.

Claim 1

[282] Ms. Stricchiola's opinion is that claim 1 of the 167 Patent is infringed by Google Ads because:

- Google Ads provides performance reports which provide advertisers with detailed reporting on their Max CPC for a keyword (the amount bid) and on the subsequent actual CPC (the price) the advertiser paid for a clickthrough (Stricchiola First Report at para 119);

- The second price auction in Google Ads automatically reduces the advertiser's entered bid amount to a lower amount that will not change their position (Stricchiola First Report at para 125); and,

- Advertisers authorize Google to perform an automatic reduction of their bid (Stricchiola First Report at para 127).

[283] Based upon my construction of claim 1 of the 167 Patent, it does not disclose an auction mechanism. It discloses: retrieving a webpage, entering a target keyword and the amount including the maximum amount, authorizing an automatic reduction, and generating reports. The claim 1 steps are directed to an advertiser. Bid collapsing is the target of claim.

[284] Google's evidence is that Google Ads does not report on any individual bidding results to advertisers. As well, Google Ads is premised upon an auction format but has a different focus from Claim 1. Google Ads focuses on searchers of the internet. The 167 Patent focuses on advertisers.

[285] Finally, Google Ads auction does not perform the function of reducing bid amounts (collapsing) and it does not report to advertisers any reductions in bid amounts. These steps are essential elements of Claim 1 of the 167 Patent.

Claim 17

[286] Ms. Stricchiola claims that claim 17 is infringed by Google at para 150 of her First Report as follows: "Google Ads allows advertisers to authorize the Google Ads system to automatically perform offer adjustments that increase or decrease an advertiser's offer based on certain criteria […]".

[287] Claim 17 is a method claim with three steps: receiving authorization, monitoring keyword offers and, implementing an identified change. I construed the phrase "offeror's offer" to mean the "bidders bid". I did not construe claim 17 as calling out the automatic adjustment of bids as the word automatic is not used in claim 17, unlike claim 1, which specifically calls out the automatic function. Claim 17 operates on a first price auction (such as GoTo).

[288] The evidence is that Google Ads do not operate on a first price auction. Further, Google's automated bidding system, as explained by Mr. Moser, does not include the element of monitoring of bids. I accept the evidence of Mr. Moser on this point over the generalizations and assumptions of Ms. Stricchiola.

[289] Based upon the construction of claim 17, monitoring is an essential element. Without it, there is no infringement.

[290] As Google Ads does not monitor bids, I conclude that there is no evidence of infringement with respect to claim 17.

Claims 28, 44, 59, 75

[291] These claims were added to the 167 Patent many years after the patent was initially filed. This presumably accounts for the fact, as acknowledged by Ms. Stricchiola, Dr. Haruvy and Dr. Parkes, that these claims differ from claims 1-27. However, the infringement analysis will be done regardless of the motivation or reason for the late addition of these claims.

[292] In her infringement analysis, Ms. Stricchiola contends that "statistics of relevance" and "statistics" referred to in claims 28, 44, 59 and 75 are equivalent to Google's CTR or pCTR.

[293] I have construed the references in the 167 Patent to statistics and statistics of relevance as being statistics in relation to keywords only, and providing information for example on the frequency of searches of a keyword, or how often an advertisement was displayed.

[294] CTRs are explained in a Google document relied upon by Ms. Stricchiola titled *AdWords Basics* as a way to get an ad to show in a higher position. As the document describes: "The Clickthrough Rate component of the ordering rules rewards advertisers that have well-targeted ads that are appealing and relevant to searchers" (Exhibit 39).

[295] The evidence establishes that Google ads systems have always focused on the quality of the advertisements as one of the key metrics assessed. Although Ms. Stricchiola attempts to draw a correlation between "statistics" as used in the 167 Patent and Google ad quality assessments, I do not agree that such a correlation can be made. The 167 Patent was not concerned with the advertisers' advertisements – it was concerned with the advertisers'

keywords and bids. It does not call out any steps or methods to assess or consider the nature or type of advertisement sought to be displayed or promoted.

[296] I accept Dr. Parkes' opinion on this issue as noted in his second report at para 284:

> The Skilled Person would not understand the click through rate (CTR) of an advertisement to be a "statistic" of claim 28 in the context of the 167 Patent. All of the "statistics" disclosed in the 167 Patent relate to the keywords, or the number of impressions of a hyperlink. The Skilled Person would not understand CTR to be a statistic "related to" impressions as claimed by Ms. Stricchiola. The Skilled Person would understand that statistics "related to" impressions would tell you something about the impressions (such as the number of impressions per week or month). This is not the case with CTR. Knowing the CTR of an advertisement does not tell you anything about the number of impressions of the advertisement (let alone of a hyperlink). The CTR of an advertisement is the number of clicks divided by the number of impressions, and without knowing the number of clicks, the CTR does not tell you anything about the number of impressions.

[297] Ms. Stricchiola attempts to draw a correlation between the statistics referred to in claims 28, 44, 59 and 75 and CTRs used by Google. However, as confirmed by Mr. Iljic, [redacted]. On this issue, I accept the evidence of Mr. Iljic's over the opinion of Ms. Stricchiola.

[298] I am satisfied that Google Ads does not consider statistics in the manner called out in claims 28, 44, 59 and 75 of the 167 Patent, therefore, there is no infringement of these claims or their dependent claims.

C. Conclusion – Infringement

[299] Based upon the above analysis, PSET has simply not met its evidentiary burden to establish infringement.

[300] In their infringement claims, PSET is claiming a second price auction, the use of ad quality, ranking of ads, and pricing the clicks on the ads. When this claim is considered against the 167 Patent, there are no claims or embodiments that capture these concepts or support such an interpretation. In fact, during cross-examination, the inventors candidly admitted under questioning that they did not know how a second price auction worked.

[301] In their closing submissions, PSET raised the issue of inducement. As this ground was not pleaded by PSET in their Statement of Claim, it is not appropriate to consider this as a ground to support their infringement argument. In any event, as infringement has not been established, it does not require further consideration.

[302] Overall, I conclude that even if the asserted claims were valid, there is no evidence to support the infringement allegations.

VIII. Territory

[303] Google raised the issue of whether its activities – which occur outside of Canada, but relate to online services – can result in the infringement of a Canadian patent. They argue that

patents are territorial and Canadian patents cannot be infringed outside of Canada. Given my findings above it is not necessary for me to address this issue.

IX. Remedies

[304] As I have concluded that all of the asserted claims in issue are invalid and there is no infringement, it is not necessary to address the issue of remedies and I decline to do so.

X. Conclusion

[305] PSET's action will be dismissed with costs payable to Google. If the parties are unable to agree on costs they can make written submissions to the Court within 30 days of the date of this Judgment.

[306] These Reasons are being issued to the parties on a confidential basis. They have 15 days after the date hereof to advise the Court whether there is a requirement to redact any confidential information.

JUDGMENT IN T-40-18

THIS COURT'S JUDGMENT IS that:

1. The action is dismissed.

2. The counterclaim is granted: Claims 1 and 17 and the dependent claims are invalid as they are anticipated or obvious; Claims 28, 44, 59, and 75 and the dependent claims and Claim 9 are invalid for lack of sufficiency.

3. The defendants have not infringed any of the 167 Patent claims.

4. Google is entitled to its costs.

5. If the parties are unable to agree upon costs, they may make written submissions not exceeding 10 pages within 30 days of the date of this Judgment. Responding submissions not exceeding 5 pages may be made within 10 days thereafter.

"Ann Marie McDonald"
Judge

FEDERAL COURT

SOLICITORS OF RECORD

DOCKET:	T-40-18
STYLE OF CAUSE:	PAID SEARCH ENGINE TOOLS, LLC v GOOGLE CANADA CORPORATION, GOOGLE LLC AND ALPHABET INC.
PLACE OF HEARING:	HEARD VIA VIDEOCONFERENCE
DATES OF HEARING:	JUNE 14-18, 21-25, 28-30; JULY 28-30, 2021
JUDGMENT AND REASONS:	MCDONALD J.
CONFIDENTIAL JUDGMENT AND REASONS ISSUED:	DECEMBER 17, 2021
PUBLIC JUDGMENT AND REASONS ISSUED:	JANUARY 12, 2022

APPEARANCES:

Marcus Klee, Scott Beeser, Jonathan Stainsby, Devin Doyle, Jonathan Giraldi, Bryan Norrie	FOR THE PLAINTIFF
Christopher Van Barr, Michael Crichton, Marc Richard, Charlotte McDonald, Alex Gloor, Marc Crandall, Natalia Thawe	FOR THE DEFENDANTS

SOLICITORS OF RECORD:

AITKEN KLEE LLP Ottawa, Ontario	FOR THE PLAINTIFF
Gowling WLG (Canada) LLP Ottawa, Ontario	FOR THE DEFENDANTS

Exhibit 3 to the Cross-Examination of Professor Tadelis

CARTELS AS RATIONAL BUSINESS STRATEGY: CRIME PAYS

John M. Connor & Robert H. Lande†

TABLE OF CONTENTS

† The authors are, respectively, Professor Emeritus at Purdue University, and Senior Research Fellow, American Antitrust Institute; and Venable Professor of Law at the University of Baltimore School of Law, and Director, American Antitrust Institute. The authors are grateful to Joshua Davis, Albert Foer, Jay Himes, John Kirkwood, Dionne Koller, Christopher Sagers, Daniel Sokol, Andreas Stephan, Robert Steiner, Maurice Stucke and the participants at the Loyola Antitrust Colloquium for extremely helpful suggestions, and to James Denvil, Gary Stapleton, and Thomas Weaver for valuable research assistance.

INTRODUCTION

Cartels have always been the highest concern of antitrust. They overcharge consumers many billions of dollars every year[1] and there is a strong consensus that they should be sanctioned heavily.[2] Yet, until now no one has ever seriously attempted to analyze whether cartel sanctions are at the optimal level. This Article is the first to undertake this formidable task. Surprisingly, it demonstrates that the combined level of U.S. cartel sanctions has been only 9% to 21% as large as it should be to protect potential victims of cartelization optimally. This means that the average level of U.S. anti-cartel sanctions should be quintupled.[3]

Until now, no comprehensive empirical study has attempted to analyze whether cartels have been sanctioned optimally because of data constraints and the complexity and number of factors involved. The United States imposes a wide variety of sanctions against those who collude. These include criminal fines for the firms involved, prison, house arrest, and fines for the corporate officials involved.[4] Victims can sue for mandatory treble damages and attorney's fees.[5] Judge Posner called this combination of sanctions the equivalent of dropping "cluster bombs" on defendants.[6] This multiplicity has led to the common—but unsupported—belief that the current level of sanctions is adequate[7] or excessive.[8]

---

1 *See infra* Part III.A.

2 Strong anti-cartel policies are not only on the agenda of progressives; most conservatives advocate sanctioning cartels heavily. *See, e.g.*, Frank A. Easterbrook, *Treble What?*, 55 ANTITRUST L.J. 95, 95 (1986). In 2004, the Bush Administration proposed and helped enact significant increases in the criminal fines against cartels. *See* Antitrust Criminal Penalty Enhancement and Reform Act of 2004, Pub. L. No. 108-237, 118 Stat. 661, 665–68 (substituting a $100 million maximum corporate fine for the existing $10 million maximum; a maximum $1 million individual fine for the existing $350,000 maximum; and a maximum ten year prison sentence for the existing maximum three year sentence).

3 Another option would be to implement ways to vastly improve the cartel detection rate. For an analysis of a number of alternatives, see *infra* Conclusions, Section A.

4 *Id.* There also are such relatively unusual or minor sanctions as disgorgement actions by the Federal Trade Commission (FTC) or the Department of Justice (DOJ). Although individual disgorgement cases can be important, they are relatively rare. *See* Einer Elhauge, *Disgorgement as an Antitrust Remedy*, 76 ANTITRUST L.J. 79, 79 (2009).

5 *See* 15 U.S.C. § 15 (2000). Prevailing plaintiffs also receive filing fees and expert witness fees. *Id.*

6 Richard A. Posner, *Antitrust in the New Economy*, 68 ANTITRUST L.J. 925, 940 (2001) [hereinafter Posner, *Antitrust*]. *See generally* Spencer Weber Waller, *The Incoherence of Punishment in Antitrust*, 78 CHI.-KENT L. REV. 207 (2003).

7 The ABA Antitrust Section, for example, recently opposed increasing the Sherman Act's criminal penalties: "Some also believe that combined criminal and civil penalties provide too much deterrence that will chill the businessperson in his decision making . . . . Whether increased criminal penalties will provide an appropriate level of deterrence . . . should be the subject of hearings and public briefings to reach the proper deterrence balance." SECTION OF ANTITRUST LAW, AM. BAR ASS'N, COMMENTS OF THE ABA SECTION OF ANTITRUST LAW ON H.R. 1086: INCREASED CRIMINAL PENALTIES, LENIENCY DETREBLING AND THE TUNNEY ACT

This Article employs a unique database to determine whether the United States' anti-cartel sanctions are optimal overall. It does this by analyzing the total, combined impact of every measurable anti-cartel sanction using the standard optimal deterrence approach.[9] This assumes corporations and individuals contemplating illegal collusion will be deterred only if the expected rewards are less than the expected costs[10] divided by the probability the illegal activity will be detected and sanctioned.[11]

AMENDMENT 11–12 (2004), *available at* http://www.americanbar.org/content/dam/aba/administrative/antitrust_law/comments_increasedcriminalpenalties.authcheckdam.pdf.

8 This view was eloquently articulated by Professors Lopatka & Page even before the criminal fine levels were significantly increased in 2004: "Even setting imprisonment aside, the federal criminal penalties are substantial. . . . [and] today may well be high enough that the optimal penalty can be imposed through criminal sanctions alone. . . . It seems likely that the combination of federal penalties is adequate." John E. Lopatka & William H. Page, *Indirect Purchaser Suits and the Consumer Interest*, 48 ANTITRUST BULL. 531, 568 (2003) (footnote omitted); *see also* ABBOTT B. LIPSKY, LATHAM & WATKINS, LLP, PRIVATE DAMAGE REMEDIES: TREBLE DAMAGES, FEE SHIFTING, PREJUDGMENT INTEREST 4–5 (2005), *available at* http://govinfo.library.unt.edu/amc/commission_hearings/pdf/Lipsky.pdf (statement to the Antitrust Modernization Commission) ("[S]o long as Section 1 and Section 2 violations can be—and in the case of cartel violations, typically are—prosecuted criminally and punished with actual incarceration for individuals and criminal fines. . . . [i]t is possible that the treble-damage claims unintentionally assume some of the characteristics of a wealth-transfer program . . . [similar to] the retributive and unwise legal methods that produced or at least inflamed the Salem Witch Trials . . . ."); *Criminal Remedies: Public Hearing Before the Antitrust Modernization Comm'n*, at 83, Nov. 3, 2005, *available at* http://govinfo.library.unt.edu/amc/commission_hearings/pdf/051103_Transcript_Criminal_Remedies.pdf (statement of Anthony V. Nanni, former Chief of the National Criminal Enforcement Section in the Antitrust Division, U.S. Department of Justice) ("[W]hen you have such large corporate fines combined with the other framework—i.e., civil treble damages—you really run the risk of pushing corporations to the brink of bankruptcy.").

9 See *infra* notes 15–22 for an explanation of the standard optimal deterrence approach. As explained throughout this paper, including in notes 28 and 32 *infra*, we believe this Article's analysis is best carried out in relatively traditional, non-behavioralist terms. Some of the remedies we propose, however, fairly might be termed "behavioralist." *See infra* Part V.A.

10 Optimal deterrence depends upon the rational conjectures or expectations of potential cartelists as to a number of factors when a cartel is being formed. Ideally, one would like to know how much would-be cartel managers or their employers expect to gain from their collusion, how likely it is they think they will be apprehended, and how large a corporate fine and how long a prison term they believe the managers and their employers will receive should they be caught. Managers may be carrying out a corporate decision, or they may be rogues. What goes on in the minds of potential cartelists is largely unexplored in the cartel literature (but for insights on this issue, see Michael O'Kane, *Does Prison Work for Cartelists?: The View from Behind Bars*, 56 ANTITRUST BULL. 483 (2011)). We only can estimate how much discovered cartels have gained in the past, what the historical rate of discovery and conviction likely has been, and how heavily corporate participants and their employees have been sanctioned. We then assume the historical outcomes match the cartelists' expectations—an admittedly rough approximation. See *infra* Part I.A for a more thorough discussion.

11 In other words, a sanction slightly larger than $300 would be necessary if a cartel expects total overcharges to reach $100 and believes there is a 1/3 chance its activities will be detected and condemned. In operational terms, the optimal penalty will be assumed to be equal to (the cartel's overcharges) ÷ (the probability the cartel will be detected × the probability the detected collusion will be sanctioned).

Our analysis begins with calculations of the rewards from collusion in a sample of seventy-five cartel cases. We then survey the literature to ascertain the probability that cartels are detected and sanctioned. We further assemble data on the size of the sanctions involved in each case in our sample. These include the corporate fines, individual fines, and payouts in private damage actions for these cartels. Finally, we determine the opportunity cost (or disvalue) of imprisonment or house arrest for the individuals convicted in these seventy-five cases.[12]

Our optimal deterrence analysis[13] concludes that the combined level of U.S. cartel sanctions has been only 9% to 21% as large as it should be to protect potential victims of cartelization optimally. Hence, despite all the existing sanctions, collusion remains a rational business strategy. Cartels are a crime that, on average, pays. In fact, it pays very well.

This Article proceeds in six Parts. Part I analyzes the optimal deterrence of cartels, including separate discussions of the necessary individual, as well as corporate perspectives and incentives. Part II analyzes the sizes of cartel sanctions in our sample of seventy-five cases: payments made in private damages actions, corporate fines, individual fines, restitution payments, and the monetary equivalents of imprisonment and house arrest for corporate officers engaged in collusion. Part III summarizes the field's empirical knowledge about the harms to society from collusion. Part IV ascertains the probability a cartel will be discovered and sanctioned. Part V combines the previously calculated figures, for our sample of seventy-five cartel cases, to produce our results.

This Article's results should be of paramount importance to anyone interested in protecting the public against collusion. Accordingly a sixth, concluding section will discuss the implications of our research for public policies towards cartels. Because current cartel sanctions are far too low, we suggest specific ways they could be increased to become more nearly optimal. Doing so would save consumers billions of dollars each year.

---

12 It is of course impossible to equate incarceration and monetary sanctions in an objective manner since this would mean computing the "value" or "cost" of time spent in prison or under house arrest. Nevertheless, this Article will examine several social science approximations of the disutility of prison time and house arrest, ascertaining and combining many different estimates in a conservative manner. *See infra* Part I.B. Consequently, the Article's overall assessment of the impact of incarceration will be both as accurate and non-controversial as possible.

13 As explained throughout this Article, we use the best available data for each part of the optimal deterrence calculation. Some information is known with certainty, but some of the required information is not available with as much precision or the degree of confidence we would like. In recognition of these imprecisions, we undertake a sensitivity analysis: We determine the highest and lowest likely values for each relevant factor and combine them into appropriate low and high estimates of the overall optimal deterrence tradeoff.

I. OPTIMAL DETERRENCE: INDIVIDUAL VS. CORPORATE PERSPECTIVES[14]

How can cartels best be deterred? Should sanctions focus upon corporations, individuals, or both? How large should each category of sanctions be relative to the harms from collusion?

A. *Overall Framework for Analysis*

The generally accepted overall approach to the optimal deterrence of antitrust violations was developed by Professor William Landes.[15] He showed that to achieve optimal[16] deterrence[17] the damages from an

14 This Part relies heavily upon and significantly extends some of the authors' earlier joint work. *See* John M. Connor & Robert H. Lande, *How High Do Cartels Raise Prices? Implications for Reform of the Antitrust Sentencing Guidelines*, 80 TULANE L. REV. 513 (2005), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=787907. This Part also relies upon John M. Connor, *Problems with Prison in International Cartel Cases*, 56 ANTITRUST BULL. 311 (2011), and Robert H. Lande & Joshua P. Davis, *Comparative Deterrence from Private Enforcement and Criminal Enforcement of the U.S. Antitrust Laws*, 2011 BYU L. REV. 315, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1565693.

15 William M. Landes, *Optimal Sanctions for Antitrust Violations*, 50 U. CHI. L. REV. 652, 656 (1983) (adapting Gary Becker's well known "theory of crime" to examine price-fixing violations that are nearly always prosecuted as felony crimes by the DOJ; for that reason, the ex ante approach to analyzing crimes is dubbed "Beckerian"). By the early 1990s, the Beckerian formulation of the problem of policies designed to deter hard-core price-fixing violations had been adopted universally by legal-economic scholars. *See* Richard A. Posner, *Optimal Sentences for White-Collar Criminals*, 17 AM. CRIM. L. REV. 409 (1979–1980) [hereinafter Posner, *Optimal Sentences*]. In addition, an alternative analysis of optimal anti-cartel policies has grown during the last decade. *See, e.g.*, Paulo Buccirossi & Giancarlo Spagnolo, *Optimal Fines in the Era of Whistleblowers: Should Price Fixers Still Go to Prison?*, *in* THE POLITICAL ECONOMY OF ANTITRUST 81 (Vivek Ghosal & Johan Stennek 2007). This newer perspective on enforcement focuses on policies like corporate or individual leniency programs that may destabilize cartels that are already formed. Thus, we view policy prescriptions arising from this body of scholarship as ex post and, far from being contradictory, as supplementary to the ex ante policies we examine in the present Article.

16 One might quite reasonably reason that, unlike the case for conduct that might violate the prohibitions against illegal monopolization, because price fixing is never in the public interest, we should attempt to design a regime that prevents all price fixing, not a regime that permits some "optimal" amount of price fixing. One might argue that we should not worry about imposing excessive penalties against cartels.

Our quest should not be complete deterrence, however, because enforcement aggressive enough to deter all cartels almost certainly would penalize and therefore discourage some honest business conduct. As with any legal system, there is some uncertainty at the margin of cartel illegality. Beneficial horizontal conduct near this line, conduct that results in efficiency gains for society, sometimes could be mistaken for illegal collusion. For this and other reasons sanctions should not be excessive; they should only be as large as necessary to deter most of the undesirable conduct. To give an extreme example, a mandatory death penalty for price fixing, if regularly imposed, surely would chill a significant amount of procompetitive behavior because most people quite understandably would avoid doing anything that could give rise to even a small probability of being mistaken for price fixing.

17 Professor Landes was not concerned with the compensation of victims. Landes, *supra* note 15. For an analysis that takes compensation into account, see Robert H. Lande, *Are*

antitrust violation should be equal to the violation's expected "net harm to others"[18] divided by the probability of detection and proof of the violation.[19] All figures should, of course, be expressed in constant dollars. Most analysts of both the Chicago and post-Chicago schools of antitrust have accepted these principles.[20] The "net harm to others" from collusion, of course, includes the overcharges that result from

Antitrust "Treble" Damages Really Single Damages, 54 OHIO ST. L.J. 115, 161–68 (1993), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1134822.

18 The logic underlying the "net harm to others" standard was explained clearly by Professors Breit and Elzinga. Their example is that of a horizontal cartel. However, in their example, the activity also produces a significant efficiency gain. Sometimes horizontal activity that produces a significant efficiency gain is labeled a "joint venture" rather than a "cartel." Other times "cartel" is simply a shorthand for horizontal activity that produces more losses than gains.

> The trick to discovering the optimal sanction is to find a rule that will force the potential cartelist to compare any cost saving from his activity with the deadweight loss triangle. If the cost saving were larger than the deadweight loss, it would be in his (and society's) interest to undertake the illegal activity. So after he deducts the monopoly profit rectangle . . . the cartelist will examine the deadweight loss (the remainder of the fine to be paid) and compare it with the value of the cost saving. The fine that is the sum of the deadweight triangle plus the profit rectangle is the correct sanction since it will encourage the "right" amount of illegal antitrust activity. Damages larger than this could lead to over-deterrence . . . .
>
> A numerical example may help to clarify the concept of the optimal antitrust sanction. Assume that a potential cartelist calculates that joining a horizontal price-fixing conspiracy will increase his profits by $100 million. He also is aware that the deadweight loss imposed on society by his activity is $50 million. If the expected value of the fine imposed is the entire amount of consumers' surplus ($150 million) would he enter the cartel? He would do so if he believed that the cartel would be accompanied by cost reductions to him greater than $50 million. If the cost saving were, say, $60 million, he would still enter the price-fixing conspiracy because he would know that his fine would be $100 million (his cartel profits) plus $50 million (the deadweight loss) leaving him $10 million more revenue than would be the case if he did not enter the cartel. In this case the cartel is accompanied by cost reductions greater than the deadweight loss it imposes on society. On efficiency grounds, it should be permitted.

WILLIAM BREIT & KENNETH G. ELZINGA, ANTITRUST PENALTY REFORM: AN ECONOMIC ANALYSIS 11–12 (1986).

19 *See* Landes, *supra* note 15, at 666–68. Thus, if the harm were 10 and the probability of detection and proof were .33, since 10/.33 = 30, the optimal penalty for this violation would be 30. This assumes risk neutrality and other common assumptions. *Id.*

20 See the discussion in Lande, *supra* note 17, at 161–68. Despite the general acknowledgement of the superiority of the Landes approach, however, many respected scholars and enforcers instead focus upon the gain to the lawbreakers, perhaps because it is simpler to observe or calculate. For a recent example see Gregory J. Werden, *Sanctioning Cartel Activity: Let the Punishment Fit the Crime*, 5 EUR. COMPETITION J. 19, 28–31 (2009). For an insightful analysis see Wouter P.J. Wils, *Optimal Antitrust Fines: Theory and Practice*, 29 WORLD COMPETITION 183, 190–93 (2006). For this Article's purposes, however, the precise optimal deterrence standard used is not crucial. Similar results would arise if this Article instead used a "gross harm to others" or a "net gain to the offenders" standard.

cartel pricing.[21] They include many other—perhaps less obvious—factors, as well.[22]

Moreover, since not every cartel is detected or successfully proven, the "net harm to others" should be multiplied by the inverse of the probability of detection and proof.[23] The Antitrust Division's amnesty program has resulted in a significantly larger percentage of cartels detected and proven in recent years.[24] Nevertheless, there is continuing evidence that, despite the enforcers' superb efforts, many cartels still operate,[25] so there is significantly less than a 100% probability that a

21 *See* Landes, *supra* note 15.

22 First, cartel market power produces allocative inefficiency—the deadweight loss welfare triangle. *See* EDWIN MANSFIELD, MICROECONOMICS: THEORY AND APPLICATIONS 277–92 (4th ed. 1982) (defining allocative inefficiency and providing a proof that it is created by monopoly pricing). Allocative inefficiency often is significant empirically. *See* discussion *infra* Part III.B. Nevertheless, it apparently has never been awarded in an antitrust case. *See, e.g.*, David C. Hjelmfelt & Channing D. Strother, Jr., *Antitrust Damages for Consumer Welfare Loss*, 39 CLEV. ST. L. REV. 505 (1991).

Second, market power can produce "umbrella" effects, the name given to higher prices charged by non-violating members that were permitted or caused by the violation's supracompetitive prices. *See* PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW ¶ 337.3 (Supp. 1992). This factor also is never or virtually never awarded. *Id.*

Moreover, there are several additional types of harms that often are caused by cartels. These include: 1) uncompensated plaintiffs' attorneys' fees and costs; 2) the uncompensated value of plaintiffs' time spent pursuing the case; and 3) the costs of the judicial system. *See* Lande, *supra* note 17, at 129–58.

In addition, cartels may have less incentive to innovate or to offer as wide an array of non-price variety or quality options. Alternatively, one could argue that cartel members will have more funds to use for socially desirable innovation. We know of no evidence, however, that these innovation effects are significant empirically.

The price fixers' own legal costs, the disruption in their own efficiency as a result of sanctions litigation, and any harm to their corporate reputation, by contrast, are not "harms to others" from collusion, and therefore should not be included in the optimal deterrence analysis.

23 "Multiplication is essential to create optimal incentives for would-be violators when unlawful acts are not certain to be prosecuted successfully. Indeed, some multiplication is necessary even when most of the liability-creating acts are open and notorious. The defendants may be able to conceal facts that are essential to liability." *See* Frank Easterbrook, *Detrebling Antitrust Damages*, 28 J.L. & ECON. 445, 455 (1985).

24 *See* Nathan H. Miller, *Strategic Leniency and Cartel Enforcement*, 99 AM. ECON. REV. 750 (2009).

25 *See generally* Douglas H. Ginsburg & Joshua D. Wright, *Antitrust Sanctions*, 6 COMPETITION POL'Y INT'L 3 (2010). The continued high number of DOJ grand juries and the recent DOJ success rate in the courts also suggests that many cartels still exist. As of the close of fiscal year 2010 the DOJ had approximately 124 pending grand jury investigations. U.S. DEP'T OF JUSTICE, ANTITRUST DIVISION WORKLOAD STATISTICS FY 2002–2011, at 4, [hereinafter WORKLOAD STATISTICS 2002–2011] *available at* http://www.justice.gov/atr/public/workload-statistics.html. Between 2001 and 2010, the DOJ filed from forty-four to sixty criminal cases per

cartel will be detected and convicted. From an optimal deterrence perspective, sanctions should be more than a cartel's "net harms to others" to account for the probability that the conduct will go unpunished. As noted earlier, if a cartel that expected to overcharge by $100 only faced a 33% chance it would be detected and proven to be illegal, the sanctions should slightly exceed $300. Without this multiplier firms would be simply undeterred from committing antitrust violations.

Ideally, optimal deterrence should be based upon the expectations of potential price fixers, not the results of others' past price fixing or the sanctions imposed on similar cartels.[26] The required expectation knowledge, however, is impossible to obtain.[27] Guessing what goes on in

year, most of which resulted in convictions. *Id.* at 4. The following table, extracted from this data, shows DOJ's success in prosecuting antitrust violations:

Total Criminal Cases	'01	'02	'03	'04	'05	'06	'07	'08	'09	'10
Filed	44	33	41	42	32	34	40	54	72	60
Won	38	37	32	35	36	31	31	47	67	41
Lost	2	1	1	1	1	–	1	4	2	1
Pending	39	34	42	48	43	44	54	57	60	55
Appeal Decisions	5	1	2	7	4	5	1	4	2	7
Grand Juries Initiated	26	26	48	21	38	38	34	32	38	12

In the opinions of a large number of judges, grand juries, and juries, the DOJ Antitrust Division has been bringing a large number of meritorious anti-cartel cases in recent years. Note that in some years the DOJ won more cases than it filed because the cases the DOJ won in any given year were often filed in an earlier year.

[26] It would be extremely useful to know potential price fixers' perceptions of the probability that they will be caught and convicted of price fixing, and their belief as to how much they will be forced to pay. Moreover, as one distinguished cartel scholar noted, "[b]ecause of overconfidence bias, prospective offenders are likely to overestimate the gain and underestimate the probability of detection and punishment." *See* Wils, *supra* note 20, at 183.

We know of no reliable information on this issue, however. Their expectations will, to some degree, be informed by their discussions with their antitrust lawyers, but there still could well be systematic differences between their expectations and reality. In addition, potential price fixers probably are likely to be risk seekers, and have other relevant psychological traits on the average. Moreover, there could be a difference between how much potential price fixers think they would be likely to earn from price fixing, and the amount a court or an economist measures after the fact. Similarly, there could be a difference between reality and their estimate, at the time of the price fixing, of the probability they will get caught and convicted, and their expectation as to how much the negotiated fine will be. In addition, optimal deterrence theory is based on the balance between the present value of expected future corporate profits from the conduct and the present value of expected future monetary sanctions.

[27] To ascertain this, one would have to interview a random sample of potential price fixers and discern their expectations. In reality, however, it would be impossible to assemble a proper random sample or to get them to respond candidly. A different way to frame the optimal deterrence issue is in terms of whether cartels usually know in advance of litigation roughly how much they will be found to have overcharged. Can most firms that are members of cartels

the minds of would-be cartelists is hazardous. Nor do we know how often potential price fixers consult with their attorneys about the likely range of outcomes.[28] The best we can do is to ascertain how much overall (in terms of a median or a mean) cartels have raised prices in the past, and how often and how much they have been sanctioned, and assume these are close proxies for the expectations relevant to the decision whether to collude.[29] In effect, we are using a general deterrence approach because a specific deterrence approach is infeasible.

B. *Corporate vs. Individual Sanctions*

Even though the preceding analysis is accepted by most of the antitrust field with relatively little controversy, it does not answer the question posed at the start of this Section: Is optimal deterrence best achieved by focusing only on the corporations involved?[30] On the individuals involved? And if so, should this be done by fines or through incarceration? Or through some combination of corporate and individual sanctions?[31]

---

predict in advance of litigation, for example, that a court will find that it overcharged 5%, as opposed to 15%?

In light of the probability that lengthy, protracted litigation could result in a high, or low, sanction result, another issue is how risk seeking or averse a particular corporation is.

More generally, one might argue that our use of the standard optimal deterrence model (which assumes risk neutrality) for entire cartels is inappropriate. After all, if the most risk-averse member of a cartel decides to turn in the cartel, the entire cartel will end. (This idea is not applicable at the decision to participate stage, however, because a cartel need not contain every firm within an industry to be largely successful.) For this reason the optimal deterrence target need only be the most risk-averse member of a cartel. It seems likely, however, that most cartelists are by nature risk seekers. Accordingly, the appropriate focus of an optimal deterrence calculation actually should be on the most risk-averse member of a group of risk seeking cartelists. Is this person/corporation net risk-neutral, net a risk avoider, or still a net risk seeker? We do not know. Experimental economics offer some promise of modeling choices of participants in cartel settings. However, to our knowledge no relevant experiments have been published on this issue.

28 *See generally* D. Daniel Sokol, *Cartels, Corporate Compliance, and What Practitioners Really Think About Enforcement*, 78 ANTITRUST L.J. 201 (2012).

29 For this reason, we readily acknowledge that we are administering an imperfect test using a surrogate for what we really would like to measure.

30 This Section draws heavily upon material in Connor, *supra* note 14, and in Lande & Davis, *supra* note 14.

31 One could attempt to analyze whether sanctions should be imposed on individuals and/or on corporations, and other issues examined in this Article, using a more explicitly behavioral approach. For an excellent behavioral analysis of related issues concerning collusion, see generally Maurice Stucke, *Am I a Price Fixer? A Behavioral Economics Analysis of Cartels*, *in* CRIMINALISING CARTELS: A CRITICAL INTERDISCIPLINARY STUDY OF AN INTERNATIONAL REGULATORY MOVEMENT 263 (Caron Beaton-Wells & Ariel Ezrachi eds., 2011), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1535720.

In light of this Article's conclusion that current cartel sanctions are significantly suboptimal, however, a more explicitly behavioral approach would not significantly enhance our analysis. Our analysis shows that current sanctions are much less than they should be to

Consideration of optimal sanctions for price fixing can be traced to Richard Posner's analysis of optimal cartel penalties.[32] According to this work, hard-core price fixing is optimally punished almost exclusively through corporate fines.[33] Only when a company is unable to pay an optimal fine should imprisonment be imposed as a last resort, and only if the individuals are unable to pay optimal fines.[34]

There are many arguments in favor of the criminalization of price-fixing offenses.[35] For example, publicity about severe sentences for price fixing may help educate other corporate executives about the true individual and corporate legal risks of being caught.[36] Publicity may also contribute to the effectiveness and costs of corporate antitrust compliance programs. Imprisonment could improve the operation of public antitrust leniency programs because, by shifting corporate officers' expectations toward high personal penalties, top executives of cartel participants are more likely to seek the immunity from

deter cartels optimally, so it is unsurprising that firms contemplating collusion do so rationally and knowingly. It is in their self interest to collude, so the explanation as to why they attempt to form cartels is relatively simple and straightforward.

On the other hand, behavioral issues would be extremely important if the overall level of sanctions were optimal or super-optimal. Under these conditions one would have to explain why corporations continue to engage in the seemingly irrational behavior of illegal collusion. Under these circumstances, one should analyze, for example, issues such as whether managers who are worried about getting fired for poor performance have an incentive to defy top management's instruction not to engage in collusion by entering into a cartel with their competitors. If sanctions were optimal or super-optimal, a behavioral analysis could help decide how to stop this from happening. In light of this Article's conclusions that sanctions currently are too low, however, no such analysis is necessary.

By contrast, many of our proposed solutions could be termed "behavioral." *See infra* Part V.A.

32 Posner, *Optimal Sentences*, *supra* note 15.

33 *Id.* The conventional wisdom in the field was well summarized in V.S. Khanna, *Corporate Criminal Liability: What Purpose Does It Serve?*, 109 HARV. L. REV. 1477 (1996) ("Thus, some justification for corporate criminal liability might have existed in the past, when civil enforcement techniques were not well developed, but from a deterrence perspective, very little now supports the continued imposition of criminal rather than civil liability on corporations.").

34 Posner, *Optimal Sentences*, *supra* note 15. Posner argued for "the substitution, whenever possible, of the fine (or civil penalty) for the prison sentence as the punishment for crime." *Id.* at 409. Posner also acknowledged that he has made "an argument . . . in the antitrust context for confining criminal (or civil-penalty) liability to the corporation, on the theory that if it is liable it will find adequate ways of imposing on its employees the costs to it of violating the law." *Id.* at 417–18. He observed: "The fine [or civil liability] for a white-collar crime can be set at whatever level imposes the same disutility on the defendant, and thus yield the same deterrence, as the prison sentence that would have been imposed instead." *Id.* at 410. Yet the fines would save the cost to society of incarcerating the lawbreakers, and also, the opportunity cost to society of the time they spend in prison instead of working productively. Posner is familiar with resistance to this claim—indeed, his Article responds in part to a criticism that contends that the threat of imprisonment is inherently greater than that of a fine. *Id.* at 413.

35 See the sources cited in Connor, *supra* note 14, for a summary of the legal-economic arguments for and against individual criminal penalties for antitrust violations, including the available game theory arguments.

36 *See infra* note 45 (the example of Alfred Taubman).

prosecution that accompanies awards of corporate amnesty. In addition, public fines on employees can be socially optimal if principal-agent problems exist such that employees fail to take enough care to avoid legal risks for the corporation and the employer is unable to impose a financial penalty as high as the required public fine.

Indeed, one could argue in the extreme that sanctions should focus mainly or exclusively upon individuals. Officials at the U.S. Department of Justice (DOJ) Antitrust Division have been moving in this direction in recent years,[37] as have some of the most respected members of the antitrust community, such as Judge Douglas Ginsburg and Professor Joshua Wright, who advocates lengthy debarment for negligent corporate officers and directors of publicly traded companies that fix prices.[38]

---

37 For example, a 2006 speech by Scott Hammond contains a statement about the Division's belief that the threat of imprisonment overshadows all other sanctions as a cause of corporate leniency applications:

> It is indisputable that the most effective deterrent to cartel offenses is to impose jail sentences on the individuals who commit them. Corporations only commit cartel offenses through individuals, so executives as well as their employers need to be deterred from engaging in such conduct. Hard-core cartel offenses are premeditated offenses committed by highly educated executives. Before deciding whether to commit the offense, those executives weigh the risk and consequences of detection against the potential financial rewards of colluding. When an executive believes that incarceration is a possible consequence of engaging in cartel activity, he is far more likely to be deterred from committing the violation than if there is no individual exposure. This conclusion is not simply based on theories of human behavior or common sense. We have first-hand accounts from cartel members of how the presence or absence of individual sanctions has directly resulted in actual deterrence and continued competition in the U.S. market and failed deterrence, collusion, and great financial harm in foreign markets.
>
> We have uncovered international cartels that operated profitably and illegally in Europe, Asia, and elsewhere around the world, but did not expand their collusion to the United States solely because the executives decided it was not worth the risk of going to jail. I am referring to cartels that had every opportunity to target U.S. consumers. The cartel members sold in the U.S. market, and they were already getting together and fixing prices everywhere else they sold. Indeed, in some cases, the U.S. market was the largest and potentially most profitable, but the collusive conduct still ceased at the border. Why? The answer, from the mouths of the cartel members and verified by our investigators, is that the executives did not want to risk getting caught and going to jail in the United States.

Scott D. Hammond, Deputy Assistant Att'y Gen. for Criminal Enforcement, Antitrust Div., U.S. Dep't of Justice, Charting New Waters in International Cartel Prosecutions, Remarks at the National Institute on White Collar Crime (Mar. 2, 2006), *available at* http://www.justice.gov/atr/public/speeches/214861.htm.

38 *See* Ginsburg & Wright, *supra* note 25. Judge Ginsburg and Professor Wright certainly do not propose repealing corporate fines for price fixing. They do, however, advocate putting much more emphasis on individual sanctions. In particular, they propose lengthy debarment for negligent corporate officers and directors of publicly traded companies. Part of their preference for individual sanctions follows from their premise that the ever increasing levels of fines for price fixing have not sufficiently deterred collusion.

We certainly agree with Ginsburg and Wright that even though corporate fines have risen

The extreme form of this argument specifically rejects the logic of optimal deterrence principles. The dominant law-and-economics model of crime posits that rational choices drive corporate decisions (including the decisions of the individuals involved) to commit crimes—a "cost/benefit analysis" of the decision. Consequently, there exists a bundle of sanctions that the legal system can (at least in theory) calculate that optimally will deter the crime. Unless there are principal-agent problems,[39] the monetary values of these individual sanctions are, in principle, perfect substitutes for one another. [40]

There certainly are counter-arguments to the desire for vastly higher individual penalties for cartelization (indeed, the United States is the only nation, among the roughly 200 countries with anti-cartel laws, that incarcerates significant numbers of cartel managers).[41] Some have expressed skepticism about the effectiveness of individual sanctions in deterring antitrust crimes. An executive summary of a Policy Roundtable on this topic sponsored by the Oraganisation for Economic Co-operation and Development (OECD) asserted: "There is no

significantly in recently years, there still is significant under-deterrence of collusion. Ginsburg and Wright do not, however, analyze the possibility that even the current levels of corporate fines are insufficient to deter price fixing optimally. Despite the higher fines of recent years, if corporations still expect to make a profit from collusion, still higher corporate sanctions might lead to optimal deterrence.

39 If the firm is a proprietorship, it does not matter whether the sanctions fall upon the individuals or the corporation. But if there is a separation between ownership and management, the personal motives of managers must be considered in evaluating the effectiveness of sanctions. The simpler versions of optimal deterrence theory assume that there are no principal-agent divergences and that the managers are risk-neutral. However, it sometimes is true that the reward structures of traditional executive compensation contracts typically give short-term, personal enrichment a greater weight than the long-run interests of stockholders.

If the profits generated by price fixing generate immediate personal rewards for such managers, but long-term losses for shareholders (incurred only after years of litigation, when the managers may no longer be with the corporation) then the optimal ratio of sanctions to illegal profits must be higher than for a proprietorship. Similarly, a higher ratio will be required if managers are risk-seeking in their corporate decision making rather than risk-averse. For these reasons, our focus on corporate-level performance in the present paper is, at best, a rather imperfect surrogate for stockholder control, managerial risk aversion, and other factors that, if we were able to derive the necessary parameters, we would otherwise incorporate.

40 "The Division does *say* that it is focused on both hammering corporations with big fines and sending their price-fixing executives to jail. But the reality is that, despite vehement Division protestations to the contrary, a key element of the Division's enforcement approach appears to be a willingness to trade people (particularly senior executives) for money." TEFFT W. SMITH, KIRKLAND & ELLIS LLP, COMMENTS FOR THE ANTITRUST MODERNIZATION COMMISSION HEARING ON CRIMINAL ANTITRUST REMEDIES 5 (2005), *available at* http://govinfo.library.unt.edu/amc/commission_hearings/pdf/Smith_Statement.pdf.

41 The only other nations we know of that have imprisoned antitrust violators at least once are Great Britain, Israel, Germany, Japan, and Ireland, but they have only done so on relatively rare occasions. Canada and other jurisdictions impose prison sentences but convert them to non-custodial sanctions. *See* Connor, *supra* note 14. However, the international trend is towards greater use of incarceration for cartelists. *Id.*

systematic evidence proving the deterrent effects of sanctions on individuals, and/or assessing whether such sanctions can be justified."[42]

Moreover, an interesting set of criticisms was leveled at the DOJ's imprisonment policies at a hearing of the Antitrust Modernization Commission. Tefft Smith, a prominent U.S. antitrust lawyer who often represents defendants, testified that, in his experience, imprisonment is the DOJ's "biggest (and most effective) stick" in cartel enforcement.[43] Nevertheless, he criticized the DOJ for offering unduly short sentences[44] and because—with exceptions[45]—the DOJ tends to prosecute mid-level sales or marketing executives rather than the most senior responsible officers of the company.[46] To the extent this is true,[47] it seriously

42 ORG. FOR ECON. CO-OPERATION & DEV., *Overview to* POLICY ROUNDTABLES: CARTEL SANCTIONS AGAINST INDIVIDUALS, 2003, at 1 (2005) [hereinafter CARTEL SANCTIONS], *available at* http://www.oecd.org/daf/competition/cartelsandanti-competitiveagreements/34306028.pdf.

43 Smith, *supra* note 40, at 7–10.

44 *Id.*

45 For example, Alfred Taubman, the billionaire Chairman of Sotheby's, was sentenced to a year and a day in prison in conjunction with the auction houses bid rigging case. *See The World's Billionaires, #655 A. Alfred Taubman*, FORBES.COM (Mar. 10, 2010), http://www.forbes.com/lists/2010/10/billionaires-2010_A-Alfred-Taubman_LWZ4.html. Taubman "entered a low-security medical prison in Rochester, Minnesota, on August 1, 2002 and, after having his sentence reduced for good behavior, was released on May 15 2003." Jill Treanor, *Taubmans Lose Hold on Sotheby's: Auction House Ends Family's Grip on 62% of Voting Rights*, GUARDIAN, Sept. 9, 2005, at 18.

As of February 2011, Taubman was alive and doing well. The day after Taubman was sentenced, the Board of Directors of Sotheby's Holdings Inc. at a "thinly attended annual meeting" elected Taubman to be a member of the Board, and his son Robert replaced him as Chairman. *A Taubman Continues to Sit on Sotheby's Board*, NAT'L POST (CANADA), Apr. 25, 2002, at FP2. In addition to positions on other corporate boards, as of 2010 he was a Trustee of the Urban Land Institute. *Profile Detail—A. Alfred Taubman*, MARQUIS WHO'S WHO 2010, http://search.marquiswhoswho.com/profile/100004075742 (last visited Oct. 30, 2012) (registration required).

His re-emergence into society may have begun in Detroit in 2005, when he accepted the first lifetime achievement award from the Detroit chapter of the Urban Land Institute. *Taubman to be Honored*, CRAIN'S DETROIT BUS., Apr. 4, 2005, at 8. Moreover, his social life has revived. "Today we are living at the dawn of the ultra-mega-uber-monster book party, celebrations so huge and elaborate that you might think you were at a wedding . . . . In April, 400 guests celebrated Alfred Taubman's book, *Threshold Resistance: The Extraordinary Career of a Luxury Retailing Pioneer* . . . at the Four Seasons." Alex Kuczynski, *Comped Lit*, N.Y. TIMES, Aug. 26, 2007, § 6 (T: Women's Fashion Magazine), at 226. Louis Auchincloss, novelist and chronicler of New York City mores, was quoted as saying "in amazement" that Taubman "comes out of jail and he's just as popular and giving as many parties as he ever did! There's no disgrace in going to jail anymore unless it's for some disgusting, disgusting crime." Larissa MacFarquhar, *East Side Story: How Louis Auchincloss Came to Terms with His World*, NEW YORKER, Feb. 25, 2008, at 54. In addition, Taubman's name will remain forever on several buildings on the campuses of Harvard, Brown, and the University of Michigan. Ariana Eunjung Cha, *Corporate Scandals Tainting Donations*, WASH. POST, Sept. 15, 2002, at A1.

46 Tefft Smith wrote:

> *First*, the individuals typically carved-out in the corporate plea agreements (which give a pass on prosecution, assuming cooperation with any Division investigative requests, to all but the "carve-outs") tend to be mid-level sales and marketing

undermines the overall effectiveness of prison as a way to prevent cartelization. Therefore, we attempted to track down the past and present positions of executives imprisoned for criminal price fixing.

Of the 152 known individuals who received a fine or prison sentence in cartel cases between 1990 and 2008, we[48] were able to determine the position held during the cartel's existence for 151 of them.[49] Of those, 40 appear to have been one of the heads of the companies for which they worked;[50] 24 appear to have occupied a corporate position that was very high, but below the level of those in the former group;[51] 77 appear to have been mid-level employees; 3 were co-owners or sole-proprietors of a business; 3 were stamp dealers; and 4 were consultants.[52] Of the mid-level employees, 35 were clearly involved in sales or marketing.[53]

Still, another problem arises from the fact that some of the corporations involved forgive or even reward their price-fixing employees—directly or indirectly, legally or not—after they "take a bullet for the team" by going to prison.[54] Although it is difficult to

executives with "direct participation" or "knowledge" and "an ability to stop" the price-fixing. They tend *not* to be the senior executives, even when sometimes (in the Division's view) the senior executives are said by the Division to have been "willfully ignorant" of the misconduct.

Smith, *supra* note 40, at 9.

47 "And so it has always been true, and I am sure it is still true, that at the end of the day you're not going to get—it is very rare to get—the big multinational or national large corporation CEO or top guy as your antitrust defendant." Nanni, *supra* note 8, at 39.

48 W. James Denvil, *What Happens to Executives Who Are Sanctioned for Their Involvement in Cartels?* (on file with the author) *available at* http://www.cardozolawreview.com/content/34-2/Connor.Lande.34.2/DenvilStudy.pdf. This research was conducted by W. James Denvil while he was a student at the University of Baltimore School of Law. He is not a trained private investigator. He conducted his searches between July 15, 2010, and March 26, 2011, using Google, Bing, LinkedIn, Facebook, corporate websites, and the Federal Bureau of Prisons Inmate Locator. He searched for the individuals by using their full names, variants of those names, the names of their employers, descriptions of their cartels, the dates of their convictions, and the recent years (i.e., 2007–2010) as search terms. Because he could only rely upon public data, much of which could be unreliable, this survey should be regarded as extremely tentative, and only suggestive of what the actual results are likely to be. We urge others to conduct a more rigorous analysis of this issue.

49 However, job titles can be misleading and may not accurately reflect an individual's true position in the company.

50 *See* Denvil, *supra* note 48. This group is comprised of individuals with the title of Chairperson, President, Owner, Co-owner, Managing Director (of a European company), CEO, or COO.

51 *See id*. This group is comprised of individuals with the title of Commercial General Manager, Operations Manager, Director, Executive Vice President, President (of a division within the company), Managing Director (of a division within the company), Vice President of Operations, Commercial Director, CFO, or Co-Managing Director.

52 *See id.* This group consists of the individuals not included in the former two groups.

53 *See id.* These individuals have the words "sales," "marketing," or "development" in their titles.

54 *See* Dan Levine, *Antitrust Convictions Don't Mean End of Job for Some Executives*, RECORDER, Apr. 12, 2010, http://www.law.com/jsp/article.jsp?id=1202447903832&rss=

determine when or whether it would be legal,[55] the authors would not be surprised if it were common for the corporations involved to pay their executives' fines directly or indirectly in the form of bonuses or promotions.[56]

We attempted to discover how often convicted corporations forgive, and even reward, employees who violate the antitrust laws, and believe the results show that it may be common. We were able to determine the present whereabouts of 35 (34%) out of 103 managers[57] known to have received a prison sentence in cartel cases between 1995 and 2010.[58] Of those 35, 9 (26%) are currently employed by the company for which they worked during the cartel, and another 9 (26%) seem to be working at a different company within the same industry.[59] The remaining 17 are either in prison, unemployed, employed in different industries, or deceased.[60] Because we were unable to discover the whereabouts of 68 of the 103 who received a prison sentence, these results might not be statistically significant. Nevertheless, if the employment statistics of the out-of-sample price fixers resembles those of the known ones, approximately half of those who served a prison

newswire (describing an executive who was sent to prison for six months for price fixing, and when released, was made a senior vice president "with more responsibility than he had before he entered prison . . . ."). One reason for this may be that "since the executives are not perceived to have ripped off shareholders for personal gain, companies often have no problem welcoming them back into their corporate suites. . . . [S]ome corporate honchos believe executives that pleaded guilty took a bullet for the team, according to white-collar lawyers and industry observers." *Id.* Indeed, they have in all likelihood enriched the stockholders because the penalty their conduct led to probably was too low.

There are also reports that some companies continue to pay employees while they are in prison. *Id.* In the opinion of Tefft Smith the Antitrust Division does not get involved in employment decisions:

> [I]n my experience, the Division appears indifferent as to what the companies do with even the carved-out individuals (let alone the other executives who may have been identified as having been directly involved in the price-fixing). They need not be fired, disciplined or even re-assigned to non-sales and marketing-oriented jobs.

See Smith, *supra* note 40, at 10.

55 *See* 1 ROGER MAGNUSON, SHAREHOLDER LITIGATION § 9:37 (2010); *see also* Pamela H. Bucey, *Indemnification of Corporate Executives Who Have Been Convicted of Crimes: An Assessment and Proposal*, 24 IND. L. REV. 279 (1991); Note, *Indemnification of Directors: The Problems Posed by Federal Securities and Antitrust Legislation*, 76 HARV. L. REV. 1403 (1963).

56 JOHN M. CONNOR, GLOBAL PRICE FIXING 419–20 (2001) (describing how during cross-examination at the famous 1998 trial of three top executives of ADM for price fixing, the lead (immunized) witness for the prosecution was made to admit that his employer had paid his entire fine and promoted him to president of one of its largest subsidiaries).

57 In several cases, individuals were sanctioned but not their very small businesses. Thus, we excluded individuals who were stamp dealers, consultants, sole proprietors, or co-owners during the cartel. Many of the 152 defendants' sentencing details are not posted on the Antitrust Division's Web site. We thank the Division for providing the missing sentencing documents.

58 *See* Denvil, *supra* note 48.

59 *See id.*

60 *See id.*

sentence for their crime currently are working for their previous employers or in the same industry.[61]

We were also able to discover the current whereabouts of four people who received fines, but no prison sentence during the period between 1995 and 2009.[62] Two of them are employed by the same company for which they worked during the cartel, one appears to be working in the same industry, and the other is working in another industry.[63]

Indeed, for executives who went to prison, our figure of 52% almost certainly significantly underestimates the percentage of price fixers who went back to the same firm or industry. Some individuals may have reached retirement age, or returned to a firm or industry without notice of this fact being published in a source that is easily web-accessible, or the notice of some individuals' re-employment may have been deleted from the Internet prior to July 15, 2010.[64] Our survey may have erroneously counted such people as not having returned to their firm or industry.

The fact that some—perhaps most—corporations do not punish, and even reward, the individuals responsible for antitrust violations is only one reason why we are not persuaded by the argument that only individual sanctions matter. First of all, the financial well-being of the affected corporations often do matter to the individuals involved, as evidenced by corporate executives who, by fixing prices, often knowingly risk imprisonment largely for the financial benefit of their employers. Moreover, the literature on antitrust law generally assumes that corporations maximize profits, which means it also assumes the interests of corporate representatives and corporations generally align.[65] A corporation that truly does not want to break the laws against price-fixing because of the sanctions involved has any number of means to ensure that its employees follow company policy.[66]

---

61 *See id.*

62 *See id.*

63 *See id.*

64 *See generally id.* (noting that the research was conducted between July 15, 2010, and September 27, 2010).

65 *See, e.g.*, RICHARD A. POSNER, ANTITRUST LAW, at ix (2d ed. 2001) (arguing there is in the antitrust field a consensus that "business firms should be assumed to be rational profit maximizers, so that the issue in evaluating the antitrust significance of a particular business practice should be whether it is a means by which a rational profits maximizer can increase its profits at the expense of efficiency" (citing Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574 (1986))).

66 *See* Posner, *Optimal Sentences*, *supra* note 15, at 418 ("[I]f [the corporation] is liable it will find adequate ways of imposing on its employees the costs to it of violating the law."). Judge Posner noted: "A corporation has effective methods of preventing its employees from committing acts that impose huge [antitrust] liabilities on it. A sales manager whose unauthorized participation in a paltry price-fixing scheme resulted in the imposition of a $1 million fine on his employer would thereafter, I predict, have great difficulty finding

There are, moreover, a number of practical problems with exclusively or heavily relying on prison sentences as a means of deterring cartels, particularly international ones. First, it is more difficult to persuade managers of cartels who reside abroad to submit to U.S. jurisdiction. While indictments of foreign residents have increased, improvements in the ability of U.S. authorities to extradite individuals for price-fixing crimes have not kept pace.[67] There are large numbers of indicted cartel managers who are fugitives residing abroad.[68] Second, obtaining convictions of cartel managers who exercise their rights to a jury trial and who are within U.S. jurisdiction has proven challenging for the DOJ. Prosecutorial losses at trial are frequent.[69] Third, the demonstration effect of imprisonment requires adequate publicity about prison sentences. As the number and length of antitrust prison sentences have increased and they have become more routine, the "shock and awe" effect may decline. To offset such a trend, the DOJ has announced ever tougher standards for incarceration. It is unclear, however, whether these have been implemented to a significant extent or are mostly bluster. Fourth, coordination among those few antitrust authorities who incarcerate executives guilty of global price fixing is rare and likely to remain so in the future.[70] Where a cartel's injuries are multi-jurisdictional, multiple corporate fines have become common. However, there are no treaties on multiple incarcerations of cartel managers, so double-jeopardy concerns may well undermine the chances that the overall level of individual sanctions could be optimal.

The following matrix illustrates some of the issues involving the public policy issues underlying decisions to impose individual or corporate responsibility:

responsible employment, and this prospect should be sufficient to deter." POSNER, *supra* note 65, at 271. Posner first published this in 1976, when antitrust fines were very low. Since he believed corporations had an adequate incentive and means to control its employees when faced with prospects of a $1 million fine, a fortiori they would do so when faced with a possible $100 million fine.

67 *See* Julian M. Joshua, Peter D. Camesaca & Youngjin Jung, *Extradition and Mutual Legal Assistance Treaties: Cartel Enforcement's Global Reach*, 75 ANTITRUST L.J. 353 (2008).

68 *See infra* Part IV.B.

69 *See* Connor, *supra* note 14.

70 *See* CARTEL SANCTIONS, *supra* note 42.

Table 1
Optimal Cartel Deterrence:
Corporate v. Individual Sanctions Matrix

Individual Executives' View of Incentives	**Corporate Perspective**		
	Corporation has little incentive to create right climate or control employees	Optimal Corporate Incentive Level	Corporation has excessive incentive to control employees and create a climate that rewards honesty
Too Low on Average	**1**	**2**	**3** Wastes corporate resources, unfair to stockholders
Optimal	**4**	**5** Ideal Balance	**6**
Too High on Average	**7** Unless risk-loving, executives have little incentive to break law	**8**	**9** Additional negative—unfair to honest employees. But firm can ameliorate by paying fines, payments, or post-conviction employment

One way to analyze these possibilities is in terms of error analysis. Type I error involves problems arising from over-deterrence (this arises most in cell 9). Since collusion is judged under a criminal "beyond a reasonable doubt" standard, these errors are likely to be rare and mostly theoretical. Nevertheless, from the corporate perspective honest behavior can be mistaken for collusion, and this could be costly to society because it would cause corporations to refrain from procompetitive practices. The resulting fines would be unfair to

stockholders and cause over-investment in collusion prevention (although the actual costs of compliance programs are likely to be very small). From an employee perspective: They face prison and fines for honest behavior. But firms can ameliorate this by paying these fines for them (legally or illegally,[71] directly or indirectly, perhaps through direct payments through foreign banks, and/or post-conviction employment). Such behavior, to the extent it is not merely theoretical, is inefficient.

Type II error involves problems from under-deterrence (this arises most in cell 1). Inadequate sanctions will fail to deter collusion optimally. From a corporate perspective: The corporation is unjustly enriched from the illegal overcharges. Untold billions of dollars will be stolen from U.S. consumers and businesses, often by foreign lawbreakers. One study covering just forty private U.S. antitrust cases from 1990–2007 documented returned overcharges of more than $18 billion.[72] From an employee perspective: Employee activity that should go into productive and competitive behavior, instead, often will be directed towards establishing or maintaining collusion, or preventing its discovery. Moreover, as noted earlier, collusion also results in allocative inefficiency and other inefficiencies.[73]

In addition to Types I and II error, a system of cartel sanctions also should consider a third type of error. Type III error occurs when the system created to decide the issues leads to increased costs to businesses, consumers, enforcers, or decision-makers.[74] In the cartel context, these costs include litigation expenses by the enforcers, plaintiffs, and defendants, and their expert witness costs. It includes the costs arising from delays, and also the value of corporate time spent on these issues. It also includes the undesirable effects on society arising from any increased business uncertainty, and the increased cost to the judicial system, which imposes additional costs on taxpayers. Quantitatively, Type III error can be significant,[75] and any policy that ignores it runs a substantial risk of departing from an optimal result.

---

71 It is difficult to determine whether the antitrust fines imposed on corporate employees are ultimately paid by the employees, or are often or usually directly or indirectly paid by their employers. *See supra* note 54 and accompanying text. It also is difficult to determine whether it would be legal for the corporation to pay these fines. This area of law is exceedingly complex and, of course, even if indemnification is illegal, this does not mean it does not occur regularly. *See* ROGER MAGNUSON, SHAREHOLDER LITIGATION § 9:37 (2010); Bucey, *supra* note 55; Note, *supra* note 55.

72 Robert H. Lande & Joshua P. Davis, *Benefits from Private Antitrust Enforcement: An Analysis of Forty Cases*, 42. U.S.F. L. REV. 879 (2008), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1090661.

73 *See supra* note 22; *infra* Part III.B.

74 *See* Alan A. Fisher & Robert H. Lande, *Efficiency Considerations in Merger Enforcement*, 71 CALIF. L. REV. 1580, 1670–71 (1983), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1684227 (introducing the concept of Type III error; defining and using these terms in a related antitrust context: merger enforcement).

75 *Id.*

We know of no way to secure the information necessary to quantify and minimize these errors. Nevertheless, we believe it is likely that optimal deterrence only can be secured by a mix of corporate and individual sanctions.[76] If violations only were subject to corporate penalties, individuals might be unduly tempted to form cartels because success would benefit them tremendously and, as has been suggested by anecdotes[77] and some research,[78] they often do not face significant internal sanctions for their illegal behavior[79] and might well even be rewarded for their suffering in prison. On the other hand, if only individual penalties existed, it could be in the interests of some corporations to establish internal incentives that failed to discourage, rewarded, or even coerced employees into engaging in illegal behavior.[80] Some corporations might prefer to offer up a few executives for multi-year prison terms rather than pay $100 million or more as a criminal fine or payout in private litigation.[81] The employees could be incentivized to risk prison by multi-million dollar bonuses, perhaps paid to foreign bank accounts or in the form of future employment. Even though these payments might be quite large for individuals, they easily could be dwarfed by the prospective fine that could be imposed under a regime oriented towards corporate fines.[82]

76 In addition, it is important for a society to create a cultural norm that cartel behavior, like stealing, is something that is strongly condemned across that society. It is important that the prohibition against price fixing become a moral or social standard that is internalized within the business community. Many people refrain from stealing because they think it is the right thing to do, not because of the threat of fine or incarceration. Attaching social stigma to the act is an important aspect of optimal deterrence. *See generally* John M. Connor, Albert A. Foer & Simcha Udwin, *Criminalizing Cartels: An American Perspective*, 2010 NEW J. EUR. CRIM. LAW 199, *available at* http://www.antitrustinstitute.org/sites/default/files/NJECL%202010.pdf; Andreas Stephan, *"The Battle for Hearts and Minds": The Role of the Media in Treating Cartels as Criminal*, *in* CRIMINALISING CARTELS: A CRITICAL INTERDISCIPLINARY STUDY OF AN INTERNATIONAL REGULATORY MOVEMENT, *supra* note 31, at 381, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1866285.

77 *See supra* notes 56–67.

78 Khanna, *supra* note 33, at 1485–86; *supra* notes 56–67.

79 Greg Werden suggests additional reasons: "This can occur as a result of defects in the design of compensation schemes, especially if the executives have short time horizons or are more willing than business enterprises to take risks. Consequently, business enterprises can incur substantial costs in monitoring their executives and complying with the law." *See* Werden, *supra* note 20, at 31–32 (footnotes omitted).

80 *Id.* at 32.

81 Suppose that instead of a corporate fine or payout in private cases a corporation could offer up to the DOJ five executives who would each be sentenced to two years in prison or under house arrest. Suppose the corporation could pay each of the individuals involved $6 million per year by depositing the appropriate sums in Swiss bank accounts, and also guarantee they would return to their position in the company upon release. This would only cost the corporation $60 million, far less than many of the larger fines that have been imposed in recent years, and far less than many of the private payouts of recent years.

82 Perhaps in part because corporations often would be able to compensate the punished individuals who "took one for the team," the "rogue manager" defense rarely has been accepted by the Antitrust Division or by the courts.

We certainly do not know how to devise a formula to compare alternative cartel sanctions. Nevertheless, it is our judgment that a financial penalty against an individual has more of an impact on deterrence than a similar penalty against a corporation, and that prison time or the loss of one's corporate position[83] often is the equivalent of a very large financial penalty. We make accommodations for these assumptions in our analysis in Part III by tripling the disvalue or deterrence effects of individual sanctions relative to corporate sanctions.

II. THE OVERALL LEVELS OF CURRENT CARTEL SANCTIONS

Violations of the U.S. antitrust laws can result in a diverse array of criminal sanctions. These include corporate fines and restitution payments, as well as prison, house arrest, and fines for the corporate officials involved. During the 1990 to 2010 period the total amount of corporate fines imposed in every DOJ criminal antitrust case was $6.174 billion.[84] The total of the individual antitrust fines imposed was $74 million.[85] The Antitrust Division also secured the restitution of $165 million in conjunction with criminal antitrust cases[86] (which largely or totally consisted of restitution to the federal government for overcharges it paid).[87] Its enforcement also resulted in sentences against 367 individuals[88] that total 186,393 days (510 years) in prison.[89] Antitrust enforcement also led to another 112 years of "house arrest or

83 Donald Klawiter, an extremely experienced practitioner and former Chair of the ABA Antitrust Section, at the American Antitrust Institute's Annual Conference, held on June 23, 2011, in Washington, D.C., noted during the session on international cartels:

> From my experience in representing corporations and their executives in these cases, two things terrify executives. The first is the possibility that they will go to jail, if even for a week. And the second is that they will . . . lose their high level positions in corporations. Indeed, I've had some confess that taking them out of the CEO job or the head of sales job is much more traumatic to them than spending a year and a half in jail. That's sort of an interesting rationale and I think an interesting fact that we should look at.

Donald Klawiter, Partner, Sheppard Mullin Richter & Hampton LLP, International Cartels Presentation at American Antitrust Institute Annual Conference (June 23, 2011) (audio available at http://www.antitrustinstitute.org/content/international-cartels-presentation-and-audio-aai-annual-conference).

84 *See* WORKLOAD STATISTICS 2002–2011, *supra* note 25, at 11. The yearly figures are reproduced and summed in Lande & Davis, *supra* note 14, at 33 tbl.1.

85 *Id.* The yearly figures are reproduced and summed in Lande & Davis, *supra* note 14, at 34 tbl.2.

86 *Id.* at 12. The yearly figures are reproduced and summed in Lande & Davis, *supra* note 14, at 35 tbl.3.

87 As the Division's Workload Statistics notes with considerable understatement, "Frequently restitution is not sought in criminal antitrust cases, as damages are obtained through treble damage actions filed by the victims." *Id.* at 12 n.15.

88 *Id.* at 12.

89 *Id.*

confinement to a halfway house or community treatment center" for 262 individuals.[90]

Now, we turn to civil sanctions secured by private plaintiffs. Cartel victims receive mandatory treble damages and attorneys' fees.[91] Final verdicts in cartel cases are exceptionally rare, however. Our 2004 search for every final verdict in a U.S. cartel case since 1890 found only twenty-five examples.[92] Nevertheless, many private cases have resulted in significant settlements. An analysis of well over 100 international cartels prosecuted between 1990 and 2008 found a total of $29 billion in announced private settlements in U.S. cases.[93] The only other estimate we have found was for a very limited sample of twenty-five large private cases filed against cartels between 1990 and 2007, which documented between $9.2 billion and $10.6 billion in cash payments (not including the value of products, coupons, or discounts).[94]

---

90 *See* U.S. DEP'T OF JUSTICE, ANTITRUST DIVISION WORKLOAD STATISTICS FY 1990–1999, at 13 (2009) [hereinafter WORKLOAD STATISTICS 1990–1999], *available at* http://www.justice.gov/atr/public/246419.pdf; U.S. DEP'T OF JUSTICE., ANTITRUST DIVISION WORKLOAD STATISTICS FY 2000–2009, at 14 (2012) [hereinafter WORKLOAD STATISTICS 2000–2009], *available at* http://www.justice.gov/atr/public/281484.pdf; WORKLOAD STATISTICS 2002–2011, *supra* note 25, at 12. However, these figures might be too high for the purposes at hand, for two reasons. First, these figures are for time sentenced, not time served. We were unable to determine how much of this time actually was served or how often sentences were reduced. For example, A. Alfred Taubman was sentenced to prison for a cartel offense for twelve months, but only served nine-and-a-half months. *See* note 45 *supra*.

Second, sometimes an investigation by the Antitrust Division results in a sentence for another crime regardless of whether an antitrust violation was uncovered. Non-price-fixing crimes can include perjury, mail fraud, contempt, obstruction of justice, and false statements. WORKLOAD STATISTICS 2002–2011, *supra* note 25, at 8 (listing these crimes under the header "Other Criminal Cases"). Since the Antitrust Division uncovered these crimes, often Antitrust Division investigators are in the best position to pursue these non-antitrust issues. Most often, these other crimes are related to an antitrust offense—such as when a cartel bribes a federal purchasing agent. Other times they are not related, and quite often, they are very difficult to classify. According to the Antitrust Division, "Other Federal Crimes such as Perjury, Mail Fraud, Contempt, Obstruction of Justice, or False Statements" apparently constituted 36% of their criminal convictions since 1990 (53% during 2008–2009).

We do not, however, know how many of the 186,393 days of prison secured by Antitrust Division enforcement were imposed for crimes that were not antitrust related. Ideally, we would subtract these before we conduct our optimal deterrence analysis. For lack of data, and to be conservative, we are ignoring these issues. The figures reported above for prison time and house arrest, therefore, will be used in our subsequent analysis even though they include some individuals serving time in whole or in part for non-antitrust offenses. And, as noted, these are time sentenced, not time served statistics. Because these statistics are larger than they should be for our purposes, their use will overestimate the probable deterrence effect of the DOJ's anti-cartel program.

91 Prevailing plaintiffs also receive filing fees and expert witness fees. *See supra* note 5.

92 *See* Connor & Lande, *supra* note 14.

93 John M. Connor, Cartels & Antitrust Portrayed: Private International Cartels from 1990 to 2008, at 51 (American Antitrust Institute Working Paper No. 09-062009), *available at* http://ssrn.com/abstract=1467310.

94 Since almost all these cases were settlements, "alleged victims" would usually be a more accurate description. *See* Lande & Davis, *supra* note 14. These figures have not been adjusted for inflation. These cartel payouts constituted a part of a larger study of forty private cases that

We have aggregated all of these types of cartel sanctions and we now analyze them according to the standard optimal deterrence model. First, we have assembled the financial penalties imposed on corporations, including the amounts they pay in corporate fines, restitution actions, and private treble damages actions.[95] Second, we have assembled the fines imposed on the individual corporate actors who were held personally responsible for cartel violations. Third, we developed monetary equivalents of time in prison (or time spent under house arrest) by approximating the disvalue, cost, deterrent value, or opportunity cost of incarceration time.[96] Admittedly, establishing the likely disincentive effect of prison in an objective, accurate, and non-

documented a total of $18–19.6 billion returned to victims or alleged victims of antitrust violations. *Id.*

95 There are three additional types of monetary costs that we have not been able to quantify. First, antitrust suits are costly to defend. The amounts antitrust defendants pay in attorneys' fees usually are confidential, however, and we know of no way to systematically estimate them. One could assume they are equal in size to the plaintiffs' attorneys' fees, which are matters of public record in class action cases, and then include them in the calculations. We do not know, however, whether this would be a close approximation. Second, antitrust suits cause corporate disruption and wasted time for the corporate executives involved. We know of no method to evaluate this type of corporate loss. Third, an antitrust conviction could harm a company's reputation and cost it business, and could decrease an individual's future income and lower their reputation and social status. We know of no way, however, to quantify such losses. In addition, society must pay to incarcerate people. We believe this cost is relatively small.

Regardless, our decision to triple the $2 million "cost" of a year in prison should more than cover adjustments that should be made for these factors.

96 Note the important difference in these two baselines: corporate actors might demand a different sum to risk prison than they would be willing to pay to avoid the risk of prison. For example, suppose someone would rather pay a $6 million fine than be imprisoned for a year. How would that person react to the question of whether they would accept $6 million in return to going to prison for a year? They might not agree to this deal. Part of the difference is the relative wealth of the actor in the two situations. A corporate actor could in theory demand an unlimited amount to accept the risk of prison, and any such payment increases his or her wealth. But the same person cannot pay an unlimited amount to avoid the risk of prison; she can only spend as much money as she has or can borrow. *See* David Cohen & Jack L. Knetsch, *Judicial Choice and Disparities Between Measures of Economic Values*, *in* CHOICES, VALUES, AND FRAMES 424, 428 (Daniel Kahneman & Amos Tversky eds., 2000).

But there is another element at play here as well. Empirical evidence shows that people's attitude toward costs and benefits depend on their perception of the status quo. *Id.* at 428–29. A person who accepts prison as the status quo may be willing to pay less to avoid it than a person who sees prison as a deviation from the status quo. A corollary is that, depending on the odds and stakes, people value avoiding losses—and are willing to take risks to do so—far more than they value gains, which they generally will not take risks to do (although, oddly, this principle may vary depending on the odds of the risk and the size of the gain or loss). *See* Daniel Kahneman & Amos Tversky, *Choices, Values, and Frames*, *in* CHOICES, VALUES, AND FRAMES, *supra* at 1, 35–36. This psychological phenomenon—and others—greatly complicates an economic analysis of behavior. So, for example, a corporate actor who perceives herself as taking steps that violate the antitrust law to return to the status quo (perhaps because she thinks her corporation is suffering from unfair competition) may be far more tolerant of risk than the same corporate actor who contemplates the same measure as a means of obtaining a perceived economic advantage. Even for a single corporate actor, then, there may be no single correct amount that represents her willingness to trade off between gain for her corporation and the risk of prison for herself.

controversial manner is impossible. Because our attempt to monetize incarceration is a relatively novel feature of this Article, we allocate the major part of this Section to this topic.

Some might contend that, because no corporate officer wants to spend any time in prison or under house arrest, they would be willing to pay virtually any amount of money to avoid the risk of prison. This is equivalent to placing an infinite (negative) value on prison time, and it implies that even a small probability of spending any time in prison or under house arrest has an infinite deterrence value. However, people do not *act* as if they infinitely disvalue the risk of getting put into prison or placed under house arrest for an antitrust offense. If they did, they would never try to form a cartel because this would put them at risk of going to prison. Rather, potential offenders act as if they tolerate the risk of prison to some extent. Perhaps they calculate, implicitly, on the basis of legal advice and what they have heard from other executives, their apparent chances of getting caught and convicted, and the prison sentence, house arrest, or fine they are likely to face, at least to some very rough degree.[97] They then balance this chance of a penalty and its likely size, again in an extremely rough way, against the rewards of cartelization. In any case, we know that often they decide to form cartels. We know they often make this decision because cartelists surely know cartels are illegal, yet the number of cartels caught in recent years has been quite significant and does not seem to be decreasing.[98]

Since the disutility of prison time is not infinite, in theory we can approximate its value, though to do so in practice is, of course, difficult and speculative. There is no one objective way to compare the deterrence effect of time spent in prison to the deterrence effect of a criminal fine because different people would trade off prison versus fines in different ways. Any "average" figure used to equate the two is necessarily imprecise and arbitrary.

The valuation of custodial time is similar to one that, regrettably, society often must undertake for any number of public policy purposes. Sometimes even a life must be valued finitely. For example, our nation cannot afford perfect safety, nor would we want every automobile to be built as safely as technically possible.[99] Similarly, even though a life is beyond value and society does not want people to drive negligently, courts do not award infinite damages for the loss of life in car crashes.

---

97 As noted earlier, direct or indirect payments of fines or rewards for imprisonment by their employers might sometimes also be a factor. *See supra* note 71 and accompanying text.

98 *See supra* note 26.

99 If society did this, it would be forced to accept increased risks from other sources (i.e., society cannot afford perfect safety).

We present five different approaches to the issue of how to evaluate the cost or value of time in prison.[100] We expect that considering the use of multiple approaches will increase the reliability of our results.

The first approach is to ascertain the valuations of lives and years of life used for various regulatory, public policy purposes.[101] In the United States, lives typically are valued at between $3 million and $10 million by federal government agencies when they set, for example, transportation or environmental policy.[102] Some of these studies are especially appropriate for our purposes because they place average values on a year of life. They generally calculate figures of $300,000 to $500,000 per person per year of life (depending upon a number of variables).[103]

Second, lower figures on average, from $1.4 million to $3.8 million for a life, are awarded under tort law, in wrongful death cases.[104]

Third, following the September 11th tragedy, Congress created the September 11th Victim Compensation Fund to award compensation to victims' families.[105] The Fund's payments constitute a prominent recent reflection of the monetary value our society places on innocent human life, even though these payouts were made under unique circumstances. The Fund's average award for a life was $2,937,861, the median award was $1,677,632, the maximum award was $7,100,000, and the minimum award was $250,000.[106] Significantly for our purposes, many of the September 11 victims had been quite affluent. Eighty-nine of the victims had annual incomes between $500,000 and $1,000,000 per year (their estates were given average awards of $4,749,654), and eight victims'

100 These presented approaches have been adapted from Lande & Davis, *supra* note 14, at 14–19.

101 For a concise essay on economic methods for evaluating "statistical lives," see Thomas C. Schelling, *Value of Life*, *in* 4 THE NEW PALGRAVE: A DICTIONARY OF ECONOMICS 793–96 (John Eatwell et al. eds., 1987).

102 *See* Joseph E. Aldy & W. Kip Viscusi, *Adjusting the Value of a Statistical Life for Age and Cohort Effects*, 90 REV. OF ECON. & STAT. 573 (2008). Recently, the Department of Transportation has used $5.8 million for the value of a life. Memorandum from Tyler D. Duvall, Assistant Sec'y for Transp. Policy, and D. J. Gribbin, Gen. Counsel, to Secretarial Officers & Modal Adm'rs (Feb. 5, 2008), *available at* http://ostpxweb.ost.dot.gov/policy/reports/080205.htm. The Environmental Protection Agency currently uses $6.9 million. *All Things Considered: Value on Life 11 Percent Lower Than 5 Years Ago* (NPR radio broadcast July 11, 2008), *available at* http://www.npr.org/templates/story/story.php?storyId=92470116.

103 *See* Aldy & Viscusi, *supra* note 102. For example, values typically decline with age, and we note that most price fixers are mature businessmen. *Id.*

104 *See* Mark A. Cohen & Ted R. Miller, *"Willingness to Award" Nonmonetary Damages and the Implied Value of Life from Jury Awards*, 23 INT'L REV. L. & ECON. 165, 166, 179 (2003) (calculations made in 1995 dollars).

105 *See* Air Transportation Safety and System Stabilization Act of 2001, Pub. L. No. 107-42, 115 Stat. 230 (codified at 49 U.S.C. § 40101 (2006)) [hereinafter "the Act"]. We are grateful to Thomas Weaver for his research involving the September 11th Victim Compensation Fund.

106 1 KENNETH R. FEINBERG ET AL., FINAL REPORT OF THE SPECIAL MASTER FOR THE SEPTEMBER 11TH VICTIMS COMPENSATION FUND OF 2001, at 110 tbl.12 (2001), *available at* http://www.justice.gov/final_report.pdf.

annual income exceeded $4,000,000 per year (their estates were given average awards of $6,379,287).[107] Although we do not know the average or typical pre-conviction annual incomes of imprisoned price fixers, we would not be surprised if the latter income levels are comparable.

A disadvantage of these first three approaches is that they address the cost or disutility of lost lives, not time spent in prison. It is likely that most people would view the prospect of spending a year in prison as not as bad as losing a year of life; after all, many prisoners with no chance at parole still resist the death penalty. Thus, the first three approaches may be regarded as an upper bound on the disutility of a year in prison.

A fourth method for approximating the disvalue of incarceration comes from examining the compensation provided to defendants who have been wrongly imprisoned. Sometimes people are wrongly imprisoned by, for example, perjured government testimony.[108] The victims potentially can recover for a variety of torts depending upon the jurisdiction.[109] Often no award will be given for imprisonment due to a simple, albeit tragic, error; some type of intentional act, malice, or malfeasance typically is required.[110] The highest payment we found for a case involving at least a year of prison was $1.164 million per year, for three years of wrongful confinement for a false conviction.[111] However, when shorter imprisonments are annualized, significantly higher awards sometimes have been made.[112]

---

107 *Id.* at 97 tbl.6.

108 *See* Limone v. United States, 497 F. Supp. 2d 143, 152 (D. Mass. 2007) (stating the FBI was aware chief witness would perjure himself); *see also* Newsome v. McCabe 319 F.3d 301, 304–05 (7th Cir. 2003) (stating the officers induced eyewitnesses to falsely identify plaintiff); Bravo v. Giblin, No. B125242, 2002 WL 31547001 (Cal. Ct. App. Dec. 18, 2002) (unpublished) (stating the investigating officer fabricated evidence).

The authors are grateful to Thomas Weaver for locating and analyzing these cases, and for performing research on this subject. *See* Thomas Weaver, *The Part That Counts: Wrongful Incarceration Awards and the Value of Human Life* (May 1, 2011) (unpublished manuscript) (on file with the authors).

109 These torts include wrongful imprisonment, wrongful conviction, wrongful confinement, malicious prosecution, abuse of process, intentional or negligent infliction of emotional distress, false arrest, or an unconstitutional depravation of their civil rights. *See* Weaver, *supra* note 108.

110 *See, e.g.*, cases cited *supra* note 108.

111 *Bravo*, 2002 WL 31547001, at *24. The suit, filed under 42 U.S.C. § 1983, yielded "damages in the amount of $221,976 for his economic losses, $3,537,000 to compensate him for 1179 days of incarceration at the rate of $3000 per day, and $1 million to compensate him for emotional distress suffered between the date of the incident and the date of his sentencing." *Id.* We arrived at the award per year of imprisonment of $1,164,515.62 in this case by the following steps: 1) multiplying $3,000 a day by 365.25 days to arrive at $1,095,750; 2) the lost earnings of $221,976, divided by 1179 days in prison comes to $188.27 per day, and when multiplied by 365.25 days, adds another $68,765.62 per year. The total award per year of imprisonment thus comes to $1,164,515.62.

112 The extreme case was *Ramirez v. County of Los Angeles*, 397 F. Supp. 2d 1208, 1215 (C.D. Cal. 2005) (noting that the investigating officer fabricated evidence). *See* Rob McKay, *Verdict of the Week: US Dist. Ct., Los Angeles*, VERDICTSEARCH, Mar. 13, 2006, at 21, *available at* http://www.kkcomcon.com/doc/Ramirez%20v%20LAPD.VS.pdf (reporting that a ten month

We should note that we have not been able to ascertain any of the falsely imprisoned defendants' incomes, but we suspect most had relatively low incomes, and none appears to have been a corporate executive or upper class professional.[113] It is possible that a jury or judge would award a corporate executive wrongfully imprisoned for price fixing a larger-than-average amount for their suffering. Alternatively, a jury might react in the opposite direction. A jury might be less sympathetic to imprisoned upper class corporate executives. Still, these results do tend to show that compensation in the neighborhood of $1 million per year appears generally to be the practical maximum that society is willing to award for a year wrongfully spent in prison.

Our fifth and final approach is to examine estimates of the disvalue of prison time made by reputable scholars. We have been able to find only two estimates for an antitrust offense that seem plausible in this context.[114] First, an Article by Professors Howard P. Marvel and others equated a year in prison for price fixing to approximately $600,000 in 2010 dollars.[115] Second, a study by Professor Kenneth Glenn Dau-Schmidt and others equated a year in prison for price fixing with a fine of approximately $1.5 million in 2010 dollars.[116] These figures are higher

sentence led to a $9 million settlement, or an annual rate of $10,800,000). Because the emotional stress and discomfort could be disproportionately high for the very fact of the government malfeasance, or greater for the beginning of a prison sentence, it is unclear whether the award would have been increased proportionately if the victim had been imprisoned for a year, or for multiple years. As noted, in these cases, moreover, it is difficult to segregate the amounts awarded for false imprisonment from the amounts awarded for one-time events or other torts. "Where the period of incarceration is shorter (e.g., less than one year), proportionately larger awards (measured by annualizing the award) have been rendered, presumably reflecting Limone's observation that the injury from incarceration may be more intense towards the beginning." Smith v. City of Oakland, 538 F. Supp. 2d 1217, 1242 (N.D. Cal. 2008); *see also* John Collins Coffee, Jr., *Corporate Crime and Punishment: A Non-Chicago View of the Economics of Criminal Sanctions*, 17 AM. CRIM. L. REV. 419, 431 (1980) ("[T]he declining marginal utility of imprisonment means that each increment of incarceration increases the perceived penalty by a less than proportionate amount. Or, reduced to its simplest terms, a two-year prison term is not twice as bad as a one-year term.").

113 *See* Lande & Davis, *supra* note 14.

114 We have found one other estimate, but it seems to value prison time at a level too low to apply to white-collar criminals. *See* Tonja Jacobi & Gwendolyn Carroll, *Acknowledging Guilt: Forcing Self-Identification in Post-Conviction DNA Testing*, 102 NW. U. L. REV. 263, 283 & n.52 (2008) (estimating value of prison at approximately $200 per day, which amounts to slightly more than $70,000 per year).

115 *See* Howard P. Marvel et al., *Price Fixing and Civil Damages: An Economic Analysis*, 40 STAN. L. REV. 561, 573 (1988). The authors equated a year in prison with a $373,000 fine. The Article appeared in the February 1988 issue, so we assume they were using 1987 dollars. The Bureau of Labor Statistics Consumer Price Index inflation calculator equates $373,000 in 1987 to approximately $677,000 in 2011. *See CPI Inflation Calculator*, BUREAU OF LABOR STATISTICS, http://www.bls.gov/data/inflation_calculator.htm (last visited Sept. 2, 2012).

116 Joseph C. Gallo et al., *Criminal Penalties Under the Sherman Act: A Study of Law and Economics*, *in* 16 RESEARCH IN LAW AND ECONOMICS 25 (Richard O. Zerbe, Jr. ed., 1994). Gallo's Article equated a year in jail with a fine of $1 million. The Bureau of Labor Statistics Consumer Price Index inflation calculator equates $1 million in 1994 with $1,486,000 in 2011.

than the national average valuations for a year of life noted earlier, perhaps because price fixers are wealthier on average and can afford to disvalue prison time much more than most people can, or perhaps because price fixers' time is more valuable on average.[117]

These five approaches yield estimates that are broadly consistent with one another. To be conservative, we have taken the highest of these estimates, $1.5 million per year, and increased it to $2 million (in 2010 dollars). We note that $2 million is as much as the lower estimates for the value of an entire human life that were discussed earlier, and is much higher than the average annual national values of life.

As discussed earlier, penalties directed against individuals might well have more deterrence effect than penalties directed against the corporations. To attempt to take this into account, and in an attempt to be conservative in our analysis,[118] we have trebled the deterrence effect of every individual penalty before adding them to the corporate penalties. This means we will use $6 million (in 2010 dollars) for the deterrence value of a year in prison.[119] We also will treble the individual fines paid in antitrust cases before we add these figures to the corporate fines, restitution payments, and payouts in private damages cases.[120] And, although we believe we should use a much lower value for house arrest than for prison time (such as $1 million or $3 million per year) for simplicity of calculations and to be conservative we will value a year of house arrest at $6 million, as well.

CPI Inflation Calculator, *supra* note 115. The authors, however, used 1982 data for much of their paper's analysis. If they meant their valuation of a year in jail to be expressed in 1982 dollars, their $1 million estimate would be the equivalent of approximately $2,282,000 in 2011. *Id.*

117 Whether the time or the life of a price fixer is more, or less, valuable than that of an average person is an interesting philosophical question this Article will not explore.

118 See also the factors listed in notes 96–97, *supra*. The incremental $4 million per year should more than compensate for these factors as well.

119 We note that valuing a year's worth of life at $6 million would mean that a twenty year prison sentence would be disvalued at $120 million, a figure far in excess of the amount society places on an individual's life.

We will use the $6 million valuation, in 2011 dollars, for the deterrence produced by a year spent in prison for price fixing even if that imprisonment occurred years ago.

We recomputed our analysis using different values for time spent in prison, such as $12 million per year, but this made no significant difference in our results. See *infra* note 250, which shows that only valuing a year in prison in the range of $1 billion would make a significant difference in our results.

120 This assumes that price fixers actually pay their own fines. It is, however, difficult to determine whether antitrust fines imposed on corporate employees ultimately are paid by the employees, or are often or usually directly or indirectly paid by their employer. *See supra* note 71.

III. CARTEL HARMS: THE "NET HARMS TO OTHERS" FROM CARTELS

The standard optimal deterrence formula shows that the total amount of cartel sanctions should equal the cartel's "net harm to others" divided by the probability of detection and proof of the violation.[121] The "net harm to others" from a cartel includes not only its overcharges, but also the allocative inefficiency[122] produced by its exercise of market power. The allocative inefficiency from cartel pricing should be added to their overcharges to get a true measure of cartels' "net harms to others."

In theory, each of these parameters should be an expectation that has been individualized for the cartel in question. For each potential cartelist we would ascertain what each thought their expected profits from cartelization were likely to be,[123] what their chances of getting caught and convicted were, and the total disvalue to them of the sanctions they thought would be imposed. This calculus would be made with due regard for how much each prospective cartel manager was risk-averse or risk-seeking.[124] As a practical matter, of course, ascertaining these required figures is impossible. The best we can do is to calculate what each figure actually has been on average in the past, and to assume that this figure is likely to be close to what the managers of potential cartels believe is likely to happen in the future. This is, of course, a highly imperfect exercise. Nevertheless, it is more likely to allow us to calculate whether cartel penalties have been set at the optimal level than any other approach we can devise.

A. *Cartel Overcharges*

In an earlier Article, we developed and presented a very different survey approach. We comprehensively and systematically examined cartel overcharges by assembling two data sets. The first consisted of scholarly publications containing cartel overcharges. With very few exceptions, we attempted to analyze every scholarly study that contained quantitative information on the price effects of private cartels.[125] We

121 *See supra* notes 14–22 and accompanying text.

122 *See supra* note 22 and accompanying text. Ideally the overcharges also should be adjusted upwards for the umbrella effects of market power. *Id.* Ideally the costs imposed on taxpayers for the government to investigate and prosecute and for courts to try cartels, and the costs to the public of incarceration, also would be included since they, too, are "net harms to others" from cartels. We do not, however, have information as to how large these omitted factors are.

123 Their expected cartel profits, moreover, would be a distribution of outcomes with assigned probabilities.

124 Another factor would be the opinion of each cartel manager as to their co-conspirators. Do they believe their co-conspirators are likely to turn them in under various circumstances?

125 *See* Connor & Lande, *supra* note 14.

separately categorized domestic and international cartels from different time periods to determine whether the increased penalties of recent years have been having significant effects. Our second data source was obtained by examining every final verdict in U.S. collusion cases that we were able to find.[126] We searched for antitrust cases in which a neutral finder of fact reported collusive overcharges in percentage terms or presented conclusions that could be converted into an overcharge percentage.

Our most recent compilation from scholarly publications found 1,517 useful estimates of cartel overcharges or undercharges in more than 200 publications that analyzed cartels that operated in 381 markets.[127] Table 2 displays the medians of all average overcharges reported over time.[128] The median cartel overcharge for all types and time periods (in a data set that includes a significant number of zeros) is 23.3%.[129] There is no strong trend in the cartel markups for all types over time. Indeed, the median since 2000 is virtually the same, 22.5%. But if one examines the international cartels separately, it is noteworthy that the median over time has been higher than for national cartels (30.0% and 17.2%, respectively), but thanks to a downward trend the international and national medians since 2000 have been similar (25.8% and 20.0%, respectively).[130] The mean overcharge figures have averaged 49%, much higher than the median figures due to the presence of some extremely large overcharges in the sample.

Our search for verdicts in cartel cases proved to be extremely difficult,[131] however, because overcharges are not calculated in criminal

126 *Id.* at 555–57.

127 *See* JOHN M. CONNOR, GLOBAL PRICE FIXING (2d ed. 2007).

128 *See infra* Appendix tbl.2. We choose to show the median overcharge percentages rather than the mean overcharge percentages because a few very high overcharges in any particular category can overwhelm a mean calculated using the larger number of low-to-medium percentage overcharges.

Another interesting statistic concerns the low number of overcharges by unsuccessful cartels. Only about 7% of the data we collected indicated that a cartel episode was unsuccessful in controlling prices significantly. We did, of course, include these observations in the median calculations that appear in Table 3, *infra*.

129 Cartel overcharges might not be passed on to the next level of distribution at the same percentage rate. An overcharge of 23.3% by a manufacturer cartel could pass through several levels in the distribution chain and result in a final consumer overcharge of more than, or less than, 23.3%.

130 It is difficult to know what to make of the downward trends in profitability for most types of cartels. The influence of the spread of, and increase in, effective anticartel enforcement is perhaps the most obvious explanation. The downward trend in overcharges among cartels that were caught by antitrust authorities tends to support the idea that cartelists find it increasingly difficult to hide their activities. Alternatively, the greater antitrust scrutiny in the United States from the 1940s and from Europe since the 1960s could prompt cartelists to refrain from full monopoly pricing increases so as to reduce their chances of detection.

131 We looked for cases by the use of computer-assisted searches of databases, searching through a large number of articles and treatises on cartels and on antitrust damages, and asking groups of knowledgeable antitrust professionals for any examples they knew of that might

enforcement against cartels[132] and because almost every private antitrust suit for damages settles[133] or is dismissed before an overcharge can be calculated by a neutral observer and made part of the public record of the case.

As a consequence, final verdicts involving cartels where a judge, jury, or commission[134] calculated an overcharge[135] are rare, and we found a disappointingly small sample size of cases—twenty-five—to analyze. However, our sample is roughly as large as the sizes of the prior surveys we report in Table 1 (which were 5–7, 12, 12, 13, 22, and 38 in number, respectively). Nevertheless, due to its small size, its results should be interpreted with caution. The results of this verdict analysis are that the twenty-five collusion episodes had a median average overcharge of 22%, and a mean overcharge of 31%.[136]

Thus, our two data sets yield median cartel overcharges of approximately 25% and 22% overall. The mean results were 49% for the economic studies and 31% for the verdicts.[137] For the economic studies' post-2000 sample, the national and international cartel median overcharges averaged 20% and 25.8%.[138] These figures will prove extremely useful when we formulate our policy recommendations in this Article's Conclusion. Part V of this Article, however, which will carry out the optimal deterrence calculations according to the standard approach, will use the actual amount overcharged by each individual cartel.

B. *The Allocative Inefficiency Effects of Market Power*

The "net harms to others" from cartels also include their allocative inefficiency effects (oftentimes called the deadweight welfare loss, or DWL).[139] Unfortunately, we do not know for very many cartels either how large their allocative inefficiency harms are or the relative size of a cartel's allocative inefficiency compared to its overcharges. We instead will select a representative ratio or range that is based on economic

contain useful information. *See* Connor & Lande, *supra* note 14, at 555–56.

132 Price fixing is illegal regardless whether, or the extent to which, defendant affected prices, because the agreement to fix prices is illegal. For this reason the amount that prices changed, or even whether prices were affected at all, is not calculated in a criminal antitrust case. *Id.* at 551.

133 *Id.* For a discussion of settlement in this context, and why settlement amounts are likely to be an extremely unreliable guide as to the size of the underlying cases' overcharges, see *id*.

134 Connor & Lande, *supra* note 14, at 551–52.

135 Moreover, many verdicts were only expressed in dollar amounts which we were unable to translate into percentages, so we reluctantly had to omit these cases. *Id.* at 556.

136 *Id.*

137 *See id.* at 561. The mean figures are significantly higher than the median figures due to the effects of extremely high overcharges.

138 *Id.* at 541.

139 For a definition of the allocative inefficiency effects of market power, see *supra* note 22.

theory and constants derived from the empirical literature on cartels and monopolies. Then, we will add the DWL to the cartel's overcharges when we implement the optimal deterrence calculations.

As an example of how adding this factor into the optimal deterrence calculations could make a difference, Judge Easterbrook, in an early paper on this topic, assumed that allocative inefficiency effects are 50% as large as overcharges on average.[140] If Judge Easterbrook was correct, this would mean that under the "net harm to others" standard, every $100 in overcharges would be presumed to be accompanied by another $50 in allocative inefficiency harm.

We located a modest number of technically impressive empirical studies specifically about cartels that provide both overcharges and DWL estimates. Sølgard computes a DWL/overcharge range of from 37% to 48% for a Norwegian cement cartel,[141] and Monke et al. find a 25% ratio for a Portuguese flour cartel.[142] Gallo et al. provide a comprehensive analysis of U.S. DOJ cartel cases; they illustrate the DWL issue using a 5.3% ratio, but their choice of parameters is not well explained.[143] Needless to say, these studies form too small a sample from which to generalize. In addition, there are many empirical studies of this issue that are concerned with market power in general, not specifically with market power resulting from cartels,[144] and one very interesting

140 *See* Easterbrook, *supra* note 23, at 455. From a theoretical standpoint, 50% is in fact the *maximum* possible percentage given a linear demand curve.

141 Lars Sølgard, Chief Economist, Norwegian Competition Authority, Speech at Seminar Hosted by the Norwegian Competition Authority: Cartel Investigations in Norway (Feb. 22, 2007), *available at* http://www.konkurransetilsynet.no/iKnowBase/Content/425749/070222_LARS_SORGARD.PDF. Four companies were convicted and heavily fined in Norway for fixing the prices of corrugated cardboard paper from 1983 to 1990. *Id.* The decision was sustained on appeal to the Supreme Court of Norway. *Id.* The chief economist of the Competition Authority favorably cites an expert opinion (apparently relied upon by the Court) that the overcharge was 70–80 million NOK and the deadweight loss was 30–40 million NOK. *Id.* Thus, the ratio was from 1.75:1 to 2.67:1.

142 Erik A. Monke et al., *Welfare Effects of a Processing Cartel: Flour Milling in Portugal*, 35 ECON. DEV. & CULTURAL CHANGE 393, 406 (1987). A careful study of total welfare effects of a government-supported cartel found that the ratio of transfer to deadweight losses was 3.6:1. *Id.* at 405 (18,456 million PTE in consumer transfers and 5150 million PTE in deadweight losses).

143 Gallo et al., *supra* note 116, at 25–71.

144 *See* John M. Connor & Everett E. Peterson, *New Estimates of Welfare Losses Due to Imperfect Competition in U.S. Food Manufacturing*, *in* AGRICULTURAL MARKETS: MECHANISMS, FAILURES, REGULATIONS 205 (David Martimort ed., 1996). The authors conclude that ten published empirical studies of the food manufacturing industries—employing a variety of data sets and methods of analysis—found that the DWL/transfer ratio was 2.5% on average but varied from 0.7% to 36%. *Id.* at 226 tbl.4. Retail food demand elasticities tend to be lower (−0.3 to −0.7) than elasticities seen in cartelized industries. *Id.* Five models based on price-leadership behavior averaged a relatively low 11% ratio. *Id.* However, these studies mostly include industries with implicit collusion and some unilateral market power. *Id.*

F. M. SCHERER & DAVID ROSS, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 667–78 (3d ed.1990), evaluates several empirical estimates of the relative sizes of the deadweight loss (0.5 to 2.0% of GNP) and transfer effects (probably at the lower end of the range of 3 to 12%) due to the exercise of market power in the whole U.S. economy in the 1950s

ratio calculated by the Canadian enforcement authorities in a merger case.[145]

Another way to determine the ratio is through the use of economic theory and logic. Many textbooks do what Judge Easterbrook did and draw diagrams that imply a ratio of 50%, but these usually are heuristic illustrations not intended to be realistic representations of markets.[146] However, economic theory produces a formula for calculating this ratio. The DWL/transfer ratio is the long-run, own-price elasticity of demand (at the collusive price) multiplied by the overcharge ratio, all of which is divided by two.[147] That is, DWL is a high share of the income transfer when the overcharge is high, and the elasticity is an absolutely large number.

Theory provides some rough guides to appropriate elasticities. We know that the elasticity of demand for products that have been cartelized is generally elastic (less than −1).[148] Following Posner's lead, a good range to consider initially is −1.0 to −2.0.[149] But we can do better by considering cartel and monopoly studies analyzing good data with the most advanced techniques.[150] We will highlight one monopoly and five cartel studies.

---

to 1970s. Economy-wide analyses tend to produce lower welfare losses than do disaggregated industry studies, but the effect on the ratio of interest is uncertain. *Id.* at 664. Despite the many caveats expressed by Scherer and Ross about these numbers, we interpret the average DWL/transfer ratio to be roughly 28%. The lowest ratio is perhaps about 8% and the highest 36%. However, these studies include many industries with implicit collusion and some unilateral market power. *See id.*

145 *See* Alan A. Fisher, Robert H. Lande & Stephen F. Ross, *Legalizing Merger to Monopoly and Higher Prices: The Canadian Competition Tribunal Gets It Wrong*, 15 ANTITRUST MAG., no. 1, Fall 2000, at 71, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1358448. The Canadian Competition Tribunal predicted that a proposed propane merger would raise prices by 8%, which came to $43 million, and also produce another $3 million in allocative inefficiency losses (a 7% ratio).

146 *See, e.g.*, SCHERER & ROSS, *supra* note 144, at 662.

147 *See* Richard A. Posner, *The Social Costs of Monopoly and Regulation*, 83 J. POLIT. ECON. 807, 816 (1975) [hereinafter Posner, *Social Costs*]. The overcharge ratio is the change in market price due to an increase in market power divided by a benchmark or but-for price. *Id.* Posner considers two types of price elasticities, one for linear demand and one for constant demand (a concave demand curve). *Id.* Constant-demand curves are most appropriate for highly differentiated products, not for the typical homogeneous products that are cartelized. Thus, we adopt the linear-demand-elasticity assumption herein. We also assume that unit costs are constant over the relevant range of output.

148 James L. Smith, *Inscrutable OPEC? Behavioral Tests of the Cartel Hypothesis*, 26 ENERGY J. 51, 53 (2005) ("[E]stimated demand elasticities numerically below −1 would constitute evidence not inconsistent with the cartel hypothesis.").

149 Posner, *Social Costs*, *supra* note 147.

150 Economists have generated thousands of empirical estimates of demand that have reliable demand elasticities. *See, e.g.*, Craig A. Gallet, *The Demand for Alcohol: A Meta-Analysis of Elasticities*, 51 AUSTRAL. J. AGRIC. & RES. ECON. 121 (2007) (compiling 132 high-quality published studies of the demand for alcoholic beverages). However, there are very few papers that contain both calculated overcharges and elasticities.

In the first cartel example, Posner calculates the DWL ratio for the first episode (1929–1931) of the global nitrogen fertilizer cartel to be 31%.[151] Second, the heavily studied Joint Economic Committee Eastern U.S. railway cartel yields DWL ratios of 26%.[152] Third, a well regarded study of the U.S. cane sugar cartel of 1890–1914 implies a DWL ratio of 12% to 13%.[153] Fourth, a 1923–1968 Norwegian cement cartel has a DWL ratio of 19%.[154] Fifth, an excellent dynamic simulation model of the U.S. lysine cartel suggests a DWL ratio of 21% to 27%.[155] In sum, five leading studies of effective cartels find that the elasticities are between −0.95 and −1.64 for effective cartels, as expected, and that the DWL ratios of 12% to 31% are strongly positively related to the overcharge rate. Finally, an impressive examination of the Alcoa U.S. aluminum monopoly during 1923–1940 concludes that demand elasticity was −2.1 and that the DWL was 62% to 66% of the income transfer.[156] The aluminum example illustrates a general finding of the cartel literature: cartels aim at achieving true monopoly power, but typically, they must settle for a weaker degree of market power.

To arrive at a reasonable DWL ratio for contemporary private cartels, we will use a 45% mean average overcharge[157] and combine it with the aforementioned −0.95 and −1.64 elasticity of demand range.[158] These parameters result in a DWL ratio of 6% to 20%.[159] Using the median overcharge of 22%[160] instead, the DWL ratio range is reduced to

151 Posner, *Social Costs*, *supra* note 147, at 820. The overcharge was 75% and the elasticity was 1.45. *Id.*

152 Glenn Ellison, *Theories of Cartel Stability and the Joint Economic Committee*, 25 RAND J. ECON. 37, 51 tbl.7 (1994) (finding an overcharge of 50.8% and elasticity of −1.59 using Model 3).

153 David Genesove & Wallace P. Mullin, *Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890–1914*, 29 RAND J. ECON. 355, 367 (1998) (computing an average annual overcharge of 13.4% and elasticity of −2.03 to −2.24 during high season; during the five most effective years, 1893–1897, the overcharge was 31.0%, implying a DWL ratio of 24% to 27%).

154 Lars-Hendrik Röller & Frode Steen, *On the Workings of a Cartel: Evidence from the Norwegian Cement Industry*, 96 AM. ECON. REV. 321, 322 (2006) (finding an overcharge of 34.5% and an elasticity of −1.47).

155 Nicolas de Roos, *Examining Models of Collusion: The Market for Lysine*, 24 INT'L. J. INDUS. ORG. 1083, 1103 (2006) (estimating an overcharge of 61.5%, and the author favors a manager's subjective notion of elasticity of −1.1 to −1.4).

156 Valerie Y. Suslow, *Estimating Monopoly Behavior with Competitive Recycling: An Application to Alcoa*, 17 RAND J. ECON. 389 (1986) (computing an overcharge of 150% and an elasticity of −2.0 to −2.1).

157 In addition to the material in this Section, this figure is based upon Connor and Lande, *supra* note 14, at 559. The literature studies' mean was 49% and the mean of verdicts was 31%. The mean for the seventy-five cartels in our study was 60.3%. *Id.*

158 *See* John M. Connor, *Price Fixing Overcharges: Revised 2nd Edition* (Working Paper Apr. 27, 2010) *available at* http://ssrn.com/abstract=1610262 (expanding and updating the study in Connor & Lande, *supra* note 14; Table 7 shows that the mean overcharge for all cartels is 46%, including many with zero price effects).

159 Applying the formula, we have 1/2 × 0.45, which is then multiplied by 1.0 or 1.65.

160 In addition to the material in this section, this overcharge percentage is based upon

3% to 10%. Combined, these alternative calculations produce range extremes from 3% to 20%.[161] That is, the allocative inefficiency associated with cartelization is between $3 and $20 for every $100 in cartel overcharges, and the "net harm to others" will be $103 to $120. Therefore, we will assume that for every $100 in cartel overcharges, there is between $3 and $20 in accompanying allocative inefficiency effects.

C. *Umbrella Effects of Supracompetitive Pricing*

When a cartel raises prices, the relevant market sometimes contains a non-colluding fringe of smaller firms that are able to raise prices due to the higher overall market price set by the cartel. Since the fringe firms did not participate in the collusion, they did not violate any law and so cannot be fined or sued successfully in a private case. Nevertheless, these "umbrella effects" are another "net harm to others" from the cartel. If a cartel raised prices by $90 million, for example, and caused the non-colluding fringe to raise prices by $10 million, the "net harm to others" from the cartel should rise to $100 million. Where this data is available, our optimal deterrence calculation takes this into account.

However, this factor might not be significant empirically, and it is likely to be difficult to ascertain, even approximately. There certainly have been powerful, if short-lived, cartels with significantly less than a 100% market share. For example, the citric acid cartel only had 60% of global production; for vitamin B1 the increase in Chinese production led to a cartel market share decline from 70% to 52%; for European industrial tubes the cartel had 75% to 85% of the market.[162] We believe, however, that effective cartels with low market shares for long periods are not common.

Including this factor explicitly in the optimal deterrence calculations could also lead to other complications. First, we cannot be sure the fringe raised prices to the same extent as the cartel. Perhaps some or all of the fringe firms decided to price somewhat lower than the cartel and thereby gain market share. Second, sometimes reports about cartel cases are not careful about market definition, and many—perhaps most—cartel cases do not contain precise market definition findings by a court. This applies both to consent orders in criminal cases and to

Connor and Lande, *supra* note 14, at 515. The literature studies' mean was 25%, and the mean of verdicts was 22%. The median for the seventy-five cartels in our study was 20%. *Id.*

161 This is a conservative resolution of the issues.

162 Iwan Bos & Joseph E. Harrington, Jr., *Endogenous Cartel Formation with Heterogeneous Firms*, 41 RAND J. ECON. 92, 92–93 (2010).

private settlements. For this reason it can be difficult to be sure which sales of non-colluding firms truly were in the same product and geographic market as the cartel. Moreover, as a practical matter almost every parameter in a consent order or private case, including market definition, is subject to a negotiation and potential compromise. No doubt, many reported cartel market shares are accurate, but there surely are other times where the size of reported relevant markets have been negotiated down or misdefined.

Although we are tempted to consider this factor in the optimal deterrence calculations through the use of an especially broad range of possible values, instead we will simply take note of this issue. We will not attempt to estimate how large cartels' umbrella effects are empirically or to take them into account in our optimal deterrence calculations.

IV. THE PROBABILITY OF CARTEL DETECTION AND PROOF OF COLLUSION

Optimal deterrence theory is concerned with the expectations of the founders of cartels as to whether any cartel they are considering forming will be detected and, if detected, proven in court to have violated the antitrust laws.[163] These individuals' predictions are formed by a variety of factors, including the perceptions and historical experience of the individuals themselves, their firms, their legal and financial advisors, and their observations of others in comparable potential price-fixing situations.[164] Since it is impossible to know the actual expectations of the "average" would-be cartelist, we instead use the closest approximations we can find: the actual record of how often cartels are detected and, once detected, proven in court to be illegal.

A. *Cartel Detection*

The first question—how likely is it that a cartel will be discovered—has been answered by researchers using three basic types of methodologies. The first is based upon quantitative economic studies. The original and most famous of these was by Bryant and Eckard.[165]

163 This subsection is based upon John M. Connor, Deterrence Power of Penalties on International Cartels (Aug. 6, 2009) (unpublished study) (on file with authors).

164 Case evidence supports the view that potential conspirators are adept at predicting the quarterly or annual profits from an effective cartel, though they might have uncertainty about the scheme's longevity. *Id.* at 9.

165 Peter G. Bryant & E. Woodrow Eckard Jr., *Price Fixing: The Probability of Getting Caught*, 73 REV. ECON. & STAT. 531 (1991). Like all similar studies, *p* is computed from samples

They estimated the confidence interval for cartels' probability of detection (p) to be 13% to 17%. Their data set consists of companies convicted for domestic U.S. price fixing during 1961–1988. This study is widely cited by scholars[166] and is approvingly cited by at least eight subsequent writers on antitrust enforcement who made their own detection estimates.[167]

Two subsequent empirical studies replicated Bryant and Eckard's approach.[168] Golub et al. sampled convicted U.S. price fixers for a period after 1988; their estimated range for p is identical with that of Bryant and Eckard.[169] Their sample includes some international cartels and a period that overlaps with the revised DOJ leniency program. Combe et al. also apply the Bryant and Eckard method of analysis to a sample of firms that were fined for infringing E.U. price-fixing prohibitions.[170] All of these convictions involved international cartels (some of them intra-E.U.), but only a small share of these infringements occurred during the time that the European Commission (EC) had adopted a formal leniency program.[171] In sum, all three studies—using different data sets—point to a probability of detection in the 13% to 17% range. The stability of p across differing time periods and jurisdictions is impressive.

Bryant and Eckard published their study in 1991, prior to the 1993 advent of the DOJ's wildly successful cartel leniency/amnesty programs which have in some form been adopted by more than twenty jurisdictions, including the European Union (EU).[172] The vast increase in *numbers* of cartels detected since 1993 could be due to an increase in the probability that cartels are detected. In a highly original paper, Miller provided an economic estimate of the post-1993 increase in the probability that cartels will be detected by the DOJ.[173] His sample

of discovered cartels. Founders of never-discovered cartels might rationally conjecture a lower p. Thus, computed sizes of p may well overstate the actual average p for all cartels.

166 A Google Scholar search on February 9, 2011, found fifty citations.

167 *See infra* Appendix tbl.3.

168 All three use essentially the same method—an event study of stock market prices—to estimate a statistically calculated 90% confidence interval of the probability of cartel detection (p). However, the three apply that method to three different samples from two jurisdictions.

169 Alla Golub et al., *The Profitability of Price Fixing: Have Stronger Antitrust Sanctions Deterred?*, (2005) (presented before the International Industrial Organization Conference 3, Atlanta, Ga. (Apr. 8–9, 2005)), *available at* http://ssrn.com/abstract=1188515.

170 Emmanuel Combe et al., *Cartels: The Probability of Getting Caught in the European Union* (Bruges Eur. Econ. Res. Papers, Working Paper No. 12, 2008), *available at* http://ssrn.com/abstract=1015061.

171 Their point estimate of p is close to 13%.

172 *See* Scott D. Hammond, Deputy Assistant Att'y Gen. for Criminal Enforcement, Antitrust Div., U.S. Dep't of Justice, Address Before the 24th Annual National Institute on White Collar Crime: The Evolution of Criminal Antitrust Enforcement over the Last Twenty Years (Feb. 25, 2010), *available at* http://www.justice.gov/atr/public/speeches/255515.htm.

173 Nathan H. Miller, *Strategic Leniency and Cartel Enforcement*, 99 AM. ECON. REV. 750 (2009).

consisted of all cartels discovered and convicted by the DOJ between January 1985 and March 2005. Comparing the pattern of pre-1993 cartel enforcement with the post-1993 period, he estimates that there was an increase of about 60% in the detection of existing cartels and a reduction of about 60% in the rate of cartel formation.[174] A possible limitation of Miller's study is that, in his sample, only 9% of the observations were international cartels.[175] Nevertheless, if one applies Miller's findings to the earlier three detection-probability studies, the post-1993 range for the probability of cartel detection becomes 20.8% to 27.2%.

A completely different method of estimating the probability that cartels are detected relies on the opinions of cartel scholars. Most have legal training or write in legal-economic publications.[176] Many have prosecutorial experience; others have worked extensively with alleged cartel defendants.[177] Those who have provided specific estimates are listed in Table 3.[178] The opinions and conclusions of these twenty-five authors predominantly suggest a 10% to 25% chance of detection, although some go as high as 33%.[179]

It is clear that some of these estimates are meant to be purely illustrative,[180] while others are from surveys or are intended to be true depictions of reality.[181] The three writings that are clearly illustrative average 29%.[182] If one takes the non-illustrative estimates and eliminates those that depend on Bryant and Eckard, the remainder are independent estimates. For the ten independent estimates that are not purely illustrative, the upper-end estimates average 25.6%, which is comfortably close to the economists' 27% high estimate.[183]

There is yet another way to estimate the average detection probability—opinion surveys. Although these surveys might not ask precisely the questions that are best for our purposes, they too suggest

174 *Id.* at 760–61.

175 As we understand these results, both changes are simultaneous after 1993. To illustrate, suppose that there are 100 cartels being formed that affect the U.S. economy each year in the years before 1993. With a known median life of seven years and no enforcement, the total stock of prosecutable cartels would reach a steady state of 700 cartels. With discovery of 15%, then a net formation of 85 lasting seven years would imply discoveries of 15 per year and a stock of 600 hidden cartels. Then, using Miller's results, with amnesty the number formations drops to 40 per year or 280 total cartels, of which about 70 are discovered per year and 210 are hidden in any given year. Thus, deterrence improves (fewer net formation and fewer hidden cartels), and detection rates per year also rise.

176 *See infra* Appendix tbl.3.

177 *See id.*

178 *See infra* Appendix tbl.3.

179 *See id.*

180 *See* Landes, *supra* note 15, at 656.

181 *E.g.*, Alan R. Beckenstein & H. Landis Gabel, *Antitrust Compliance: Results of a Survey of Legal Opinion*, 51 ANTITRUST L.J. 459 (1982).

182 *See* POSNER, *supra* note 65, at 47; Landes, *supra* note 15; Werden, *supra* note 20, at 27–29.

183 *See infra* Appendix tbl.3.

low cartel detection rates. For example, in the survey by Feinberg of antitrust lawyers working in Brussels, only 5% disagreed with the statement, "[t]he [EC] fails to detect most [price-fixing] violations," whereas 62% agreed with the statement.[184] A large-scale 2006 survey of competition lawyers working in the United Kingdom (UK) and Brussels asked how many times one of their clients had, upon seeking legal advice, abandoned or changed a possible cartel practice because the clients feared an antitrust investigation, and how many of their clients had been the subjects of an adverse cartel ruling by the UK's Office of Fair Trade. The result was that 22% were said to have been in violation of cartel laws.[185] This is, of course, a minimal indicator of detection because some participants in secret cartels do not seek legal advice.

Professor Daniel Sokol recently conducted another very interesting survey.[186] He asked a sample of 234 antitrust lawyers,

> In the past 2 years, by total number of matters, how often have clients come to you with hard-core cartel issues that to your and/or their knowledge never got investigated by U.S. government (federal and state) enforcers as opposed to situations where the underlying behavior ultimately led to U.S. investigation of your client?

If the "Not Applicable" responses are eliminated, 52% of the lawyers said this had happened to them at least once.[187]

All told the above methods yield estimates for *p*: 1) 20.8% to 27.2%, 2) 25.6%, and 3) non-quantifiable but low estimates that are roughly consistent with the first two estimates. In the interest of being conservative, for the remainder of this Article we adopt a relatively high 25% to 30% probability that cartels will be detected.[188]

---

184 Robert M. Feinberg, *The Enforcement and Effects of European Antitrust Policy: A Survey of Legal Opinion*, 23 J. COMMON MKT. STUD. 373 (1985). Other interesting results were: 1) 95% agreed that price fixing was intentional and for profit gain, and 2) 100% agreed that the greatest deterrents are a high probability of detection and high EU fines. *Id.*

185 DELOITTE & TOUCHE LLP, THE DETERRENT EFFECT OF COMPETITION ENFORCEMENT BY THE OFT (2007), *available at* http://www.oft.gov.uk/shared_oft/reports/Evaluating-OFTs-work/oft962.pdf.

186 *See* Sokol, *supra* note 28.

187 *See id.* at 239 tbl.14.

188 We believe our methodology has been overly conservative and that the actual chances a cartel will be detected are lower than 25–30%. As an indication of how conservative our methodology is, Ginsburg and Wright recently performed an analysis very similar to ours, including analyzing both the Bryant and Eckard, as well as the Miller studies, and concluded that 25% was their best estimate as to the rate of cartel detection. *See* Ginsburg & Wright, *supra* note 25, at 8.

B. *Probability a Detected Cartel Will Be Convicted*

Even if a cartel is detected, its chances of being convicted are less than 100%. The DOJ asserts that in 95% of its cases, indictments end in convictions.[189] Indeed, the evidence is so damning in most cases that nearly all defendants negotiate a guilty plea.[190] On the other hand, when accused individual price fixers choose to litigate a criminal price-fixing case, the government wins only approximately half the time.[191] Thus, discovered cartelists that are able to afford the best legal defense team and are adept at hiding or obfuscating the most incriminating evidence might well judge their chances of conviction to be less than the DOJ's 95% figure.[192]

From 2005 to 2009, of the 87 individuals charged with international price fixing, 64 pled guilty and 4 were found guilty.[193] On the other hand, 7 were acquitted, 11 became fugitives, and 1 indictment was dismissed.[194] Therefore, in total, from 2005–2009, 68 of 87 (78%) were convicted. For the entire 1990–2009 period the corresponding figure is 158 of 222 (71%).[195] Therefore, a high estimate of how often detected cartelists escape conviction would be the 22% to 28% who were not convicted in DOJ proceedings.

However, some or all of the non-convicted defendants could have been innocent. Others could have been guilty, but perhaps the DOJ simply could not prove their guilt sufficiently to meet the high standards for felony convictions. There is no way to know how many of those who were not convicted actually formed a cartel, and that this cartel was detected, but they nevertheless got away with their crime. At a minimum, however, we believe we can fairly make a presumption concerning the fugitives from prosecution. A total of 11 of the 87 defendants from 2005–2009, and 47 of the 222 from 1990–2009, were fugitives.[196] We believe it is reasonable to presume that it is more likely

189 *See* WORKLOAD STATISTICS 1990–1999, *supra* note 90, at 7; WORKLOAD STATISTICS 2000–2009, *supra* note 90, at 8.

190 Connor, *supra* note 14, at 328.

191 *Id.* (finding that only fifteen of twenty-eight indicted individuals were convicted).

192 *See id.*

193 *See* Connor, *supra* note 14, at 539 tbl. 3. The Antitrust Division's official statistics, reported *supra* note 25, cannot, however, be used to derive comparable won/lost ratios for domestic cases. For the 1990–2009 federal fiscal years, we can determine that there were 929 individuals indicted for Sherman Act section 1 criminal offenses; of those, 57% were fined, 38% were imprisoned, and 28% were subject to other forms of confinement. But these three types of sanctions are not additive. While nearly all those who were imprisoned were also fined, we cannot determine what proportion of those fined were also imprisoned or otherwise confined. Therefore, the DOJ does not trumpet the number of fugitives.

194 *Id.*

195 *Id.*

196 *Id.*

that a fugitive is a price fixer who fled, rather than an innocent person who could not prove their innocence.[197] Therefore, on this basis there is (using data from the two time periods) a 47/222 = 21%, or 11/87 = 13%, chance that detected price fixers will get away with their crime.[198]

By contrast, the DOJ reports that from 2005–2009 they won 124 cases against corporate and individual defendants, mostly through plea agreements, and lost seven.[199] This is a 95% success rate; much higher than their 77% success rate for the same period when prosecuting individuals. Does this mean that the corporations are significantly more likely to be convicted than individuals? Yes, if one counts any corporate fine at all as a government "success." However, we cannot help but wonder whether every DOJ "win" is truly a win. Almost all of the DOJ "wins" are plea agreements or consent orders. No doubt, many corporate or individual defendants simply agree to a "slap on the wrist" consent order rather than endure the significant legal expenses and corporate disruption involved in taking the DOJ to court. No doubt many of the token DOJ "wins," which secured only minimal fines and no prison time, were really defendant victories. Ideally we would find and use in our calculations the percentage of detected cartels that not only were convicted, but that also received significant sanctions. Unfortunately, there is no way to tell which of DOJ's alleged "wins" are truly wins, and which ones mostly, in reality, should instead be categorized as being DOJ losses.[200]

It seems likely, however, that individuals are less likely to plead guilty even to a token fine than are corporations. Corporations might readily agree to a "slap on the wrist" fine as part of a settlement with the DOJ because to them small fines are almost like parking tickets, and some large corporations receive similar "costs of doing business" frequently.[201] Moreover, corporate managers are paying fines with other people's (i.e., the stockholders') money. As a matter of ethics they would deny this makes a difference, but unless the corporate officer owns a large share of a company's stock, the principle/agent literature suggests

197 Innocent people sometimes flee. This is why one can only presume that fugitives actually fixed prices.

198 These figures and ratios are for individuals, not for corporations, and most of our sanctions are corporate, not individual. We will, however, assume that the conviction rates for individuals apply to corporations, as well.

199 *Id.*

200 Perhaps one should draw a very low arbitrary line, such as making the assumption that any DOJ fine (and private settlement, as well) for less than 1% of the volume of commerce involved was "really" a defendant victory. Or perhaps one should classify these settlements into groups, such as 0–1% of affected commerce, 1–3%, 3–6%, etc., and then we could argue over the point at which the settlements are likely to be genuine victories.

201 There are exceptions, of course. Corporate felony convictions can bar a firm from bidding for federal contracts for a number of years, and this could be a major blow to firms that depend on such sales for a significant portion of their revenues.

he or she is more likely to let personal motives affect what is best for the owners.[202] By contrast, an individual has more to lose and may be more risk-averse. If an individual admits to a felony, even one resulting only in a small fine, their personal record has been stained, perhaps with dramatic results for the person involved. For these reasons, individuals are, on average, more likely to resist than a large corporation. If we are correct, the DOJ's conviction rate for individuals is a better reflection of the DOJ's real won/lost record than the corporate statistics.

We also believe that the DOJ's 95% conviction rate indicates that the Antitrust Division is risk-averse, and usually indicts only when it has a relatively large chance of conviction in the event an alleged price fixer insists on a trial. There are a number of times, for example, when the DOJ began a cartel investigation, but never filed an indictment, yet private plaintiffs secured a significant settlement against these same corporations.[203]

To be conservative, however, we will base our final conviction estimate on the statistics for individual convictions, and assume that 23% to 28% (high estimate) or 20% to 23% (low estimate) of detected cartels are not convicted. In our final calculations we will round these numbers slightly downwards, to 20%.[204] Note that the probability of a cartel being detected (25% to 30%) *and* convicted (80%) then becomes 20% to 24% (depending on whether low or high estimates are used).

V. THE OPTIMAL DETERRENCE CALCULATIONS

As noted in Part I, under the optimal deterrence approach, cartel sanctions should be equal to:

> *Net harms to others ÷ (Probability of detection × Probability of conviction)*

202 This topic, also studied under the titles "managerial capitalism" or "managerial utility," is reviewed by Alan Hughes, *Managerial Capitalism*, *in* 3 THE NEW PALGRAVE: A DICTIONARY OF ECONOMICS 293–95 (John Eatwell et al. eds., 1987).

203 *See, e.g.*, *In re* Automotive Refinishing Paint Antitrust Litig., 177 F. Supp. 2d 1378 (E.D. Pa. 2001); Robert H. Lande & Joshua P. Davis, *Benefits from Antitrust Private Antitrust Enforcement: An Analysis of Forty Cases*, 42 U.S.F. L. Rev. 879 (2008), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1090661 (noting government investigation, but no government case filed; private recovery of $106 million in cash).

204 We believe that the 20% estimate (which means that 80% of detected price fixers are convicted) substantially understates the probability that detected individuals or corporations will escape conviction. However, the only reliable data on this point we have been able to find concerns the fugitives, so we do not feel comfortable assuming, for example, that only 50% of detected cartels are convicted.

We also note that DOJ's risk aversion and the fact that many of their wins are only token victories probably mean that the 20% figure we selected probably underestimates the percentage of detected cartels that truly escape punishment for their crime.

We have been able to ascertain approximations for each of the required quantities for seventy-five cartels that have been sanctioned in the United States since 1990.[205] We illustrate how we carried out the optimal deterrence analysis and calculations using the lysine cartel as an example.

A. *The Lysine Cartel as an Example*

1. Background on the Lysine Cartel

The lysine cartel was one of the earliest large international cartels to be heavily sanctioned in multiple ways.[206] It dated back to mid-1992. The U.S. Department of Justice began an investigation in late 1992 that culminated in a June 27, 1995 raid, where more than seventy FBI agents simultaneously raided the headquarters of Archer-Daniels-Midland Company (ADM) and a number of ADM officers' homes.[207] Within a very short time, investigators had also raided the offices of four other companies that manufactured or imported lysine.

During this cartel's existence the average manufacturers' delivered price of lysine in the United States rose from $0.68 per pound to a plateau of $0.98 (October to December 1992), fell again to $0.65 (May 1993), and rose quickly again to above $1.00 for most of the remainder of the conspiracy period.[208] Early in this cartel's existence an ADM Vice President was caught on tape saying that their recently concluded agreement would generate $200 million in joint profits annually in a global market for lysine that generated from $500–700 million in annual sales.[209] His prediction turned out to be astonishingly accurate.[210]

---

205 Although we started with a larger universe of cartels, we were forced to eliminate many from our sample because the necessary data was not available, was insufficiently reliable, or some legal actions were unresolved. Every one of our final group is an international cartel. Although we are unable to state with certainty that all of the assembled data on these seventy-five cartels are perfect and complete in every respect, we believe all of it to be generally reliable and accurate. As an example of its potential inadequacy, although we looked diligently for settlements in private cases and believe we found every significant settlement, there surely have been settlements that we missed, especially secret settlements and opt-out settlements too small to have made the legal, general or trade press. By contrast, class action settlements usually cannot be secret and almost always are reported in the legal, general, or trade press.

206 *See* Connor, *supra* note 56. As will be apparent from the Conclusions, *infra*, the lysine cartel was one for which actual sanctions were relatively close to the optimum.

207 This Article's analysis of the lysine case is based upon John M. Connor, *Global Cartels Redux: The Lysine Antitrust Litigation*, *in* THE ANTITRUST REVOLUTION 300, 300 (John E. Kwoka, Jr. & Lawrence J. White eds., 5th ed. 2009).

208 *Id.* at 12.

209 *Id.* at 13.

210 *Id.*

Ultimately the lysine cartelists pled guilty, and in late 1996 incurred U.S. fines that totaled $95.55 million.[211] The DOJ also prosecuted four lysine executives in a highly publicized jury trial held in Chicago in the summer of 1998.[212] Three of the four were found guilty and were heavily sentenced, to a total of ninety-nine months in prison.[213] The fourth defendant, a managing director of Ajinomoto of Japan, remains a fugitive.[214]

Within a year of the FBI raids, more than forty civil antitrust suits were filed in U.S. federal courts.[215] Approximately 400 plaintiffs were certified as a single federal class of direct purchasers, and in July 1996, the federal class in *Amino Acid Lysine Antitrust Litigation* settled with the three largest defendants for $45 million.[216] The two other defendants settled for almost $5 million about a year later.[217] There also were significant numbers of indirect purchaser suits and opt-out suits which have been very difficult to trace, but these payments have been estimated to total more than $25 million, and to produce total payments in the U.S. private suits of approximately $82.5 million.[218]

2. Optimal Fine Calculations for the Lysine Cartel

What should the overall level of sanctions have been, ex-ante, for the Lysine cartel?[219] Before one could calculate this using the "net harm to others" approach, however, it is necessary to account for inflation or the time value (opportunity cost[220]) of money.[221] Because we are

211 *Id.* This includes $94.3 million in corporate fines and $1.25 million in individual fines, which we tripled to give more weight to individual sanctions relative to corporate sanctions. For a discussion of this tripling, see *supra* Part II.

212 Connor, *supra* note 207, at 1.

213 *Id.* The cartel also was fined by the antitrust authorities of Canada, Mexico, and the European Union a total of at least another $121.5 million. *Id.* at 2.

214 *Id.* at 2.

215 *Id.*

216 *Id.* The settlement was approved in late 1996, before the federal fines were announced, which is very unusual. *Id.*

217 *Id.*

218 *Id.*

219 This number is only illustrative because society must as a practical matter focus upon general deterrence, not specific deterrence. We could never hope to know the mindsets of particular corporate executives well enough to calculate the penalty that optimally would prevent those individuals from cartelizing, the most we can do is to calculate a good overall, general deterrence penalty and then implement it generally. For an analysis of these issues, see *supra* Part I.

220 "Opportunity cost" is a fundamental economic concept positing that the value of any economic choice actually made is approximately equal to the next best alternative course of action not taken. For example, the value of an afternoon's leisure to an individual might be approximated by the income foregone in employment. Similarly, the cost of consuming for a household today might be the future income from investing the same amount in some financial instrument.

attempting to determine how much purchasers were harmed by paying supra-competitive prices for their products or services, we should analyze the opportunity cost issue from the victims' perspective and attempt to place the victims in the position in which they would have been had no violation occurred. Adjusting for the time value of money can raise the amounts involved significantly, especially when there is a long lag between the collusive period and fine or the court approval of a settlement.

It is impossible to know what would have happened to the overcharges had the violation not occurred. Consumer-victims or business-victims might have invested the overcharges they were forced to pay to the cartel in, for example, Treasury bills. Alternatively, suppose a victim had been harmed and believed it would recover from defendant in *X* years. A reasonable course of action for that victim might be to obtain an *X* year loan for the amount of the damages at the prevailing consumer loan or business loan interest rates, both of which would exceed the prime interest rate. Moreover, since the overcharges were involuntary (and illegal!), it would be fair to resolve doubts over the correct adjustment rate in favor of the victims. The members of the cartel, by contrast, might have invested the overcharges instead.

A conservative approach to these issues consistent with principles of financial economics is to approximate the opportunity cost to the victims of being deprived of their money for a period of time by using the prime rate of interest plus one percent.[222] For simplicity, we will use as our initial year the midpoint year of the cartel,[223] and as the fine year the year in which the first corporate defendant plead guilty.[224] The

221 Neither fines nor payments made in private cases contain prejudgment interest. However, once a private case results in a verdict or a court-approved settlement, post-judgment interest begins to accrue. *See* Lande, *supra* note 17.

222 The prime rate of interest includes a component that anticipates what lenders expect inflation to be over the loan period. Another portion of the prime rate is an average low-risk rate of return to be earned by borrowers. The one percent is added to account for the fact that borrowers expect to earn profits on the investment above a low-risk investment rate of return.

223 This approximates the mean date that buyers' funds were transferred to the owners of the cartels. If we had the data we would instead assess the magnitude of the cartel overcharges on a yearly basis, and would separately take into account the date of each of the imposed fines and settlements with each cartelist, and make the adjustments accordingly. This would be slightly more accurate because cartels do not overcharge the same percentage every year, and because some fines and some settlements—particularly opt-out settlements—take place years later. As a practical matter, however, we rarely have the necessary information. We do, however, have good information concerning the starting and ending dates for all seventy-five cartels in our sample.

Normally, overt collusion stops on the date subpoenas are served or inspections are carried out by an antitrust authority. In some cases collusion may have stopped years earlier. Other times the firms continue implicit collusion even after the explicit collusion is uncovered and the formal (proven) collusion ends.

224 This too is conservative, for two reasons. In some cartel cases the late-pleading participants take a year or two to plead after the first defendant does so. Second, defendants

terminal year for settlements in private suits is the year in which the federal class settlement or other private case receives preliminary judicial approval.[225] Although this approach is perhaps too low and thus too conservative from the "net harm to others" perspective, it does have the advantage of approximating the value of the overcharges to the cartelists, who of course continue to have use of the victims' money interest-free until they pay their fines or damages in private suits. After the net present value of the fines or settlements is calculated, we adjust the value of money, due to general inflation, to the year 2010, employing the annual Producer Price Index calculated by the U.S. Bureau of Labor.[226] Expressing all penalties in 2010 dollars permits us to make meaningful comparisons across conspiracies and punishments that took place at different times.

If we restrict all data and calculations to the United States, for the lysine cartel the optimal penalty ((net harms to others) ÷ (probability of detection × probability of proof)) can be calculated as follows:

- The net harms to U.S. direct purchasers were $80 million, expressed in 1993–1995 dollars.[227] To apply the "net harm" or investment-opportunity-cost adjustment, we use Federal Reserve Bank prime rates of interest for the years 1995 and 1996, plus 1%, or 12.22%. Thus, the damages were $80 million,[228] which is the sum that the victims ought to have received when the cartel operated, and is equivalent to $119.8 million in 2010 dollars.
- These overcharge figures should be multiplied by 1.03 to 1.20 to account for the additional allocative inefficiency harms (deadweight loss) of market power.[229]

increasingly pay their fines in up to six installments spread over five years. Thus, by using the initial fine date we are over-inflating the effect of fines to some extent. But this assumption makes the calculations simpler.

225 This date is conservative because in many instances there are opt-outs from the primary class, and opt-out suits typically take months or years to negotiate beyond the class approval date.

226 *See CPI Inflation Calculator*, *supra* note 115. We use the Producer Price Index for intermediate materials, rather than the Consumer Price Index, because most cartelized products are inputs sold to manufacturers. If we had used the Consumer Price Index, however, the results would be similar.

227 *See* Connor, *supra* note 207, at 302.

228 The actual overcharge amount is $80 million. To this should be added foregone profit of $9.8 million which should have accrued between the dates of the actual overcharges and 1996. Another way of looking at the $9.8 million is that it represents income to the cartelists on the $80 million in illegal monopoly profits held in the companies' treasuries. By rights, this income belonged to the victims all along. This total of $89.8 million is the figure that we convert to 2010 dollars.

229 For an explanation of the allocative inefficiency adjustment, see *supra* Part III.B.

- The average probability of cartel detection, the evidence shows, is 25% to 30%.[230]
- The average probability the enforcers will be able to prove that the cartel violated the law has been estimated to be 80%.[231]

Therefore, for the Lysine cartel, the optimal U.S. penalty (in millions of 2010 dollars) was:

($119.8 × 1.03) ÷ (0.30 × 0.80) (low estimate)
or
($119.8 × 1.20) ÷ (0.25 × 0.80) (high estimate)

= $514–719

The optimal penalty should be compared to the actual U.S. sanctions that were imposed on the Lysine cartel. When expressed in terms of millions of 2010 dollars they were:

$114	Fines (converting $98.55 million in fines in 1996–2010 dollars)
+ 99	Private Suits (converting $82.5 million in recoveries in 1996–2010 dollars)[232]
+ 50	Prison-Equivalent for ninety-nine months of U.S. prison time at $500,000 per month[233]
$263	Total Sanctions

Thus, even though the lysine cartel was heavily sanctioned in the United States in three ways (by fines, prison for top executives, and by private litigation), the combination of the sanctions that were imposed is only 37% to 51%[234] as large as the overall amount of sanctions that should have been imposed from the perspective of optimal deterrence.

230 *See supra* Part IV.A.

231 *See supra* Part IV.B. Another issue concerns the distinction between "technical" convictions and "real" convictions. Some of DOJ's reported convictions may be technical convictions that amounted only to "slaps on the wrist" and produced only token fines. Perhaps, we should have attempted to find and use in our calculations the percentage of detected cartels that not only were convicted, but that also received significant sanctions. Because of the subjectivity of classifying fines this way, we did not, however, attempt to make this distinction.

232 Only the first settlement was in 1996, but to be conservative we assumed that all of the payments in every private case were made in 1996.

We of course can only count settlements known to us through our searches of the legal and general media. We readily acknowledge the existence of secret settlements, especially involving opt-out cases. However, every class action must be approved by a court, so no class action settlement can be secret. Publicly traded corporations often are required to report significant income or losses on their balance sheets and cannot, for example, simply state in its annual report that it paid or received a significant, but secret, sum in an antitrust case. Still, we surely missed some settlements.

233 For the analysis of the monetary equivalent of prison sentences, see *supra* Part II.

234 Depending upon when and how the figures involved are rounded, this range could also be expressed as 37% to 49%.

B. *Calculating Overall Optimal Deterrence Using Every Cartel in Our Sample*

We have undertaken the same analysis for all seventy-five cartels for which we have been able to ascertain the necessary data.[235] The overall results show that, on average, the value of the imposed U.S. sanctions has been much less than they should have been for society to obtain optimal deterrence against cartelization. If mean average figures are used, the total value of the imposed sanctions were only 15.8% to 20.8% of their optimal level. If median figures are used, the imposed sanctions averaged only 9.2% to 12.1% of optimality.[236]

One outlier, E-Rate Federal Internet Program, may have been sanctioned more than the optimal amount (our results show 125% to 175% of optimality).[237] A second cartel, PVC Windows Coverings, was probably optimally sanctioned (we estimate 88% to 124%). The other seventy-three were sanctioned much less than optimally. Moreover, half of the seventy-five were sanctioned less than 10% of the optimal amount. It certainly is possible that some of the individual firms in the

235 Data employed and calculations are available at http://www.cardozolawreview.com/content/34-2/Connor.Lande.34.2/AntitrustStudyRawData.pdf.

236 These results might, moreover, be too high for a methodological reason we have not yet discussed: for a variety of reasons, many of our sales figures might be overly small. The correct sales data would tend to lower the calculated ratios. This is because affected sales figures derived from seemingly reliable sources often are larger than the sum of the affected sales employed by the DOJ in sentencing the members of cartels. There may be quite defensible reasons for this. For example, because of the high degree of reliability of evidence needed to convict corporations for crimes, the DOJ may reduce the time periods, geographic region, or scope of products employed for calculating sales during collusion to that which can be proven "beyond a reasonable doubt." On the other hand, prosecutors sometimes may uncritically accept arguments made by defendants that diminish the scope of the affected market because of time pressures in settling guilty plea agreements, or because the government lacks the resources necessary to disprove defendant assertions.

An example is the *Central Indiana Ready-Mix Concrete* case. *In re* Ready-Mixed Concrete Antitrust Litig., 261 F.R.D. 154 (S.D. Ind. 2009). Concrete for pouring is a relatively simple product; the counties involved and the time period were not issues in the case. A sales figure of $680 million for all seven firms involved in the cartel was reported in the local press; all seven paid civil settlements. The sales information purportedly came from transcripts of a jury trial of two executives (they were convicted) and from the testimony of the plaintiffs' class expert in fairness hearings (plaintiffs prevailed). Sales according to DOJ documents were much less. One participant was granted amnesty; two others were not charged, most likely because of cooperation agreements. The DOJ used a smaller geographic market definition than for civil plaintiffs. When one adds up the affected sales from the DOJ sentencing memoranda for the four companies that were criminally convicted of price fixing through plea agreements, the total is $391 million. Taking into account the fact that two of the smallest cartel members were not convicted because of bankruptcies, the DOJ's total market affected sales is as much as 40% lower than the affected sales proven by the private litigants. *See* E-mail from John Connor to Scott Gilchrist, Attorney, Cohen & Malad, LLP (Aug. 24, 2011, 10:25 AM) (on file with author).

237 This cartel was unusual for many reasons, including its record-breaking number of incarcerations. Moreover, because the affected sales of several school-district bids are unavailable, we believe that the total affected sales is significantly underestimated.

seventy-five cartels were optimally or excessively sanctioned due to circumstances unique to those firms.[238] From a deterrence perspective, however, would-be cartelists are unlikely to focus upon outliers rather than the norm. They are much more likely to be guided by what happened on average to the vast majority of cartels that affected the roughly $1 trillion in affected sales (about $2 trillion in 2010 dollars) in the cases covered by our study.[239]

Our analysis is confined to effects within the United States. For each cartel, only United States overcharges, sales, corporate and individual fines, restitution payments, prison and house arrest time, and payouts in private cases were considered. For some of these cartels, particularly the more recent ones, the European Commission's fines have been as large as or larger than those in the United States.[240] If managers were assessing whether to form an international cartel, their probable overcharges in Europe, as well as the E.U.'s sanctions, should, of course, have been considered in addition to those imposed by the United States. It is indeed unfortunate that, regardless what they might conclude about the expected profitability of operating in Europe or

238 Even if individual firms appear to have been sanctioned more than the amount calculated under the overall optimal deterrence approach, this could have been due to a number of factors that make the sanctions not excessive. Fundamentally, every firm in a cartel is jointly responsible for entirety of the cartel's overcharges. For this reason, it would be reasonable to attribute the entirety of a cartel's overcharges to an individual cartel member before carrying out the optimal deterrence calculations (although we have not done this in this Article). Only if this were done and the optimal deterrence calculations showed that the sanctions were excessive could there be true over-deterrence.

Moreover, the alleged over-deterrence could result from a cartel not producing profits as high as its instigators had hoped. Perhaps if the cartel had been as profitable as its planners had hoped, the overall penalty level might have been too low. Further, we used reported or provable affected sales in our calculations. As noted *supra*, note 234, reported or provable sales often are lower than the true amounts.

As we noted in Part I, the overall level of sanctions cannot be set, in advance, for particular individuals or corporations. The best we can do is to set the overall sanctions level for mean or median cartels, not for the outliers.

239 One interesting factor that helped drive these conclusions is the relatively small effect of prison sentences. Their mean value per case was a relatively modest $13.6 million, or 17% of the average fine (the median is zero because for the majority of the cartels in the sample (forty-eight out of seventy-five) there was no imprisonment). *See supra* note 231. Even though we valued the deterrence from a three-year sentence at $18 million (which is more than most estimates of the value of an entire life), this pales in comparison to the possible rewards from cartelization. *See supra* note 119. Nevertheless, the absence of a criminal sanction correlates with an exceedingly small overall sanction. Almost all of the fifteen cartels with actual sanctions that were less than 2% of optimal penalties had no criminal sanctions imposed. *See supra* Part IV.B. The absence of a criminal conviction means that the private sanctions cannot come close to providing optimal sanctions. By contrast, the E-Rate cartel case involved 626 months worth of prison, which constituted 85% of the sanctions in that case. For this data, see the online appendix, *Antitrust Study Raw Data*, at http://www.cardozolawreview.com/content/34-2/Connor.Lande.34.2/AntitrustStudyRawData.pdf.

240 *See* John M. Connor, *Has the European Commission Become More Severe in Punishing Cartels? Effects of the 2006 Guidelines*, 32 EUR. COMPETITION. L. REV. 27 (2011).

elsewhere, the combined level of U.S. sanctions are woefully inadequate to deter them from operating in the United States.

Recent developments have not negated the policy import of our results. For example, criminal fines and prison sentences have risen since the mid-2004 Antitrust Criminal Penalty Enhancement and Reform Act (ACPERA) amendment went into effect.[241] A GAO report on ACPERA shows that total criminal fines have risen by 51%, on average, and total jail time by 56% since ACPERA went into effect.[242] But these increases could well be explained by higher affected sales of cartels that colluded after 2004. Moreover, the GAO data refers to fines corrected for inflation on all cartels, both international and domestic, with fiscal years 2005–2010 being compared to 1994–2003. However, for international cartels over a comparable period we find that even though real fines did increase, real settlements and the value of prison declined so much that penalties per cartel declined by 38%.[243] The explanation for this overall decline is that private settlements are, on average, the largest sanction in terms of the magnitude of their deterrence effects.

CONCLUSIONS

"If three is the wrong number, it is too small."

– *Judge Frank Easterbrook*[244]

The primary goal of this Article has been to determine whether the overall level of U.S. anti-cartel sanctions is optimal. This Article demonstrates that when the deterrence effects of every measurable sanction are considered (including corporate and individual fines, payments in private cases, restitution payments, and an allowance for incarceration), the overall level of anti-cartel sanctions is far too low. To protect victims optimally, the collective level of existing sanctions should be multiplied by a factor of five. Specifically, we find that on average the total value of imposed sanctions have been only 9% to 21% as large as they should have been.[245] In other words, only if, on average, cartel sanctions were approximately five times as large as they are today,

241 *See supra* note 2.

242 *See* U.S. GOV'T ACCOUNTABILITY OFFICE, CRIMINAL CARTEL ENFORCEMENT: STAKEHOLDER VIEWS ON IMPACT OF 2004 ANTITRUST REFORM ARE MIXED, BUT SUPPORT WHISTLEBLOWER PROTECTION 21–22, 24 (2011), *available at* http://www.gao.gov/products/GAO-11-619.

243 *See id.* at 59–62.

244 *See* Easterbrook, *supra* note 2, at 95.

245 If mean figures are used, the total value of the imposed sanctions has been only 15% to 21% of the optimal level. If median figures are used, the imposed sanctions averaged only 9% to 12% of optimality.

and if these higher amounts were imposed by the courts on price fixers,[246] would consumers be optimally protected from becoming cartel victims.

To arrive at this conclusion we made many assumptions and estimates. As noted throughout this Article, we believe that every time we made necessary assumptions and estimates we chose alternatives that were conservative (i.e., they would tend to increase the relative size of the imposed sanctions relative to their optimal level).[247] Similarly, as noted, we have attempted to ascertain every relevant piece of data for every cartel in our study as accurately as possible.[248] Nevertheless, even if some of our assumptions or estimates are off, or if some of our cartel data is inaccurate, our conclusion that sanctions should be increased at least fivefold is quite robust. It is unlikely to be wrong by very much. It is very unlikely that the overall existing level of sanctions only should be doubled.[249]

One of our controversial assumptions was to value the deterrence effects of a year in prison or under house arrest as the equivalent of a $6 million sanction. We readily admit this figure is arbitrary and that reasonable people could select a different amount. Although we believe $6 million is more than the average that a year of confinement should be valued at, one could argue that in light of how hard people try to avoid prison, how much defendants spend in legal fees to avoid prison, how wealthy many price fixers are, and how time spent in prison might lower individuals' future income and social status, we should be using a significantly higher figure.

However, even assuming a year in confinement produced the deterrence equivalent of $12 million or $24 million would not change our conclusions significantly. Even the assumption that a year of confinement produced $365 million in deterrence would not mean that existing sanctions are adequate. Only if a year of confinement were

246 It is possible, however, that some courts might find ways to avoid imposing dramatically higher sanctions. For example, courts might not want to impose prison sentences five times as high as the current ten-year maximum sentence for price fixing. As a practical matter courts might be able to find ways not to do so.

247 Similarly, to conservatively assess whether the current overall levels of sanctions are optimal, we used full or high estimates of the sizes of existing sanctions at every opportunity. By contrast, an Article dealing with related topics, Lande & Davis, *supra* note 14, made low assumptions about the recoveries from private cases, a methodology that tended to understate the magnitude of the benefits from private litigation.

248 Complications include the fact that many of the cartels at issue cover more territory than the United States, and that it is difficult to disentangle U.S. effects from transnational effects.

249 An additional factor must, moreover, be considered whenever a cartel is international in scope: Fines and private damages actions brought under the U.S. antitrust laws reflect only purchases made by buyers in the United States. *See* F. Hoffmann LaRoche Ltd. v Empigran S.A., 542 U.S. 155 (2004). If a significant percentage of the cartel's sales and profits are generated outside the United States, sanctions based solely upon what happens in the United States will result in significant under-deterrence.

assumed to have the same deterrence value as an outlandish $4.4 billion to $6.3 billion fine would our overall conclusion change.[250] Only under this fantastic assumption could we fairly conclude that the current level of sanctions is sufficient. Under any reasonable assumption about the deterrence value of prison and house arrest, the current level of sanctions is far too low.

For our sample of seventy-five recent cartels that operated in the United States and internationally, their median overcharge was approximately 19% of their sales. We also found that they were sanctioned almost the exact same amount—a median sanction of approximately 17% of their sales. If they had been certain they would be caught, forming most cartels would have been a close call, because the benefits (19%) would have been only slightly larger than the costs (17%).

Unfortunately, the best evidence is that, historically, cartels in the United States have faced only a 20% to 24% chance of being discovered and convicted. Thus the "costs" of being punished are reduced to an expected 4% of sales, not 17%. This is an important reason why U.S. sanctions imposed on cartels would have had to have been on average five times higher to truly discourage most firms from colluding.

We found only one unusual cartel (out of seventy-five for which we could assemble the necessary information) for which the totality of sanctions was approximately optimal, and possibly somewhat supra-optimal.[251] A second cartel was probably optimally sanctioned.[252] The other seventy-three cartels, however, were suboptimally sanctioned, many substantially.

Concerns about over-deterrence are simply inappropriate. We believe that one reason there currently are so many cartels operating in the United States (and, indeed, the world) is that even though firms do not have all the specific data or analysis presented in this Article, prospective cartelists do have a rough appreciation that their chances of getting caught and convicted are relatively small, and that the penalties they would be likely to face if this happened would probably be modest.

250 Calculated as follows (in 2010 dollars): Total U.S. overcharges in our sample of seventy-five cases were $182 billion. To account for the allocative inefficiency effects of market power we multiplied this by 1.03 to 1.20. *See supra* Part III.B. This result ($187–218 billion) was divided by 20% to 24% (the chances of a cartel being detected and convicted). *See supra* Part IV.B. This means that our optimal sanctions goal is $779–$1090 billion.

The actual sanctions (in 2010 dollars) were $20.5 billion in settlements, plus $5.1 billion in fines, which totals approximately $26 billion. The prison and house arrest total was 2031.8 months, or 169.32 years.

The current amount of sanctions for these seventy-five cartels could be sufficient to deter collusion optimally only if the sum of $26 billion and 169.32 years in prison and under house arrest equals between $779 billion and $1090 billion in sanctions. This would occur only if each year of prison or house arrest has the sanction equivalent of $4.45–$6.28 billion.

This analysis assumes that fines and private recoveries remain unchanged.

251 *See* discussion of the E-Rate Federal Internet Program cartel *supra* Part V.B.

252 *See* discussion of the PVC Window coverings cartel *supra* Part V.B.

Coupling these low and uncertain probabilities with the relatively high prospects of significantly higher prices over a substantial period, many prospective cartel managers conclude that the risk is well worth taking. In other words, we believe that many or most prospective cartelists share the intuition behind the opinion voiced by Judge Easterbrook at the beginning of this section that crime pays. In the spirit of Judge Posner's battlefield imagery, the "cluster bombs" that constitute the current anti-cartel sanctions have been duds.

A. *Effects of Results on Cartel Sanctions and Detection*

There are two general strategies for improving the deterrence power of antitrust enforcement against cartels. One could increase the sanctions. The other possibility would be to raise the probability of detection and conviction. The proposals that follow do both.[253]

Perhaps the most straightforward policy conclusion that follows from our study would be to quintuple the overall current U.S. cartel sanction levels. A modest, ultra-conservative step in the right direction would be to double the average sanction level. This would almost certainly beneficially deter collusion and thereby save victimized consumers and businesses billions of dollars per year. Nevertheless we recognize that even a decision to double existing sanctions[254] is political in nature and is almost certain to be greeted with strong opposition. This political reality has prompted us to consider alternative policy prescriptions.[255] We instead propose nine steps that perhaps might be perceived as somewhat less controversial by those convinced that the nation's antitrust traditions are wise public policy. Only the last two would require new legislation.

First, the budget of the Antitrust Division should be increased significantly and earmarked for cartel enforcement. If the Division were able to pursue more investigations, it surely would detect and prove more cartels. As part of its use of these funds, the Division would have

253 Some of the proposals that follow, such as numbers 5, 6, and 8, fit well into the framework of conventional optimal deterrence theory. Others, such as numbers 2, 3, 4, 7, and 9, could perhaps better be termed behavioral in nature.

254 If sanctions were doubled, this study could be re-done after a few years. Perhaps, for example, even doubled levels of sanctions would cause many of the most risk-avoiding cartel members to avoid collusion or turn in existing cartels. If the results of this future optimal deterrence study showed that the overall level of cartel deterrence had not increased to an acceptable level, the sanctions could be increased still further.

255 Some of the proposals that follow fairly could be termed "behavioral," even though this paper's overall approach has been to employ the standard optimal deterrence model. See *supra* note 31 for why this is appropriate.

to commit to bring more cases where they were less than certain of victory.[256]

Second, our modest and very imperfect survey of imprisoned price fixers shows it may not be unusual for a corporation to retain and even reward employees who violate the antitrust laws.[257] We found that approximately half of those who served a prison sentence for their crime subsequently found employment for their previous employer or another employer in the same industry.[258] Too often, the corporate attitude towards price-fixing felons has been that they "took a bullet for the team" and should be rewarded. Such felons ought to be stigmatized, not awarded a badge of honor. The DOJ should re-do our study and, if the problem is in fact a significant one, as part of its settlement negotiations, should require corporations never to hire people who have ever been convicted of an antitrust violation in the same industry.[259] Similarly, convicted price fixers should agree, as part of their sentence negotiations, never to work for a firm in the same cartel again. This means that convicted price fixers will lose their jobs and be prevented from direct or indirect future employment with their employer or with other firms in the same industry, a sanction that may be very powerful indeed.[260]

Third, the Department should require convicted corporations to agree not to pay the fines incurred by their employees, directly or indirectly, or to compensate them for time spent in prison or under house arrest, directly or indirectly.[261] It is unclear how often this occurs, but it should never happen.[262]

256 For example, in 2010, the DOJ won forty-one cartel cases and lost only one. *See supra* note 25. The public interest probably would have been better served, however, if their budget had allowed them to bring one hundred cartel cases, even if they lost ten.

257 *See supra* notes 48–64. We repeat our caveat as to the extremely tentative nature of any conclusions based upon this survey, and urge others to perform a more rigorous analysis of this issue.

258 *See supra* note 48.

259 This proposal should be extended to prohibiting future service contracts with the former employer lest the convicted employee become an employee in the guise of a "consultant." For additional compliance related possibilities, see *Competition Law Compliance*, OFFICE OF FAIR TRADING, http://www.oft.gov.uk/OFTwork/competition-act-and-cartels/competition-law-compliance (last visited Sept. 26, 2012).

260 Some believe that the loss of one's job often can be even a more powerful sanction than imprisonment. *See supra* note 84. The DOJ should conduct its own survey as to what happens to convicted price fixers after they leave prison, a survey that would be much more rigorous than the preliminary one we were able to carry out and report in Part I.B.

261 Making this condition a standard clause in plea agreements is quite feasible and places the burden of monitoring on the employer. Corporations rarely, if ever, violate their plea agreements and, presumably, would be subject to penalties if they did so.

262 An analogous proposal that goes much further was made by Judge Ginsburg and Professor Wright. They believe negligent corporate officials should be debarred from working for any publicly traded corporation. *See* Ginsburg & Wright, *supra* note 25. Since their proposal would apply to the negligent corporate officials who should have prevented the antitrust violation, not just to those convicted of the offense, and it would bar them from employment at

Fourth, the Antitrust Division already has a "Wall of Shame" on its Web page—a list of every company that has paid more than $10 million in antitrust fines.[263] This should be extended to individuals for several years after their conviction. The DOJ could host, for example, a web page containing the names and photos of people given sentences of at least 6 months in prison.

Fifth, cartel fines are calculated using a formula promulgated by the U.S. Sentencing Commission.[264] The lynchpin of this formula is its estimate "that the average gain from price-fixing is 10 percent of the selling price."[265] However, in Part III.A we presented the results of two sets of data that show average cartel overcharges of 49% and 31%, and median overcharges of 25% and 22%, for the economic study and the verdict data sets, respectively.[266] A conservative, yet quite important, step the U.S. Sentencing Commission could take[267] would be to double its presumption that cartels raise prices by an average of 10%. This could increase fines substantially.

Sixth, the DOJ could change its administrative practice of awarding fine discounts from the *bottom* of the Guideline's range and start instead from the *top* of the range. We expect that this change also should result in average corporate fines that are much larger than their current levels.[268]

Seventh, the DOJ could require stricter corporate compliance programs. Some, for example, have advocated the use of corporate monitors for convicted defendants.[269] Currently, the DOJ does not require those admitted into the leniency program to have or implement compliance programs, and it certainly is possible that the widespread

any publicly traded company, not just the companies that employed them when they violated the antitrust laws, their proposal would go much further than simply preventing these punished executives from returning to their former employers. It would, however, require new legislation. A much milder—and not totally dissimilar—sanction is in effect today. Firms that fix prices can be barred from bidding on contracts with the U.S. government. We believe this does not happen very often, but it could be done more frequently.

263 *Sherman Act Violations Yielding a Corporate Fine of $10 Million or More*, U.S. DEP'T OF JUSTICE, ANTITRUST DIVISION, http://www.justice.gov/atr/public/criminal/sherman10.html (last updated July 31, 2012).

264 U.S. SENTENCING GUIDELINES MANUAL § 2R1.1(d)(1) (2005).

265 *Id.* § 2R1.1 application n.3. For an explanation how this 10% presumption results in the current fine levels, see Connor & Lande, *supra* note 14, at 522–24.

266 *See supra* Part III.A (quoting Connor & Lande, *supra* note 14, at 541). For the most recent years the figures were slightly lower—the thirty post-1990 domestic U.S. observations had a mean overcharge of 26.2% and a median overcharge of 24.5%. *Id.*

267 Technically, Sentencing Commission changes to the Guidelines are subject to Congressional approval, but historically, these resolutions have been approved unanimously.

268 Because fines are almost always a matter of negotiation, the fines might not double simply because the U.S. Sentencing Commission's formula indicates they should double.

269 *See* D. Daniel Sokol, *Behavioral Remedies for Cartels? End to Fines for Leniency Applicants and the Case for Corporate Monitors* (Jan. 15, 2012) (unpublished manuscript) (on file with the authors).

use of corporate monitors could help deter collusion.

Eighth, legislation could add prejudgment interest to both private treble damage actions and criminal fines.[270] This would increase the effective size of these sanctions substantially, especially for durable cartels or cartelists that use delaying tactics during plea bargaining or litigation. Even though any legislation that increased sanctions is likely to face strong opposition, this change has the advantage of being a change that intuitively should strike many people, including Judges Easterbrook[271] and Posner,[272] as reasonable.

Finally, the United States could implement a whistleblower-reward, or bounty system, for individuals who turn in cartels, and perhaps even for corporations.[273] Bounty proposals have the potential to enhance cartel detection and to destabilize cartels even more than the current leniency and amnesty programs. The bounties could be introduced gradually, and could be limited to individuals.[274] If this approach is not successful, some have advocated that it be introduced on the corporate level.[275] If, for example, the annual discovery rate of cartels does not decline after the other proposals in this section have been in effect for a number of years, a bounty might be awarded to corporations that turn in cartels, even if they had once been a member of the cartel. Perhaps amnesty recipients could be given 10% of all the other cartel participants' fines in egregiously harmful cases (for example, bounties could be limited to cases where affected sales exceeds $1 billion, or where the cartel members were recidivists).[276]

---

270 The U.S. Sentencing Commission could add prejudgment interest to current cartel penalties without new legislation.

271 As Judge Easterbrook noted in *Fishman v. Estate of Wirtz*, 807 F.2d 520, 583–84 (7th Cir. 1986) (Easterbrook, J., dissenting):

> [T]he time value of money works in defendants' favor. Antitrust cases can be long-lived affairs. This one has lasted 14 years, 2 1/2 of which passed between the finding of liability and the award of damages. During all of the time, the defendants held the stakes and earned interest. . . . To deny prejudgment interest is to allow the defendants to profit from their wrong, and because 14 years is a long time the profit may be substantial.

Virtually the entire profession of financial economists would agree with these principles.

272 See Judge Posner's opinion in *Patton v. Mid-Continent Systems, Inc.*, 841 F.2d 751, 752 (7th Cir. 1988) (discussing the appropriateness of contact damages: "[T]he major inadequacies being that pre- and post-judgment interest rates are frequently below market levels . . . .").

273 The UK's Office of Fair Trading and the Korean Fair Trade Commission already have these policies in place for individuals.

274 *See* William E. Kovacic, *Private Participation in the Enforcement of Public Competition Laws*, *in* 2 CURRENT COMPETITION LAW 167, 173–75 (Mads Andenas et al. eds., 2004); *see also* Cécile Aubert et al., *The Impact of Leniency and Whistleblowing Programs on Cartels*, 24 INT'L J. INDUST. ORG. 1241 (2006).

275 *See* Giancarlo Spagnolo, *Leniency and Whistleblowers in Antitrust*, *in* HANDBOOK OF ANTITRUST ECONOMICS 259 (Paolo Buccirossi ed., 2008).

276 If 10% proves to be an insufficient bounty, it could be increased to 20%, or whatever fine level proved to be optimal. It might even be optimal to give all of the fines collected from a

B. *Effects on Other Parts of the Antitrust System*

This Article's conclusions should have consequences far beyond the basic issue of whether the current levels of cartel sanctions should be raised. For example, in 1977 the U.S. Supreme Court granted standing only to direct purchasers of supracompetitively priced products, in large part because of its fear that suits by indirect purchasers would lead to "duplicative" payments.[277] The majority of states reacted by enacting "Illinois Brick Repealers" to permit injured indirect purchasers to sue for damages.[278] It often is asserted that these state laws lead to six-fold damages[279] (in addition to possible criminal penalties), and therefore, to over-deterrence. In light of this Article's conclusion that the current overall level of anti-cartel sanctions—a total that includes payments in indirect purchaser cases—should be increased at least five-fold, the Court's fear is unwarranted. On the contrary, indirect purchaser suits and state indirect purchaser laws should lead to more nearly optimal deterrence.

Moreover, as a general matter, many respected scholars believe that judicial fears that the private treble damages remedy is excessive—even before the other cartel sanctions are considered—systematically biases the results of antitrust litigation in defendants' favor.[280] Many believe

cartel to the amnesty recipient!

277 Illinois Brick Co. v. Illinois, 431 U.S. 720 (1977).

278 *See* Robert H. Lande, *New Options for State Indirect Purchaser Legislation: Protecting the Real Victims of Antitrust Violations*, 61 ALA. L. REV. 447, 448 (2010).

279 There have been a number of variations of the argument that the combination of "treble" damages for direct purchasers, plus another "three" for indirect purchasers, plus disgorgement, plus fines of two-fold damages, can lead to six-fold, eight-fold, or more overall damages paid by a cartel or monopoly. *See, e.g.*, Michael L. Denger, *A New Approach to Cartel Enforcement Remedies Is Needed, 2002 ABA Spring Antitrust Meeting* 15 (meeting held Apr. 24–26, 2002) (unpublished draft) (on file with the authors). This fear shaped the ABA's proposal in this area. *See also* Richard M. Steuer, *Report on Remedies*, 2005 A.B.A. SEC. ANTITRUST REP. 3 (One of the "key features" of their proposal is that "[t]here would be no duplicative recovery under the new cause of action . . . . the proposed statute would eliminate the possibility of duplicative recovery.").

280 As former FTC Chairman William E. Kovacic observed,

> [A] court might fear that the US statutory requirement that successful private plaintiffs receive treble damages runs a risk of over-deterrence. A court might seek to correct such perceived infirmities in the anti-trust system by recourse to means directly within its control—namely by modifying doctrine governing liability standards or by devising special doctrinal tests to evaluate the worthiness of private claims . . . . The courts will "equilibrate" the antitrust system in one of three ways. Judges will: Construct doctrinal tests under the rubric of "standing" or "injury" that make it harder for the private party to pursue its case; [a]djust evidentiary requirements that must be satisfied to prove violations; or [a]lter substantive liability rules in ways that make it more difficult for the plaintiff to establish the defendant's liability.

See Kovacic, *supra* note 274, at 173–75.

that a fear of over-deterring or unduly penalizing defendants often causes judges to favor defendants when they formulate substantive antitrust rules, when they measure ambiguous factual situations against these rules, and when they devise appropriate standing rules.[281] Similarly, in otherwise close private cases judges might unduly resolve ambiguities in defendants' favor when they compute damages because they believe the resulting award—after the mandatory trebling—will be excessive. A fortiori, a remedy system that includes not only "excessive" private damages but also incarceration and corporate fines could cause virtually every area of antitrust to develop unduly in defendants' favor. This result would be desirable only if the sanctions, when considered together, are indeed excessive. However, this Article demonstrates that for cartels, by far the most common and important type of private case, the opposite is true. Courts should resist any temptation to be lenient on lawbreakers out of a fear that they are being sanctioned too heavily.

Although we have cited critics of antitrust who are concerned about over-deterrence, at the same time, there are others who exhibit a great deal of complacency—sometimes tinged with triumphalism – that U.S. enforcement is the oldest, best developed, and most effective in the world. Pride in the antitrust idea, one of our country's most successful peaceful policy export, is understandable. But justified delight in our accomplishments can become prosecutorial hubris tantamount to obliviousness in light of the continuing high rates of cartel detections and the results of this Article's analysis. To truly protect American consumers and businesses from tremendous illegal overcharges, vigilance and increased efforts are crucial.

In short, the inquiry undertaken by this Article is not just relevant to the crucial issue of whether the overall level of cartel sanctions should be changed. Almost every piece of the extraordinarily complex and interconnected antitrust system is affected by the field's belief as to whether the current level of cartel sanctions is optimal. We believe that almost every portion of the antitrust system should be re-examined in light of this Article's analysis and conclusions.

---

281 *Id.* See also Stephen Calkins, *Equilibrating Tendencies in the Antitrust System, with Special Attention to Summary Judgment and to Motions to Dismiss*, *in* PRIVATE ANTITRUST LITIGATION 185 (Lawrence White ed., 1988), and the sources cited therein, particularly the reference to a similar analysis by Areeda and Turner, *id.* at 191. Professor Calkins discusses how many areas of antitrust law might have developed more narrowly because of the effects of damages awards that the courts believed were at the threefold level. *Id.* at 191–95. He concludes that "class actions probably would be more easily certified were there no trebling." *Id.* at 197. Professor Calkins also demonstrates why "it seems probable that trebling is a factor in" causing courts to scrutinize "damage claims more rigorously than they once did." *Id.* at 198. "Plaintiffs would find standing rules more hospitable in a single damage world." *Id.*; *see also* Stephen Calkins, *Summary Judgment, Motions to Dismiss, and Other Examples of Equilibrating Tendencies in the Antitrust System*, 74 GEO. L.J. 1065 (1986).

APPENDIX

Table 1
Summary of Economic Surveys of Cartel Overcharges

		No. Cartels	Mean %	Median %
1.	Mark A. Cohen & David T. Scheffman[282]	5–7	7.7–10.8	7.8–14.0
2.	Gregory J. Werden[283]	13	21	18
3.	Richard A. Posner 284	12	49	38
4.	Margaret Levenstein & Valerie Suslow[285]	22	43	44.5
5.	James M. Griffin[286]	38	46	44
6.	OECD (excluding peaks)[287]	12	15.75	12.75
Total (simple average)		**102–104**	**30.7**	**28.1**
Total (weighted average)		**102–104**	**36.7**	**34.6**

282 Mark A. Cohen & David T. Scheffman, *The Antitrust Sentencing Guideline: Is the Punishment Worth the Costs?*, 27 AM. CRIM. L. REV. 331 (1989).

283 Gregory J. Werden, *The Effect of Antitrust Policy on Consumer Welfare: What Crandall and Winston Overlook* 1–9 (Econ. Analysis Group, Antitrust Div., U.S. Dep't of Justice, Discussion Paper EAG 03-2, 2003), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=384100.

284 POSNER, *supra* note 65.

285 Margaret Levenstein & Valerie Suslow, *What Determines Cartel Success?* 16 (Univ. of Mich. Bus. Sch., Working Paper 02-001, 2002).

286 James M. Griffin, *Previous Cartel Experience: Any Lessons for OPEC?*, *in* ECONOMICS IN THEORY AND PRACTICE: AN ECLECTIC APPROACH 179 (L.R. Klein & J. Marquez eds., 1989).

287 ORG. OF ECON. CO-OPERATION & DEV., REPORT ON THE NATURE AND IMPACT OF HARD CORE CARTELS AND SANCTIONS AGAINST CARTELS UNDER NATIONAL COMPETITION LAWS (2002), *available at* http://www.oecd.org/dataoecd/16/20/2081831.pdf.

Table 2
Median Average Episodic Overcharges, by Year and Type

Cartel Episode End Date	Membership		Legal Status		Bid-Rigging	Classic Price Fixing	All Types
	Nat'l	**Int'l**	Found Guilty	**Legal**			
	Median percent [a]						
1780–1890	19.3	**50.8**	16.0	**25.0**	16.2	21.3	20.3
1891–1919	24.5	**57.3**	24.8	**41.5**	39.0	35.0	36.8
1920–1945	4.6	**31.6**	38.9	**27.6**	34.0	30.0	30.0
1946–1973	15.0	**38.9**	14.3	**20.4**	13.3	19.0	15.2
1974–1989	16.8	**37.4**	23.0	**7.5**	21.8	16.9	20.0
1990–1999	14.9	**24.8**	22.8	**11.7**	16.0	23.0	22.2
2000–2009	20.0	**25.8**	23.3	**17.5**	18.5	24.1	22.5
ALL YEARS	**17.2**	**30.0**	**22.8**	**26.0**	**18.6**	**25.0**	**23.3**

Sources: Appendix Tables 1 and 2, summarized in J. Connor, *Price Fixing Overcharges Master Data Set*, spreadsheet dated July 2009.

[a] Medians of the point estimates or, where appropriate, of the midpoint of range estimates. Includes many zero estimates. See Table 4 for the numbers of observations in each cell.

Table 3
Studies and Opinions as to the Probability of Cartel Detection

Source	Probability	Comment
Alan R. Beckstein & Gabel H. Landis[288]	Less than 0.50	A large anonymous survey of antitrust lawyers in the ABA, most working in the United States; the mean response was 3.6, where 5=strongly agree, 4=agree, and 3=neither agree nor disagree.
William M. Landes[289]	0.33	Merely an illustration, but a seminal work on optimal deterrence that may influence many adherents of optimal deterrence theory.

288 Alan R. Beckstein & Gabel H. Landis, *Antitrust Compliance: Results of a Survey of Legal Opinion*, 52 ANTITRUST L.J. 459, 487–516 (1982).
289 Landes, *supra* note 15, at 657.

R.M. Feinberg[290]	Less than 0.50	An anonymous confidential survey of antitrust lawyers working in Brussels and observing the EC; the mean response was 4.4, where 5=strongly agree and 3=neither agree nor disagree.
United States Sentencing Commission[291]	0.10	Contains the transcript of 1987 testimony of DAAG for Antitrust, Ginsburg; probably refers to domestic cartels of 1970s and 1980s.
Gregory J. Werden & Marilyn J. Simon[292]	Less than 0.10	Appears to be a general, subjective opinion of Antitrust Division professional prosecutors.
Mark A. Cohen & David T. Scheffman[293]	0.33	No hint as to the source, but may have been influenced by Landes (1983).
Jean-Claude Bosch & E. Woodrow Eckard Jr.[294]	0.13–0.17	A quantitative estimate derived from an event study of U.S.-prosecuted cartels 1961–1988.
Mitchell A. Polinsky & Steven Shavell[295]	0.138–0.165	Refers to U.S. arrest rates for some of the most common felonious property crimes (burglary, auto theft, and arson); may be overstated if victims of such crimes fail to report some occurrences.

290 Feinberg, *supra* note 184, at 379.

291 *Sentencing Options: Hearing Before the U.S. Sentencing Comm'n* 15 (July 15, 1986), *available at* http://www.src-project.org/wp-content/pdfs/testimony/ussc_testimony_prepared_19860715/0008752.pdf (statement of Douglas H. Ginsburg, Assistant Attorney Gen., Antitrust Division, U.S. Dep't of Justice).

292 Gregory J. Werden & Marilyn J. Simon, *Why Price Fixers Should Go to Prison*, 32 ANTITRUST BULL. 917, 926 (1987).

293 Cohen & Scheffman, *supra* note 282.

294 Jean-Claude Bosch & Woodrow E. Eckard Jr., *The Probability of Price Fixing: Evidence from Stock Market Reaction to Federal Indictments*, 73 REV. ECON. & STAT. 309 (1991).

295 Mitchell A. Polinsky & Steven Shavell, *The Economic Theory of Public Enforcement of the Law*, 38 J. ECON. LITERATURE 45, 70 (2000).

Office of Fair Trading[296]	0.30	An anonymous survey of U.S. antitrust lawyers in private practice (with a "low response rate") asked about the increase in cartel activity "if the Division stopped enforcing Section 1 of the Sherman Act." Results were originally summarized in the FY2001 DOJ report to Congress.
Richard A. Posner[297]	0.25	An illustration of an optimal deterrence calculation by a leading antitrust jurist.
Organisation of Economic Co-Operation and Development[298]	0.13–0.17	OECD accepts Bosch and Eckard (1991).
Emmanuel Combe et al.[299]	0.129–0.133	Replicate Bosch and Eckard's (1991) method using data from EU-prosecuted cartels from 1969 to 2002.
Bush et al.[300]	0.10–0.33	A summary of most of the sources in this table above.
Alla Golub et al.[301]	0.13–0.17	This paper replicates the Bosch and Eckard (1991) model using U.S. cartels from a later period and finds few differences in deterrence.
Terry Calvani[302]	0.13–0.17	In an Article on cartel enforcement an experienced antitrust official cites Bosch and Eckard (1991) with approval.

296 DELOITTE, THE DETERRENT EFFECT OF COMPETITION ENFORCEMENT BY THE OFT 20 (2007), *available at* http://www.oft.gov.uk/shared_oft/reports/Evaluating-OFTs-work/oft962.pdf (prepared for OFT).

297 POSNER, *supra* note 65, at 47.

298 ORG. OF ECON. CO-OPERATION & DEV., *supra* note 287, at 18–19.

299 Combe et al., *supra* note 170.

300 Brief for Bush et. al. as Amici Curiae Supporting Respondents, F. Hoffman-LaRoche v. Empagran, 542 U.S. 155 (2004) (No. 03-724).

301 Golub et al., *supra* note 169.

302 Terry Calvani, *Enforcement of Cartel Law in Ireland*, *in* 6 CAMBRIDGE YEARBOOK OF EUROPEAN LEGAL STUDIES ch. 4, at 77 (John Bell & Claire Kilpatrick eds., 2005).

Wouter P.J. Wils[303]	Less than 0.33	Cites with approval Bosch and Eckard (1991), but author believes that the U.S. probability has increased since 1961–1988 and that it is lower in the EU than the United States; this is a "conservative" upper limit for the EU.
Maarten Pieter Schinkel[304]	0.15	Cites only Bosch and Eckard (1991), but considers it "controversial as well as dated."
Maurice E. Stucke[305]	Unknown, but possibly 0.13–0.17	"Nobody knows." However, the author also favorably cites USSG (1986), OECD (2002), and Bosch and Eckard (1991).
Paolo Buccirossi & Giancarlo Spagnolo[306]	0.15	The author's "prudent" assumption for their simulation analysis.
J. Chen & J.E. Harrington[307]	0.1–0.3	In illustrating the effect of detection probability of cartel formation, the authors chose this range.
Office of Fair Trading[308]	21.7% caught of those seeking advice	Results of a survey of 234 competition-law lawyers in the UK and Brussels for the years 2004–06 asking what proportion of their clients were convicted of illegal cartel conduct (295) by the UK's OFT compared to the 1361 instances where a client abandoned or changed a possible cartel agreement "because of the risk of OFT investigation."

303 Wouter P.J. Wils, *Is Criminalization of EU Competition Law the Answer?*, 28 WORLD COMPETITION 117, 130 (2005).

304 Maarten Pieter Schinkel, *Effective Cartel Enforcement in Europe* 25 (Amsterdam Ctr. of Law & Econ. Working Paper No. 2006-14, 2006), *published in* 30 WORLD COMPETITION: LAW & ECON. REV. 539 (2007), *available at* http://www.ssrn.com/paper=948641.

305 Maurice E. Stucke, *Morality and Antitrust*, 2006 COLUM. BUS. L. REV. 443, 457.

306 Buccirossi & Spagnolo, *supra* note 15, at 95.

307 Joe Chen & Joseph E. Harrington, *The Impact of Corporate Leniency on Cartel Formation and the Cartel Price Path*, *in* THE POLITICAL ECONOMY OF ANTITRUST, *supra* note 15, ch. 3, at 76.

308 DELOITTE, *supra* note 296, at 50–54.

Nathan H. Miller[309]	0.21–27.5	An empirical study of U.S. cartel prosecutions shows that detection rates rose 62% because of the revised 1993 Leniency Program; this increase is applied to Bosch and Eckard's estimate of p.
Renato Nazzini & Ali Nikpay[310]	Less than 0.20	"The authors' own anecdotal observations suggest that the OFT fully investigates less than 20 percent of all cases in which it has a reasonable suspicion that the competition rules have been breached."
Gregory J. Werden[311]	0.25	Part of an illustration of optimal fines for typical EU cartels.
Peter Ormosi[312]	10–20%	Calculations for Europe based on a large number of factors.

309 Miller, *supra* note 24.
310 Renato Nazzini & Ali Nikpay, *Private Actions in EC Competition Law*, 4 COMPETITION INT'L POL'Y 111, 111 (2008).
311 Gregory J. Werden, *supra* note 20, at 27–29 (2009).
312 Peter L. Ormosi, *How Big Is a Tip of the Iceberg? A Parsimonious Way to Estimate Cartel Detection Rate* (Ctr. for Competition Policy, Working Paper No. 11-6, 2011), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1851309.

Exhibit 4 to the Cross-Examination of Professor Tadelis

CARDOZO LAW

Cardozo Law Review

Cardozo Law Review was established in 1979 and consistently ranks among the top law journals in the country.

Key links

Cardozo Law Review

Email | cardozo.l.rev@gmail.com

Cardozo Law Review publishes six issues per year, which contain articles and student notes on a variety of legal topics. The *Law Review* is dedicated to advancing empirical, interdisciplinary, and philosophical scholarship, as well as exploring cutting-edge national issues.

Additionally, the *Law Review* sponsors several **symposia and lectures** each year that delve into unique and intellectually stimulating areas of legal scholarship. Recent symposia and lecture topics have included election law, immigration, comparative constitutionalism, and law and economics.

Finally, ***de•novo***, the *Cardozo Law Review*'s online companion, is the home for shorter articles on timely legal issues, as well as responses to articles that appear in the print journal. *De•novo* also hosts "mini-symposia" that feature articles from professors and practitioners across political and sociological lines in order to create a lively debate on a selected legal issue. In addition to articles written by professors and practitioners, *de•novo* also publishes several student notes each year.

CARDOZO
Benjamin N. Cardozo School of Law

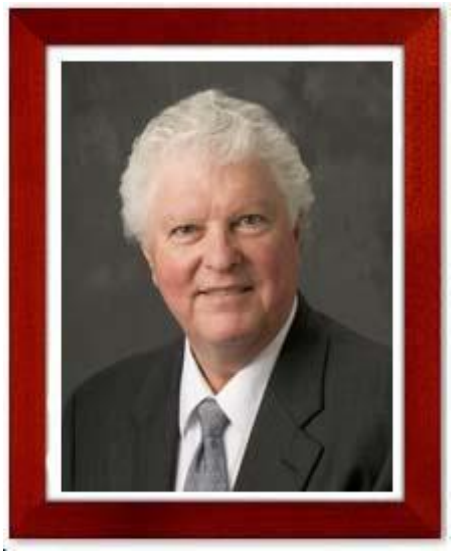
Jacob Burns Institute for Advanced Legal Studies

55 5th Avenue
New York, NY 10003
lawinfo@yu.edu
212.790.0200

Exhibit 5 to the Cross-Examination of Professor Tadelis

Former Agricultural Economics Faculty

JOHN M. CONNOR (1943 -)



Since 1983 a professor of Industrial Economics at Purdue University's Department of Agricultural Economics, he holds a bachelor's degree in mathematics from Boston College, an MA from the University of Florida, and a Ph.D. from the University of Wisconsin. From 1965 to 1968 he served as a volunteer in the U. S. Peace Corps teaching mathematics in two African nations. From 1976 to 1979 he was an Adjunct Professor at the University of Wisconsin, and from 1979 to 1983 he was Head of Food Manufacturing Research in the Economic Research Service of USDA.

Dr. Connor specializes in research and teaching of empirical industrial economics and antitrust policy. For 25 years he studied the market structure and performance of the food manufacturing and distribution industries; for the past 15 years the focus of his research has been international price-fixing cartels and world-wide antitrust enforcement.

Connor is the author of 19 books and monographs and more than 600 other scholarly publications in economics and law. His publications are widely read. In March 2011, the Social Science Research Network reported on the number of downloads of working papers, preprints and offprints listed by 358,000 authors listed in SSRN. Connor had 11,241 downloads of his 58 papers, which ranked him 86th out of more than 11,000 economists in the SSRN repository. In addition, his research has been cited in more than 75 law-review articles and five court decisions.

His latest book, *Global Price Fixing*, received two national writing awards in 2002 and 2003. New editions of *Global Price Fixing* appeared in 2007 and 2008. A review of the book in the *Journal of Economic Literature* calls it:

"….an invaluable reference for those with a deep interest in the economics of cartels, and a very readable and accessible narrative…on international cartels."

Dr. Connor is Senior Advisor to the American Antitrust Institute, and he consults for government antitrust authorities and law firms in cartel cases. In 2009, his lifetime achievements were recognized when he was awarded the highest honor of the Agricultural and Applied Economics Association, the title of Fellow.

Exhibit 6 to the Cross-Examination of Professor Tadelis

(https://www.antitrustinstitute.org/)

Mission and History

Mission

Devoted to fair competition that protects consumers, businesses, and our society

SUPPORT AAI (HTTPS://WWW.ANTITRUSTINSTITUTE.ORG/SUPPORT/)

The American Antitrust Institute (AAI) is an independent, nonprofit organization devoted to promoting competition that protects consumers, businesses, and society. We serve the public through research, education, and advocacy on the benefits of competition and the use of antitrust enforcement as a vital component of national and international competition policy.

The American Antitrust Institute is a 501(c)(3) not-for-profit organization (Tax ID #52-2093834). All donations are tax deductible. Learn how you can support AAI.
(https://www.antitrustinstitute.org/about-us/supporting-aais-research-education-and-advocacy-mission/),

History

View less −

AAI created and shaped the modern progressive antitrust movement

AAI was founded in 1998. It was the first public interest research, education, and advocacy organization to provide a counterpoint to conservative influence in antitrust enforcement and competition policy.

The organization's original directors were AAI's founder Albert Foer (https://www.antitrustinstitute.org/people/albert-a-foer/), Robert Lande (https://www.antitrustinstitute.org/people/robert-h-lande/), and Jonathan Cuneo (https://www.antitrustinstitute.org/people/jonathan-cuneo/). Foer served as President until his retirement in 2014, and was succeeded by then Vice President Diana Moss (https://www.antitrustinstitute.org/people/diana-l-moss/). Moss served until 2023 and was succeeded by Vice President Randy Stutz (https://www.antitrustinstitute.org/people/randy-stutz/). The growth and progress of AAI has been made possible by the dedicated and long-term work of leadership and staff with expertise in law, economics, policy, and institutions.

AAI created and shaped the modern progressive antitrust movement, with a focus on preserving a competitive economy through vigorous public and private enforcement of the antitrust laws. Protecting markets, competition, innovation, consumers, and workers is fundamental to ensuring economic and entrepreneurial freedom and the underlying democratic values embodied in our market system.

U.S. antitrust enforcement in the 1960s and 1970s embraced the importance of protecting vulnerable market participants and entrants through vigorous enforcement. By the 1980s, however, antitrust had almost reversed course. Advocacy by the conservative law and economics movement steered antitrust policy in a non-interventionist direction marked by lax merger control and forbearance from policing monopolistic and other anticompetitive practices.

The conservative movement has been, and continues to be, rooted in excessive deference to unsubstantiated efficiencies and pro-business justifications for mergers and abusive conduct. It has arguably resulted in increases in concentration at economy-wide, industry, and market levels and the growth of dominant firms with significant seller and buyer market power. Today, a burgeoning body of economic and policy work has linked these trends to a general decline in competition, likely due in no small part to the lax enforcement policies that have held sway for the past five decades.

AAI seeks to return balance to antitrust enforcement and competition policy (https://www.antitrustinstitute.org/issues/competition-policy/) through new ways of thinking that keep antitrust relevant in the face of increased and more sophisticated strategic behavior, technological change, and globalization. AAI applauds pro-enforcement efforts of the Antitrust Division of the U.S. Department of Justice (https://www.justice.gov/atr), the Federal Trade Commission (https://www.ftc.gov/), the State attorneys' general, and "private attorneys' general." Still, resources devoted to public antitrust enforcement are well below their level of 30 years ago and attacks on public and private antitrust enforcement continue to mount. AAI meets these challenges by educating the public and media on the importance of reinvigorated antitrust enforcement and working collaboratively with enforcement agencies, sector regulators, Congress, academia, industry, and other advocacy groups.

AAI continues to be part of the bipartisan tradition that has supported the antitrust enterprise for over 125 years. AAI works across a network of the legal, economic, business, policy, and academic communities to develop positions that reflect clear-headed, informed, multidisciplinary, and progressive thinking about antitrust and competition policy. It benefits from the counsel of a diverse Advisory Board (https://www.antitrustinstitute.org/people/role/advisory-board/) composed of highly regarded experts and a network of friends and supporters.

For a more detailed history of AAI, see How AAI Became an Institution (https://www.antitrustinstitute.org/bert-foer-tribute-issue-of-the-antitrust-bulletin/).

Diversity & Inclusiveness

View less −

The American Antitrust Institute (AAI) is committed to **diversity, equity, and inclusion (DEI)** in all aspects of its work. We believe that DEI is essential to fulfilling our mission of promoting competition that protects consumers, businesses, and society. Within our organization, a diverse workforce and leadership team are better able to understand and represent the needs of all stakeholders, and to develop and advocate for policies that benefit everyone.

AAI is also committed to promoting DEI in the antitrust field as a whole — working to raise awareness of the importance of DEI in antitrust enforcement and competition policy, while advocating for policies that promote DEI in the antitrust field.

See our efforts in DE&I → (/issues/diversity-and-inclusion/)

MAKE A DONATION (/DONATE/)

AMERICAN ANTITRUST INSTITUTE

About AAI (http://www.antitrustinstitute.org/about-us/)

Mission and History (http://www.antitrustinstitute.org/about-us/mission-and-history/)

Our People (http://www.antitrustinstitute.org/about-us/our-people/)

Awards (/about-us/awards/)

Careers & Internships (http://www.antitrustinstitute.org/about-us/internships-and-research-fellows/)

Our Work (/our-work/)

News & Events (/news-events/)

Support AAI (http://www.antitrustinstitute.org/about-us/mission-and-history/supporting-aais-research-education-and-advocacy-mission/)

Contact Us (http://www.antitrustinstitute.org/contact-us/)

Join Our Mailing List (/join-aai-email-list/)

Terms of use (http://www.antitrustinstitute.org/about-us/terms-of-use/)

Exhibit 7 to the Cross-Examination of Professor Tadelis

PUBLIC

Robert H. Lande Biography

Robert H. Lande

Robert H. Lande is the Venable Professor of Law at the University of Baltimore School of Law, and also is a co-founder and Director of the American Antitrust Institute. He formerly worked at Jones, Day, Reavis and Pogue and at the Federal Trade Commission. He has authored or co-authored more than 60 U.S. and nine foreign publications, including articles in the California Law Review, Journal of Law and Economics, Duke Law Journal, and the Georgetown Law Journal. Seven of his articles have been republished in books or collections of articles.

Professor Lande has been quoted in the media hundreds of times about antitrust issues, and has discussed antitrust on TV in the United States, France, the United Kingdom and China. He has spoken at national events sponsored by the American Bar Association, Association of American Law Schools, National Association of Attorneys General, American Antitrust Institute, American Economic Association and INCECOPI (Peru). He has testified before the US House of Representatives Judiciary Committee, the US Senate Commerce Committee, and the Antitrust Modernization Commission. He has given competition advice to enforcement officials from four foreign nations. He is a past chair of the Association of American Law Schools Antitrust Section and has held many positions in the ABA Antitrust Section. He is an elected member of the American Law Institute and a member of the the Washingon, D.C. bar. He can be reached at rlande@ubalt.edu.

Updated January 3, 2024

Exhibit 8 to the Cross-Examination of Professor Tadelis

COMPASS LEXECON

William M. Landes

Co-Founder

PhD in Economics, Columbia University

Contact

Office: +1 312 322 0200

Fax: +1 312 322 0218

Email

Services

William Landes, Co-Founder of Lexecon, is the Clifton R. Musser Professor of Law and Economics at The Law School at the University of Chicago. Prior to joining the University of Chicago Law School, he taught in the Economics Departments of Stanford University, the University of Chicago, Columbia University, and the Graduate Center of the City University of New York and fellow of the American Academy of Arts and Sciences. Landes is an expert in the application of economics and econometrics to legal problems including intellectual property, torts and antitrust and has written widely in these fields. He has appeared as an expert before courts and various government agencies. Landes formerly served as the President of the American Law & Economics Association and as an editor of the *Journal of Law and Economics* and the *Journal of Legal Studies*. In addition, he co-authored *The Economic Structure of Tort Law* and *The Economic Structure of Intellectual*

COMPASS LEXECON

University.

Education

- PhD in Economics, Columbia University

William's selected cases

COMPASS LEXECON

Rate Adjustment Proceeding

Learn how Compass Lexecon"s experts influenced the CRB to set fair royalty rates for mechanical licenses, including ringtones, benefiting Copyright Owners.

Entertainment & Media | Intellectual Property

Join our mailing list to receive the latest news and industry insights

Your email address Submit

Resources
- Professionals
- Services
- Industries
- Cases
- Insights

Company
- About us
- Careers
- Locations
- Contact us

Legal
- Privacy Policy
- Terms & Conditions
- Cookie Policy

COMPASS LEXECON

COMPASS LEXECON

Continents

Nobel Prize winners

Offices

Our story

Established in 1977, Compass Lexecon has grown to become a globally pre-eminent economic consulting firm.

Read our story

COMPASS LEXECON

1977
Established

700+
Employees

170+
PhDs

4
Continents

2
Nobel Prize winners

25
Offices

Our story

Established in 1977, Compass Lexecon has grown to become a globally pre-eminent economic consulting firm.

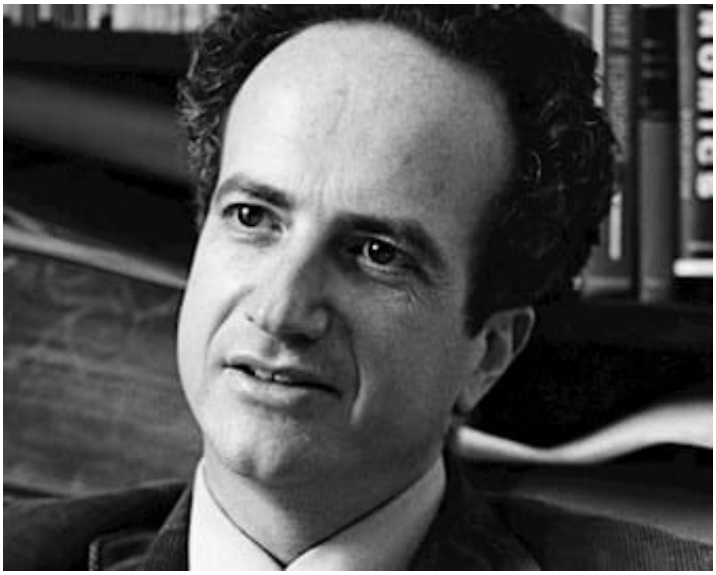
Read our story

Exhibit 9 to the Cross-Examination of Professor Tadelis

Nobel Laureate

Gary Becker

Becker won the 1992 Nobel Prize in Economic Sciences for having extended the domain of microeconomic analysis to a wide range of human behavior and interaction, including nonmarket behavior.



A Pioneer of the Economics of Human Behavior

One of the foremost economics scholars of the 20th century, the late Gary Becker transformed the field by extending its impact to the study of human behavior. By applying economic frameworks to real-world issues such as crime, discrimination, and the economics of the family, he upended the definition of economics.

Born in the small coal-mining town of Pottsville, Pennsylvania, Becker spent his formative years in Brooklyn, New York, where his family's frequent discussions of politics and justice sparked a lifelong interest in social issues.

At Princeton University, he found his future calling when he "accidentally took a course in economics." Intrigued by the "mathematical rigor of a subject that dealt with social organization," as he later wrote, he graduated from Princeton in just three years and went on to pursue a graduate degree in economics at the University of Chicago.

At the University of Chicago, he studied with the late Milton Friedman, whose influence on Becker was profound. "He emphasized that economic theory was not a game played by clever academicians, but was a powerful tool to analyze the real world," Becker later wrote. "His course was filled with insights both into the structure of economic theory and its application to practical and significant questions."

Inspired by those ideas, Becker focused his research on improving the world. His first major publication, *The Economics of Discrimination*, examined the economic implications of discrimination. He went on to apply economic tools to the study of labor markets, birth rates, addiction, education, crime, and other social issues.

In the early part of his career, Becker's work on social problems was controversial. "For a long time, my type of work was either ignored or strongly disliked by most of the leading economists," Becker wrote. "I was considered way out and perhaps not really an economist."

But over time his ideas gained traction, eventually becoming mainstream.

"Before Becker, economics was about topics like business cycles, inflation, trade, monopoly and investment," economist Lawrence H. Summers wrote in *Time* magazine shortly after Becker's death. "Today it is also about racial discrimination, schooling, fertility, marriage and divorce, addiction, charity, political influence—the stuff of human life."

Becker taught at the University of Chicago from 1954 until 1957, when he joined the faculty at Columbia University and the National Bureau of Economic Research. He returned to the University of Chicago in 1970 and spent the rest of his career here, joining the Chicago Booth faculty in 2002.

In addition to his scholarly work, Becker wrote monthly columns for *Businessweek* and ran a popular blog in collaboration with Richard Posner, a judge on the federal court of appeals and a senior lecturer at the University of Chicago Law School.

"It [economics] is judged ultimately by how well it helps us understand the world, and how well we can help improve it."

—Gary Becker

Nobel Prize–Winning Impact

Gary Becker's Nobel Prize–winning work broke new ground by crossing disciplinary boundaries and applying economic principles to answer pressing questions about human behavior.

Published in 1957, *The Economics of Discrimination* examined racial discrimination from the perspective of markets. This landmark work ultimately upended the way economists and other social scientists think about wage inequities.

In his work on crime and punishment, Becker applied free-market principles to the question of why people commit crimes, arguing that the way to reduce crime was to make the costs outweigh the benefits.

His 1964 book *Human Capital* advanced groundbreaking theories about the payoff of investments in education and worker training.

In his 1981 book *A Treatise on the Family*, he argued that bringing more women into the workforce increases the value of their time and thus reduces their desire to have children—work that has had an impact on development policies worldwide.

Milestones

1951

Graduated from Princeton University with a bachelor's degree in mathematics

1954

Joined the economics faculty at the University of Chicago

1955

Earned a PhD from the University of Chicago

1957

Published *The Economics of Discrimination*

1970

Rejoined the University of Chicago faculty after 12 years at Columbia University

2000

Received the National Medal of Science

2002

Joined the Chicago Booth faculty

2007

Was awarded the Presidential Medal of Freedom

Featured Insights

How Gary Becker Saw the Scourge of Discrimination

Becker's analysis would extend the reach of economics, and completely reshape the field—and social-science research in general, but it took decades to do so.

CBR - Economics

Is It Time to End the Cuban Embargo?

Gary S. Becker and Richard Posner call for an end to the policy that dates to 1960.

CBR - Economics

From 2009: Approaching Global Warming

How to analyze climate change policies.

CBR - Economics

CHICAGO BOOTH
THE UNIVERSITY OF CHICAGO BOOTH SCHOOL OF BUSINESS

Explore

Why Booth
Degree Programs
Executive Education
Faculty & Insights
Alumni
Events
Recruiting
Admissions
About
Careers at Booth
Media Relations
Chicago Booth Magazine
Chicago Booth Review

REQUEST INFORMATION
APPLY
GIVE

Other Offerings

UNDERGRADUATES
TAKING CLASSES
ALL OTHER OFFERINGS

Global Locations

Harper Center

Gleacher Center

Executive MBA Chicago

Executive MBA London

Executive MBA Hong Kong

Executive Education Global Locations

Follow

The University of Chicago Booth School of Business

5807 S. Woodlawn Ave.
Chicago, IL 60637 USA

Phone: 773.702.7743

CONTACT

Exhibit 10 to the Cross-Examination of Professor Tadelis

2015

Antitrust Enforcement Regimes: Fundamental Differences

Keith N. Hylton
Boston University School of Law

ANTITRUST ENFORCEMENT REGIMES: FUNDAMENTAL DIFFERENCES

Boston University School of Law Working Paper No. 12-41
(August 23, 2012)

Keith N. Hylton
Boston University School of Law

Antitrust Enforcement Regimes: Fundamental Differences

Keith N. Hylton*

August 2012

Abstract: Since China has modeled its antitrust regime on that of the EU, there are essentially two antitrust regime types: the U.S. and the EU. This chapter is a brief comparative study of the two regimes. I focus on three categories in which fundamental differences are observed: enforcement, legal standards, and procedure. Within each of the three categories, I narrow the focus to a specific illustrative feature. With respect to enforcement, the EU imposes gain-based penalties while the U.S. imposes harm-based penalties. In predation law, the U.S. has a marginal cost standard and the EU has an average cost standard. With respect to procedure, the U.S. is a common law system, while the EU's procedure is closer to the civil law system in its allocation of power between the courts and the enforcement agency. These differences have profound implications for the welfare consequences of global antitrust enforcement.

* Boston University, knhylton@bu.edu. I thank Danny Sokol for helpful comments. Adam Mayle provided excellent research assistance. This is a forthcoming chapter in Roger D. Blair and D. Daniel Sokol, Oxford Handbook of International Antitrust Economics (Oxford University Press).

There are more than 100 antitrust enforcement regimes around the world (see, e.g., Hylton n.d.). Because of this, it is difficult to say much about antitrust statutes globally without running into the difficulty that such large numbers immediately imply for comparative projects. Comparisons are possible on a global scale, but they are necessarily limited to statistical summaries of the major features of the enforcement regimes.[1]

In spite of the large number of antitrust enforcement regimes, there are three that are recognized as extremely important in global commerce: the U.S., the EU, and China.[2] Moreover, China has modeled its antitrust law regime on that of the EU (Farmer 2010, 35-36). Given this, there are essentially two antitrust regime types that dominate global commerce: the U.S. and the EU.

This chapter is a brief comparative study of the two major antitrust regimes. Even limiting my study to two competition regime types, there are many details in which the two types diverge. I will not compare the regimes in terms of all of the details. Instead, I will focus on three major areas in which fundamental differences are observed: *enforcement*, *legal standards*, and *procedure*.

Within each of the three categories, I narrow the focus to a specific illustrative feature. In the enforcement category, I discuss penalty provisions. In the legal standards category, I examine predatory pricing law as a central feature illustrating fundamental differences between the two regimes. Finally, I summarize broad differences in procedure under the regimes.

With respect to enforcement, the EU and U.S. regimes differ in that the EU imposes fines that are based on the violator's gain while the U.S. imposes harm-based penalties. In predation law, the U.S. has adopted a marginal cost standard and the EU has adopted an average cost standard. With respect to procedure, the U.S. is a thoroughly common law system, while the EU's procedure is closer to the civil law system in its allocation of power between the courts and the enforcement agency. These differences have profound implications for the welfare consequences of global antitrust law enforcement.

Enforcement

The economic theory of enforcement prescribes punishment schemes that maximize society's welfare, by reducing the sum of the costs of offensive conduct and the costs of enforcement. In this part, I briefly review the theory, and use it to address core differences in the antitrust enforcement policies of the U.S. and the EU.

---

[1] For an empirical study of competition law enforcement regimes around the world, see Hylton and Deng 2007. For comparative analysis, Hylton and Deng group countries into regions (Europe, North America, etc.) to compare the general stance of antitrust enforcement across regions.

[2] The primacy of the U.S., EU, and China antitrust regimes is evidenced by the fact that news stories discussing major mergers often focus on the approval processes in these three regimes. The coverage of the Google and Motorola Mobility merger exemplifies this trend (Bartz and Chee 2012; Whitney 2012).

Optimal Enforcement Policy

The theory of optimal antitrust enforcement is traceable to Gary Becker's article on the economics of punishment (Becker 1968). Becker argued that an efficient system of punishment would seek to internalize the social costs associated with offensive conduct. Internalization is accomplished by shifting the costs suffered by victims to the offender in the form of a penalty. Becker considered the implications of his argument for antitrust, arguing that the social costs arising from antitrust violations should be internalized by those engaging in anticompetitive conduct (Becker 1968, 198-199). Later, Landes (1983) provided a more detailed application of Becker's analysis to antitrust. In the antitrust context, internalization requires the punishment authority to shift the costs suffered by consumers, in terms of monopolistic overcharges or restrictions in supply, to the monopolizing firm in the form of a penalty.

Consider the case of a firm that takes some action that enables it to gain monopoly pricing power, and at the same time generates efficiencies in production or sale. For example, the action could be a merger that creates or enhances monopoly power and at the same generates efficiencies – say, by cutting redundant worksites. Alternatively, the action could be an exclusive dealing arrangement that forecloses a rival firm and at the same time reduces costs in the supply chain.

The economic effects of the firm's conduct can be examined in Figure 1. Before the firm takes the action, the market is competitive, with price equal to marginal cost ($p_0 = c_0$). After the firm takes the action, it gains the power to raise its price to the monopoly level p_1. However, costs fall, as a result of the firm's action, from c_0 to c_1.

The optimal enforcement policy in this case is to impose a penalty on the firm equal to the sum of the wealth transfer from consumers and the forgone consumer surplus, which is represented by the sum of areas *T* and *D* in Figure 1 (Landes 1983; Hylton and Lin 2010).

Why would setting the penalty equal to the sum of the consumer wealth transfer and the forgone consumer surplus ($T+D$) be optimal? The reason is that it aligns the firm's incentives with society's incentives. If the penalty is equal to the sum of the consumer wealth transfer and the forgone consumer surplus, then the firm will choose to take the monopolizing action when and only when it enhances society's wealth.

Suppose the monopolizing action generates an efficiency gain, shown by the area *E* in Figure 1. The firm's action is welfare enhancing for society as long as the efficiency gain is greater than the forgone consumer surplus, that is, $E > D$. The firm will choose the monopolizing action if the gain the firm gets from monopolizing is greater than the expected penalty for monopolization. The firm's gain is the sum of the transfer from consumers and the efficiency gain, $T+E$. Thus, if the penalty is equal to the sum of the

transfer and the forgone consumer surplus ($T+D$), the firm will take the monopolizing action only when it is welfare enhancing ($E > D$).[3]

Although I have used the terms "monopolization" and "monopolizing firm", this analysis applies equally to cartels. If a cartel has an efficiency basis, then it will lead to an increase in price, generating a wealth transfer from consumers, and a reduction in supply costs, generating an efficiency gain. For simplicity, I will use the term "monopolizing firm" for the remainder of this paper, even in instances in which the monopolizing act is the decision by a group of firms to create a cartel.

This analysis implies that the optimal antitrust enforcement policy internalizes, to the monopolizing firm, the harm suffered by consumers. To be sure, this is the optimal policy because there is an efficiency gain resulting from the firm's monopolizing action. If there were no efficiency gain ($E = 0$), then the optimal policy would set the penalty in order to completely deter the violator's conduct (Becker 1968, 180; Hylton 1996, 197-198). Such a complete-deterrence penalty would have to be at least as large as the wealth transfer from consumers (Becker 1968, 198-199; Landes 1983, 656). In other words, when the firm's monopolizing action does not generate an efficiency gain, the optimal punishment policy is complete deterrence, which is accomplished by ensuring that the firm cannot profit from monopolization.[4] Any penalty greater than the wealth transfer (T) satisfies the complete deterrence – or, equivalently, gain elimination – objective.

Thus, there are two general approaches a punishment authority can take under an optimal punishment regime. One is to *internalize consumer harm*. The other is to *deter completely by eliminating the expected profits* from anticompetitive conduct. The internalization approach is appropriate for conduct that is either efficient or has a significant chance of being efficient. The complete deterrence approach is appropriate for conduct that is unambiguously inefficient.

I have already provided examples of monopolizing conduct that may be efficient: mergers and exclusive dealing. As for conduct that is unambiguously inefficient, the obvious example is the price-fixing agreement (Posner 2001, 39). The standard price-fixing agreement involves no efficiency motivation; it is simply an arrangement to transfer wealth from consumers to producers (see, e.g., Leslie 1993).[5] Under these conditions, the optimal punishment policy is to set a fine sufficient to eliminate the prospect of gain from the price-fixing cartel's actions.

In the case of a price-fixing cartel, the internalization policy still satisfies the optimal punishment goal, because the consumer harm is the same as the producers' gain from

---

[3] To see this, note that the firm will take the monopolizing action when $T+E >$ penalty. If the penalty is equal to $T+D$, then the firm will monopolize when $T+E > T+D$, or when $E > D$.

[4] On the distinction between internalization and complete deterrence policies, see Hylton 1998, 425-433

[5] It is possible for a price-fixing agreement to be efficient. If the agreement is efficient, and the efficiency gains are enjoyed by the cartel, the cartel may have an incentive to continue the agreement even when faced with a damages remedy. *See* Becker 1968, 199 ("If . . . certain constraints of trade raise the level of economic welfare, fines could fully compensate society for the harm done, and yet some constraints would not cease, because the gain to participants would exceed the harm to others.").

price-fixing. For this reason, Becker concluded that the optimal punishment policy for antitrust is one that internalizes the consumer harm (Becker 1968, 199).

While it is true that the two policies suggested by enforcement theory, harm internalization and complete deterrence, can be satisfied in the antitrust setting by a penalty that internalizes consumer harm, the different policy goals of the internalization and deterrence approaches should be kept in view. The reason is that there may be instances in which the internalization approach is administratively infeasible. For example, the internalization approach requires the punishment authority to produce a precise estimate of the consumer wealth transfer and the forgone consumer surplus. The data necessary to generate such an estimate may be unavailable. In such a case, where the risk of error is substantial, it is important to identify the precise policy basis for punishment.

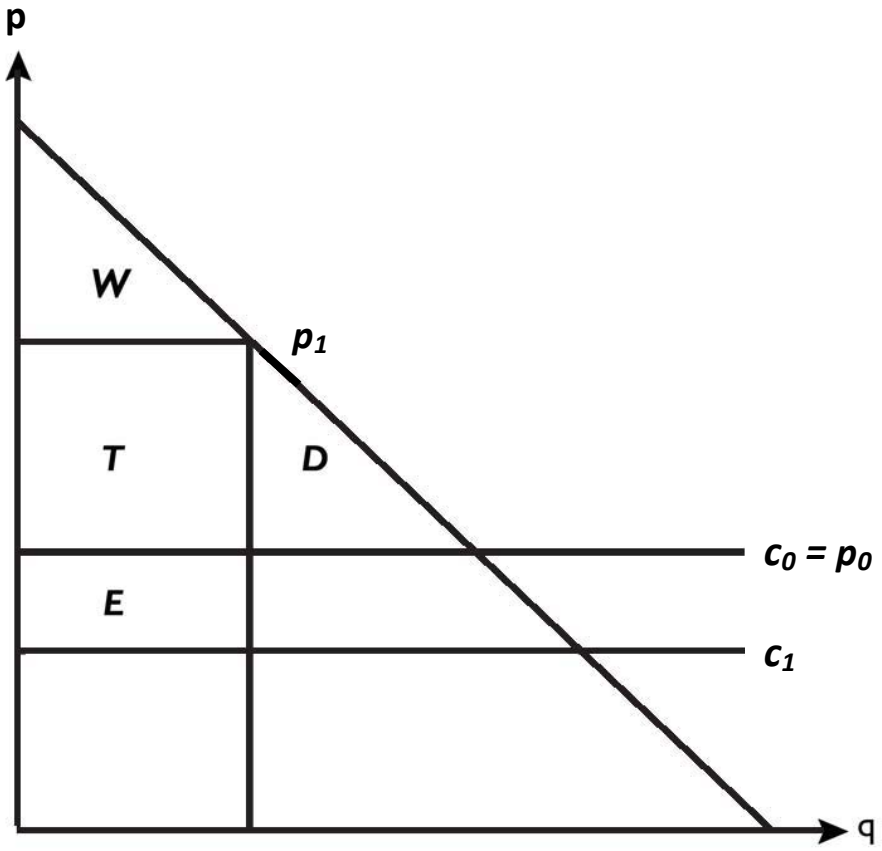


Figure 1: Monopolization with Cost Reduction

Actual Punishment Policies

The actual punishment policies of the U.S. and EU systems do not closely follow the prescriptions of optimal punishment theory. However, their key elements are distinctive and can be associated with the goals of harm internalization and complete deterrence.

Consider U.S. enforcement policy. The Sherman Act sets out a fixed maximum penalty of $100 million (Federal Trade Commission n.d., 1). In addition, U.S. law permits courts to deviate from the fixed maximum penalty in the statute by imposing a penalty equal to twice the loss imposed by the violator on consumers (18 U.S.C. §3571; Hylton 2003, 49). In addition, private lawsuits enable plaintiffs to sue for treble damages (see, e.g., Hylton 2003, 48-49). Private lawsuits outnumber government lawsuits by a 10-1 ratio (Salop and White 1986, 1003).

An optimal consumer harm-based penalty would divide the consumer harm, which is the sum of the wealth transfer and the forgone consumer surplus, by the probability that the monopolizing firm will be punished. Thus, going back to Figure 1, the optimal penalty is equal to $T+D$, divided by the probability of punishment. The reciprocal of the probability of punishment is therefore the "optimal multiplier" that should be applied to the consumer harm penalty.

American law, in contrast, imposes the transfer, T, as a penalty on the monopolizing firm, and employs either a multiplier of two under the fines enforcement law, or a multiplier of three for private lawsuits. These multipliers may or may not be optimal, depending on the relationship between the probability of punishment and the statutory multipliers of two and three. For example, if the probability of punishment – say because detection is difficult – is 1/10, then the multipliers on fines and damages would be too low to efficiently internalize the consumer harm.

Although the statutory multipliers may not be optimal, the American approach is broadly consistent with the harm internalization approach to punishment. The penalty for an antitrust violation varies directly with the magnitude of the consumer harm. Thus, the American antitrust punishment system is one in which violators pay a penalty that is proportional to consumer harm, where the proportionality factor is greater than one.

Now consider EU antitrust enforcement policy. The EC Treaty provides that the penalty for an infringement will be based on the total sales of the violator, with a maximum penalty equal to ten percent of the total revenue of the violator (Kaczorowska 2008, 873). Article 83(2) of the EC Treaty instructs that penalties should "ensure compliance with the prohibitions laid down in Article 81(1) and Article 82" (European Community Treaty 2002, art. 83). The relevant treaty provisions have been interpreted by enforcement officials and commentators to support the deterrence rather than the internalization objective (*Showa Denko v. Commission*, para. 58, cited in Wils 2007, 205 n.57; *Archer Daniels Midland v. Commission*, para. 49).[6]

[6] EU competition policy's emphasis on deterrence was also apparent when the EU Commission imposed double the basic fine against Microsoft in the EU-Microsoft cases (*see* Economides and Lianos 2010, 372).

The lesson suggested is that the EU penalties are designed to deter prohibited conduct completely by eliminating the prospect of earning profits through conduct that violates the antitrust laws. The EU penalties are *not* designed to vary directly with the amount of consumer harm; they are structured to vary directly with the amount of profit that a firm gains from conduct deemed to violate EU competition law.

Thus, the penalty provisions under U.S. and under EU antitrust law reflect the major economic theories of punishment, complete deterrence and internalization. The EU has embraced complete deterrence as the objective of enforcement. The U.S. has adopted the harm internalization approach.

One might be inclined to conclude that the EU has the relatively inefficient enforcement system. Penalties, such as those under the EU system, that run proportional to revenues punish efficient as well as inefficient monopolizing conduct with equal severity. The U.S. system enables firms that intend to engage in efficient monopolizing acts to go forward and reap the rewards, as long as those rewards are sufficiently greater than the harm to consumers.

On closer inspection, the comparison, on efficiency grounds, of enforcement provisions in the U.S. and EU requires consideration of more minute features. First, consider the distinction between *cartelization* (anticompetitive agreements among firms, such as price-fixing) and *monopolization* (anticompetitive single-firm conduct).

Linking monetary fines to revenues, as the EU system does, may be efficient overall when applied to cartelization. If the cartels do not have an efficiency basis, as the enforcement structures in the U.S. and in the EU assume, then the complete deterrence objective is optimal. It happens that a policy of imposing penalties that are a multiple of consumer harm will achieve the complete deterrence objective just as effectively as a policy that aims to eliminate gains. But the policy of complete deterrence, rather than harm internalization, remains the correct policy on social welfare grounds.

If a cartel does have an efficiency basis, then the EU system would impose inefficiently large penalties – assessing a penalty based on the sum of the wealth transfer and the efficiency gain ($T+E$) instead of limiting the penalty to the wealth transfer (T). However, both punishment systems are based on the assumption that the vast majority of cartelization cases are inefficient wealth transfers. If this assumption is correct, then the EU system's average degree of inefficiency in cartel punishment cases may be less than that of the U.S. The reason is that the EU would impose the correct level of the fine in the vast majority of cases, and an inefficiently large fine in a minority of cases. Under the same assumptions, the U.S. fine, limited to the transfer, might fall short of providing optimal deterrence because it fails to fully internalize consumer harm.

To clarify this argument, consider the following hypothetical. Suppose a firm that is (correctly) found guilty of participating in an inefficient price-fixing conspiracy has a total annual revenue of $20 billion. Suppose the firm's gross gain from the conspiracy

(*T*) is $150 million, and the forgone consumer surplus (*D*) is $75 million. The penalty under U.S. law could be as high as $300 million – since the rules permit courts to deviate from the $100 million Sherman Act limit up to twice the gross gain from the conspiracy. The penalty under EU law could be as high as $2 billion (10 percent of revenue). If the likelihood of being punished for the conspiracy is 1/10, the amount necessary to internalize consumer harm is $2.25 billion, and the amount necessary to eliminate the prospect of gain is $1.5 billion. The Sherman Act penalty of $300 million would be too low, from either the internalization or the complete deterrence perspective. The EU rules, however, could result in an optimal penalty for deterrence purposes. If the EU penalty were set at a level between $1.5 and $2 billion, it would eliminate the gain prospectively, which is sufficient to meet the optimal punishment goal in this scenario. If the EU penalty were set at its maximum of $2 billion, it would come close to internalizing the consumer harm (the wealth transfer and the foregone consumer surplus).

Moreover, the U.S. enforcement system imposes imprisonment as a punishment on relevant officials in price-fixing cases, while the EU system limits punishment to the imposition of monetary fines (Connor 2001, 89). The imprisonment of firm officials may offset the tendency toward underdeterrence suggested in the numerical example just offered, but it depends on many factors that are shrouded in uncertainty. Imprisonment concentrates the risk of punishment on the minority of actors who are responsible for the price-fixing agreement. If those actors are not aware of the risk of imprisonment, they will not be deterred. If they are aware of the risk, they are likely to be deterred, but even this conclusion is uncertain because it depends on the actors' risk preferences and psychological attachments. For the relevant actors, the expected cost of imprisonment, given the low probability, may appear to be small in comparison to the rewards.

In addition, imprisonment generates excessive litigation, as firms and targeted officials spend huge sums to avoid the punishment. Some of the executives imprisoned have long productive records as employees in their industries; they are not common criminals. Society forfeits the value of their services by locking them up.

These points of comparison suggest that the broad-brush, gain-based penalties under the EU system may be superior on social welfare grounds to the penalty system enacted under the US antitrust laws for cartelization cases. At the least, it is unclear a priori whether the U.S. or the EU has the socially preferable system for punishing cartel activity.

With respect to monopolization, a different assessment seems appropriate. In the monopolization context, the EU fine system appears to be inferior on welfare grounds to the U.S. punishment system. The EU system aims to strip the gains from monopolizing conduct, which deters both efficient as well as inefficient conduct. The U.S. punishment system, in contrast, discriminates between efficient and inefficient monopolizing conduct. The discrimination process is not perfectly optimal, but it is probably superior on social welfare grounds to the EU punishment system.

To clarify this argument, return to Figure 1. Under the U.S. system, the monopolizing firm would be required to pay a penalty that is likely to be a multiple of the consumer harm – a multiple of two under public enforcement, and three under private enforcement. If the efficiency gain from its conduct is sufficiently large, the U.S. punishment system will not deter a firm from engaging in efficient monopolizing conduct.

For example, suppose, in an particular instance of monopolization, the wealth transfer is $150 million, the forgone consumer surplus is $75 million, and the efficiency gain is $200 million. Assume also that the likelihood of being punished is fifty percent – which reflects the greater likelihood of detection and enforcement in monopolization cases. Thus, the firm's gain from monopolization, which is equal to the sum of the wealth transfer and the efficiency gain, is $350 million. Moreover, since the efficiency gain exceeds the forgone consumer surplus, this is a case of efficient monopolization. The penalty imposed under the U.S. system would be either $300 million (twice the transfer) under public enforcement or $450 million (three times the transfer) under private enforcement. In either case (public or private enforcement), the monopolizing firm would not be deterred under the U.S. punishment system. The reason is that the expected penalty, whether under public or under private enforcement, would be less than the gain from monopolization: under public enforcement, the expected penalty would be $150 million, which is less than the $350 million gain from monopolization; and under private enforcement the expected penalty would be $225 million, which is also less than the $350 million gain from monopolization. The monopolizing firm would not be deterred by the threat of penalization under the U.S. system, and this is the efficient result.

Under the EU system, the monopolizing firm would be required to pay a penalty that is likely to be a multiple of the sum of the transfer and the efficiency gain ($T+E$). Thus, in a case of efficient monopolizing conduct, the firm definitely would lose rather than gain after taking the penalty into account. For example, suppose a firm with annual revenue of $20 billion engages in a monopolizing act in some part of its business. The transfer from consumers is $150 million, the forgone consumer surplus is $75 million, and the efficiency gain is $200 million. The firm's total gain from monopolization is therefore $350 million. As in the previous example, this is a case of efficient monopolization. The EU could impose a fine on the firm as high as $2 billion.[7] If it chooses a fine greater than $700 million, though still well below the penalty ceiling, it would completely deter the monopolizing act, an inefficient outcome.

The general picture that emerges from this comparison is that the U.S. antitrust law enforcement system is compatible with the efficiency goal while EU law is not. EU law has followed Bentham by setting fines sufficiently large to wipe out the gains from conduct deemed unlawful. U.S. law has followed Becker by keeping damages closely tied to the harms suffered by consumers.

[7] I have assumed that the firm is large (revenue $20 billion). Obviously, if the firm's revenue is not much more than the amount it receives from the monopolizing activity, the EU fine may be far less than the amount required to deter.

The inefficiency of the EU enforcement system is not limited to the EU. China's competition law enforcement system is modeled on the EU's provisions (Wei 2011, 812-815), as are the enforcement rules of several other competition enforcement authorities, including Pakistan and Singapore (Wilson 2011, 111; Ong 2007, 109-110). The prevalence of the EU enforcement model indicates a strong preference for enforcement provisions that threaten fines tied to gain, as a form of trade tariff, rather than fines tied to consumer harm.

Substantive Standards: Predation

The second important area in which to examine the differences between the major competition regimes is substantive law. There are many parts of substantive law to examine. At the most general level, however, the antitrust laws fall into one of two categories: prohibitions on cartelization, or prohibitions on monopolization. Since the substantive policies with respect to cartelization are virtually the same in all competition regimes, I will focus on monopolization law; specifically, predatory pricing.

A predatory pricing claim is an assertion by one firm, the predation target or victim, that it has been injured by the low prices of another, the incumbent dominant firm. The victim typically argues that the dominant firm cut its price during a predatory campaign, in order to force the victim to sustain losses that would compel it to leave the market. After the victim leaves the market, the dominant firm, if all goes according to the theory, raises its price to the monopoly level. During the period in which the dominant firm prices at the monopoly level, it recoups the losses that it suffered during the predatory pricing campaign.

Predatory pricing claims in the U.S. fall under Section 2 of the Sherman Act. The law under Section 2 puts high hurdles in the way of predation plaintiffs. In order to avoid a summary judgment in a predation lawsuit, the plaintiff must present evidence that the defendant set its price below some reasonable proxy for marginal cost, and that the market structure is such that it would permit the defendant to recoup its losses from the predation campaign, after the victim has been forced out of the market (*Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 222-224). Both requirements are difficult to meet.

The EU law on predatory pricing imposes a lighter burden on the plaintiff. Under EU law, unlawful predation is established if the evidence shows that the defendant sets its price below average variable cost (*see, e.g., AKZO Chemie BV v. Commission*, para. 71). If the defendant set its price below average cost but above average variable cost, then predation can be established if the evidence suggests that it was accompanied by an intention to exclude the plaintiff (*AKZO Chemie BV v. Commission*, para. 72). The evidence required to prove predatory intent includes objective factors, such as the duration of the predatory period and the number of units sold at the allegedly predatory price. The evidentiary requirements suggest that, in practice, a significant burden falls on the defendant to disprove predatory intent when price is below average cost and above average variable cost.

The difference between the U.S. and EU standards on price predation reduces to this: the U.S. uses a marginal cost test and the EU uses an average cost test. Figure 2 illustrates the fundamental differences between the U.S. and EU with respect to predation. Price cuts below the marginal cost curve (MC) are predatory in the U.S., provided market structure evidence shows the plausibility of recoupment (*Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 210). In the EU, price cuts below the average cost curve (AC) are predatory. Since marginal cost is nearly impossible to measure precisely, the U.S. law encourages courts to examine reasonable proxies to marginal cost (*Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 223). However, the goal is to use a measure that approximates marginal cost (*United States v. AMR Corp.*, 1115-1116).

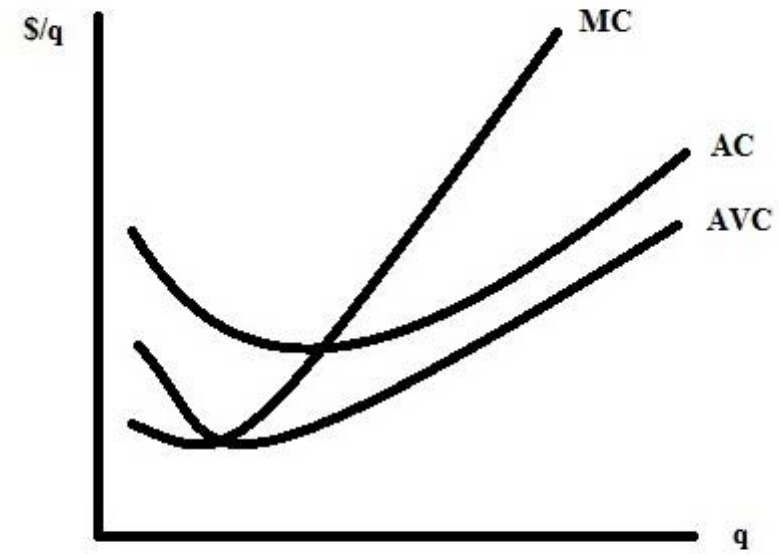


Figure 2: Marginal Cost, Average Cost, and Average Variable Cost

The U.S. marginal cost test, as Areeda and Turner (1975, 701-702) argued, is the efficient standard. When price is above marginal cost, a cut in price moving it in the direction of marginal cost, along with an associated increase in consumption, enhances social welfare. Conversely, when price is below marginal cost, a cut in price moving it away from marginal cost reduces social welfare. The reason is that price, under ideal conditions, reflects the marginal benefit to society from producing an extra unit of a good. Marginal cost, under ideal conditions, reflects the resource cost to society of supplying an additional unit of a good to the market. As long as price exceeds marginal costs,

society's welfare can be enhanced by expanding consumption. Hence, a marginal cost test for predation is consistent with a policy of enhancing society's welfare.

The average cost test of the EU disregards the efficiency principle and creates a price umbrella based on the dominant firm's average total cost. As long as a rival firm can match the average total cost of the dominant firm, it is shielded from additional price pressure under the EU law. In addition, given the uncertainties in measuring cost, and the amount and duration of losses a predatory target firm must experience before it is financially compelled to leave the market, the EU's average-cost standard effectively shields relatively inefficient firms from vigorous price competition.

The precise function or objective of the EU predation standard has never been set out clearly. One possible justification is that an average cost pricing standard enhances the set of options to consumers, by preventing efficient price predation. If this is indeed the purpose of the standard, it is unlikely that it enhances consumer welfare. It preserves relatively inefficient rivals and forces consumers to pay higher prices. The consumers themselves probably would have chosen to pay lower prices for fewer retail options, as they often do when price-cutting firms displace higher-priced rivals.

Another possible justification for the EU average-cost standard is that it preserves employment, by reducing the frequency with which price competition leads to the exit of firms that are the victims of predatory pricing campaigns. Less efficient firms are more likely to survive under the EU standard. Perhaps in a state with a generous welfare system, the policy of preserving less efficient firms is less expensive for taxpayers than a policy that allows them to be driven out of business by efficient price predation. In other words, the average-cost standard may be, in essence, a public welfare policy. The efficiency of such a policy cannot be determined without taking into account the relative inefficiencies of the administrative state.

Procedure

The third important difference between U.S and EU antitrust law is procedure. I refer to procedure in the broadest sense; from the processes by which the legal standards evolve to the methods used to determine the validity of evidence.

Development of Law

The U.S. is unique among competition law regimes in that its law is developed through the common law process. The Sherman Act says relatively little (Sherman Act 2004, §§ 1 & 2). It can be accurately summarized by saying that it prohibits price-fixing and monopolization. The detailed rules that have developed in American antitrust law have almost all come out of the courts.

The EU law has been set out in a relatively sparse treaty. The European Community Treaty ("EC Treaty"), however, is more detailed in its statement of prohibitions than is the Sherman Act (European Community Treaty 2002, art. 81). Moreover, the precise

meanings of the key competition provisions of the EC Treaty, Article 81 (now Article 101 of the Treaty on the Functioning of the European Union ("TFEU")) and Article 82 (now Article 102 of the TFEU), have been developed for the most part by the European Commission. For example, the EU Commission has been instrumental in clarifying "Market Definition" under EU Competition Law, which has ramifications for both Article 81 and 82 enforcement (*see* Report Prepared for the Competition Directorate-General of the European Commission 2005, 5). The EU courts defer to the European Commission on issues concerning the interpretation of the competition rules (Marsden 2009, 27).[8] As a result, the system of EU courts is not the primary body that interprets the meaning of the competition provisions of the EC Treaty.

These differences in the processes under which competition law develops have important implications. The American process essentially grafts onto the common law an additional branch called antitrust law without changing the process by which cases are litigated or decided in any substantial manner. Under the common law process, courts independently develop a framework for applying legal rules and modifying them in light of facts or policy arguments.

The common law process permits both sides in litigation, plaintiff and defendant, to present their positions on the meaning of the law and the state of the evidence to an impartial observer, the court. The court inevitably has some degree of discretion in both matters. Cases that are easily decided on the basis of the statutory text, or on the basis of earlier decisions, do not continue for a long time in court. Judges dismiss them, or decide them quickly, or the parties settle. The vast majority of legal disputes that spend enough time in court to come before judges have sufficient uncertainty surrounding them that the judge inevitably has discretion to decide what the law requires as between the litigating parties. The discretion that judges have had under the U.S. antitrust laws is equivalent to the discretion they have had under the common law for centuries.

The discretion given to judges under the American process has permitted courts to consider the social consequences of their decisions, and to issue judgments that effectively reduce the social costs of the rules they administer. Common law judges, in the course of examining the consequences and implications of their decisions, trade off the social costs of false convictions and false acquittals. Judges do not have to discover the relevant social costs on their own; the litigating parties have strong incentives to bring this information directly to the judges' attention. Most likely because of this constant process of weighing cost tradeoffs, American antitrust law has tended toward adopting efficiency-based legal standards, such as the marginal cost based test in predatory pricing law.

[8] The European Court of First Instance reviews Commission decisions by a "manifest error of assessment" standard which considers "whether the facts on which the Commission's assessment was based were correct, whether the conclusions drawn from those facts were not clearly mistaken or inconsistent and whether all the relevant factors had been taken into account." This limited standard of review is deferential to the Commission.

The dominant role of the enforcement agent imparts a different tendency to the law's evolution in EU. The enforcement agent will tend to interpret the law not in a manner that impartially weighs or trades off social costs from false convictions and false acquittals, but in a manner that minimizes its own enforcement costs. The dominance of per se standards based on relatively simple and abstract rules in the EU can be explained by their utility to the enforcement agent (Evans and Ahlborn 2008, 29).

The importance of administrative facility and the relative detachment of courts from the law-generation process put the EU system closer to a civil law model rather than the common law model of the U.S. Given that EU courts are staffed largely with judges drawn from civil law countries the tendency for a civil law system to develop is natural (Eurofound 2011; Apple and Deyling 1995, 1).

The EU rule on predatory pricing, for example, is an administratively simple rule. It does not require the complainant to generate a reasonable proxy of the defendant's marginal cost, which is a difficult undertaking, both for the enforcement agent and for the court. It is not efficient as a rule governing competition. However, it is efficient on administrative grounds in comparison to the American rule.

Although I have referred to American and EU law, the EU pattern has been replicated in China and other countries. Thus, one could say that competition law outside of the U.S. is largely shaped by demands of the enforcement agent. Competition law in the U.S. is shaped by the traditions of common law courts.

Evidence and Procedure

The deference policy that EU courts have adopted with respect to the EC's competition enforcement decisions implies important differences in the assessment of evidence and findings of fact. In the U.S., antitrust enforcement agencies and private plaintiffs have to present evidence to a court and attempt to persuade the court that a conspiracy has occurred, or that a defendant should be deemed a monopolist. In the EU, the complainants face the same requirements but in the presence of the enforcement agent, rather than a court. If the enforcement agency has made a decision or is inclined to pursue a case, all evidence and policy arguments will be examined under the influence of that decision.

The distinction between the U.S. and EU systems is one between a court-centered process and an agency-centered process. In a court-centered process, as in the U.S., the plaintiff, whether a private plaintiff or an enforcement agency, knows that it will have to persuade a skeptical court of the validity of its arguments, and must prepare its case with this in mind. Under an agency-centered process, the agency has a relatively small likelihood of ending up in a court in a particular case. A complainant approaching the enforcement agency will have to persuade the agency that it has a colorable claim, which is less than the burden, to prove that a violation of the law has occurred, borne by a private complainant in a court.

Moreover, there is an important difference between the agency enforcement processes in the U.S. and in the EU with respect to the degree of separation between prosecutorial and adjudicative functions. The agency enforcement process in the U.S. is observed in practice when the FTC brings an enforcement action against an entity. Firms often approach the FTC with antitrust complaints against market-dominating competitors. If the FTC chooses to proceed with a complaint, it will either go to a federal court or to an administrative law judge, depending on the statute it seeks to enforce. While the FTC must bring enforcement actions for Clayton Act violations in federal court, the FTC may enforce Section 5 of the FTC Act through internal administrative litigation before an administrative law judge, a process known as Part III proceedings (Antitrust Modernization Commission 2007, 129). In either case, the FTC must prepare its arguments and evidence to withstand questioning by an independent official. Moreover, the FTC General Counsel's office is independent of the FTC Commissioners, the final decision-making body within the agency process (Coate and Kleit 1995, 1-2). There are substantial walls of separation between the prosecutorial and adjudicative arms of the FTC's enforcement process.

The agency process in the EU is distinguishable in the sense that the European Commission considers complaints and conducts investigations, but there is no separation between the prosecutorial arm and the adjudicative arm within the EC agency process. Evidence and arguments are not subjected to an independent assessment until they are appealed to the EU courts, which have adopted a deference policy with respect to such matters. And since the EU courts have adopted a deference policy, there really is no point at which the EC's evidence and arguments are subjected to a rigorous and independent evaluation of merit.

The enforcement system in China is based on the EU model. Moreover, the absence of democracy and of basic "rule of law" norms prevent the court system in China from becoming reliably independent of the interests of the government.

The upshot is that outside of the U.S., competition law procedure effectively combines prosecutorial and adjudicative functions. This is a very substantial chasm between the enforcement regimes. The procedures adopted under the EU model are inconsistent with fundamental requirements of due process in American law.

The differences in substantive law probably pale in importance when compared to the differences in procedure. If the EU adopted efficiency-based legal standards, or an efficiency-based approach to punishment, the potential improvements in welfare could easily be vacated through the absence of reasonable due process safeguards in the enforcement process.

Conclusion

With more than 100 competition law regimes, there are countless ways in which antitrust law regimes can vary around the world. However, two models are dominant in terms of their effects on global commerce: the U.S. and the EU. This chapter has compared those

models in terms of enforcement, substantive law, and procedure. With respect to enforcement and substantive law, the U.S. has evolved toward an efficiency-based system while the EU has not. The procedural differences are perhaps more important than the differences in substantive law. The EU process of law development is closer to the civil law model while the American process is safely within the common law tradition. Retail-level procedural issues, such as the treatment of evidence, reveal stark differences, the most significant of which being the relatively weak separation of prosecutorial and judicial functions within the EU enforcement process.

References

Antitrust Modernization Commission (2007), *Report and Recommendations*, <http://govinfo.library.unt.edu/amc/report_recommendation/amc_final_report.pdf>.

Apple, James G. and Robert P. Deyling (1995), *A Primer on the Civil-Law System*, Federal Judicial Center, <http://www.fjc.gov/public/pdf.nsf/lookup/civillaw.pdf/$file/civillaw.pdf>.

Areeda, Phillip and Donald F. Turner (1975), Predatory Pricing and Related Acts Under Section 2 of the Sherman Act, *Harvard Law Review* 88: 697-733.

Bartz, Diane and Foo Yun Chee (2012), Google gets U.S., EU nod to buy Motorola Mobility. *Reuters*, Feb. 14. <http://www.reuters.com/article/2012/02/14/us-google-motorola-eu-idUSTRE81C1HE20120214>.

Becker, Gary S. (1968), Crime and Punishment: An Economic Approach, *Journal of Political Economy* 76: 169-217.

Coate, Malcolm B. and Andrew N. Kleit (1995), The Political Economy of Federal Trade Commission Administrative Decision Making in Merger Enforcement, Federal Trade Commission Working Paper 210, <http://www.ftc.gov/be/workpapers/wp210.pdf>.

Connor, John (2001), *Global Price Fixing: Our Customers Are the Enemy*, Boston, MA: Kluwer Academic Publishers.

Economides, Nicholas and Ioannis Lianos (2010), A Critical Appraisal of Remedies in the E.U. Microsoft Cases, *Columbia Business Law Review* 2010(2): 346-420.

Eurofound (2011), European Court of Justice, <http://www.eurofound.europa.eu/areas/industrialrelations/dictionary/definitions/EUROPEANCOURTOFJUSTICE.htm>.

European Community Treaty (2002), Official Journal of the European Communities C 325/39, <http://eur-lex.europa.eu/en/treaties/dat/12002E/pdf/12002E_EN.pdf>.

Evans, David and Christian Ahlborn (2009), The Microsoft Judgment and its Implications for Competition Policy Towards Dominant Firms in Europe, *Antitrust Law Journal* 75: 887-932.

Farmer, Susan B. (2010), The Impact of China's Antitrust Law and Other Competition Policies on U.S. Companies, *Loyola Consumer Law Review* 23: 34-53.

Federal Sentencing Guidelines, U.S. Code. Title 18, 1984.

Federal Trade Commission (FTC) (n.d.), *An FTC Guide to the Antitrust Laws*, <http://ftc.gov/bc/antitrust/factsheets/antitrustlawsguide.pdf>.

Hylton, Keith N. (n.d.), *Antitrust World Reports*, http://antitrustworldwiki.com.

Hylton, Keith N. (1998), Punitive Damages and the Economic Theory of Penalties, *Georgetown Law Journal* 87: 421-471.

Hylton, Keith N. (2003), *Antitrust Law: Economic Theory and Common Law Evolution*, New York, NY: Cambridge University Press.

Hylton, Keith N. and Fei Deng (2007), Antitrust Around the World: An Empirical Analysis of the Scope of Competition Laws and Their Effects, *Antitrust Law Journal* 74: 271-341.

Hylton, Keith N. and Haizhen Lin (2010), Optimal Antitrust Enforcement, Dynamic Competition, and Changing Economic Conditions, *Antitrust Law Journal* 77: 247- 276.

Kaczorowska, Alina (2008), *European Union Law*, New York, NY: Routledge-Cavendish.

Landes, William M. (1983), Optimal Sanctions for Antitrust Violations, *University of Chicago Law Review* 50: 652-678.

Leslie, Christopher R. (1993), Achieving Efficiency Through Collusion: A Market Failure Defense to Horizontal Price-Fixing, *California Law Review* 81: 243-292.

Marsden, Phillip (2009), Checks and Balances: EU Competition Law and the Rule of Law, *Competition Law International* 5(1): 24-29.

Ong, Burton (2007), Competition Law Takes Off in Singapore: An Analysis of Two Recent Decisions, *Competition Policy International* 3(2): 101-131.

Posner, Richard A. (2001), *Antitrust Law*, Chicago, IL: University of Chicago Press.

Report Prepared for the Competition Directorate-General of the European Commission (2005), *Market Definitions in the Media Sector: Comparative Legal Analysis*, July. <http://ec.europa.eu/competition/sectors/media/documents/chapter_1_ec.pdf>.

Salop, Steven C. and Lawrence J. White (1986), Economic Analysis of Private Antitrust Litigation, *Georgetown Law Journal* 74: 1001-1064.

The Sherman Act, 15 U.S.C. § 1 *et seq.* (2004).

Wei, Dan (2011), China's Anti-Monopoly Law And Its Merger Enforcement: Convergence and Flexibility, *Journal of International Economic Law* 14: 807-844.

Whitney, Lance (2012), Chinese Regulators drag heels on Google-Motorola Deal, CNET, Mar. 20. <http://news.cnet.com/8301-13578_3-57400661-38/chinese-regulators-drag-heels-on-google-motorola-deal/>.

Wils, Wouter P.J. (2007), The European Commission's 2006 Guidelines on Antitrust Fines: A Legal and Economic Analysis, *World Competition* 30: 197-231.

Wilson, Joseph (2011), Crossing the Crossroads: Making Competition Law Effective in Pakistan, *Loyola University Chicago International Law Review* 8: 105-125.

Cases Cited

Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993).

Case 62/86**,** *AKZO Chemie BV v. Commission* [1991] ECR 3359.

Case 289/04, *Showa Denko v. Commission* [2006], Opinion of Advocate General Geelhoed, ECR 05859.

Judgment of the Court of First Instance, Case 329/01, *Archer Daniels Midland v. Commission* [2006] ECR 03255).

United States v. AMR Corp., 335 F.3d 1109 (10th Cir. 2003).

Exhibit 11 to the Cross-Examination of Professor Tadelis

Crime and Punishment: An Economic Approach

Author(s): Gary S. Becker

Source: *Journal of Political Economy*, Mar. - Apr., 1968, Vol. 76, No. 2 (Mar. - Apr., 1968), pp. 169-217

Published by: The University of Chicago Press

Stable URL: https://www.jstor.org/stable/1830482

Crime and Punishment: An Economic Approach

Gary S. Becker*

Columbia University

I. Introduction

Since the turn of the century, legislation in Western countries has expanded rapidly to reverse the brief dominance of laissez faire during the nineteenth century. The state no longer merely protects against violations of person and property through murder, rape, or burglary but also restricts "discrimination" against certain minorities, collusive business arrangements, "jaywalking," travel, the materials used in construction, and thousands of other activities. The activities restricted not only are numerous but also range widely, affecting persons in very different pursuits and of diverse social backgrounds, education levels, ages, races, etc. Moreover, the likelihood that an offender will be discovered and convicted and the nature and extent of punishments differ greatly from person to person and activity to activity. Yet, in spite of such diversity, some common properties are shared by practically all legislation, and these properties form the subject matter of this essay.

In the first place, obedience to law is not taken for granted, and public and private resources are generally spent in order both to prevent offenses and to apprehend offenders. In the second place, conviction is not generally considered sufficient punishment in itself; additional and sometimes severe punishments are meted out to those convicted. What determines the amount and type of resources and punishments used to enforce a piece of legislation? In particular, why does enforcement differ so greatly among different kinds of legislation?

* I would like to thank the Lilly Endowment for financing a very productive summer in 1965 at the University of California at Los Angeles. While there I received very helpful comments on an earlier draft from, among others, Armen Alchian, Roland McKean, Harold Demsetz, Jack Hirshliefer, William Meckling, Gordon Tullock, and Oliver Williamson. I have also benefited from comments received at seminars at the University of Chicago, Hebrew University, RAND Corporation, and several times at the Labor Workshop of Columbia; assistance and suggestions from Isaac Ehrlich and Robert Michael; and suggestions from the editor of this journal.

The main purpose of this essay is to answer normative versions of these questions, namely, how many resources and how much punishment *should* be used to enforce different kinds of legislation? Put equivalently, although more strangely, how many offenses *should* be permitted and how many offenders *should* go unpunished? The method used formulates a measure of the social loss from offenses and finds those expenditures of resources and punishments that minimize this loss. The general criterion of social loss is shown to incorporate as special cases, valid under special assumptions, the criteria of vengeance, deterrence, compensation, and rehabilitation that historically have figured so prominently in practice and criminological literature.

The optimal amount of enforcement is shown to depend on, among other things, the cost of catching and convicting offenders, the nature of punishments—for example, whether they are fines or prison terms—and the responses of offenders to changes in enforcement. The discussion, therefore, inevitably enters into issues in penology and theories of criminal behavior. A second, although because of lack of space subsidiary, aim of this essay is to see what insights into these questions are provided by our "economic" approach. It is suggested, for example, that a useful theory of criminal behavior can dispense with special theories of anomie, psychological inadequacies, or inheritance of special traits and simply extend the economist's usual analysis of choice.

II. Basic Analysis

A. The Cost of Crime

Although the word "crime" is used in the title to minimize terminological innovations, the analysis is intended to be sufficiently general to cover all violations, not just felonies—like murder, robbery, and assault, which receive so much newspaper coverage—but also tax evasion, the so-called white-collar crimes, and traffic and other violations. Looked at this broadly, "crime" is an economically important activity or "industry," notwithstanding the almost total neglect by economists.[1] Some relevant evidence recently put together by the President's Commission on Law

[1] This neglect probably resulted from an attitude that illegal activity is too immoral to merit any systematic scientific attention. The influence of moral attitudes on a scientific analysis is seen most clearly in a discussion by Alfred Marshall. After arguing that even fair gambling is an "economic blunder" because of diminishing marginal utility, he says, "It is true that this loss of probable happiness need not be greater than the pleasure derived from the excitement of gambling, and we are then thrown back upon the induction [*sic*] that pleasures of gambling are in Bentham's phrase 'impure'; since experience shows that they are likely to engender a restless, feverish character, unsuited for steady work as well as for the higher and more solid pleasures of life" (Marshall, 1961, Note X, Mathematical Appendix).

Enforcement and Administration of Justice (the "Crime Commission") is reproduced in Table 1. Public expenditures in 1965 at the federal, state, and local levels on police, criminal courts and counsel, and "corrections" amounted to over $4 billion, while private outlays on burglar alarms, guards, counsel, and some other forms of protection were about $2 billion. Unquestionably, public and especially private expenditures are significantly understated, since expenditures by many public agencies in the course of enforcing particular pieces of legislation, such as state fair-employment laws,[2] are not included, and a myriad of private precautions against crime, ranging from suburban living to taxis, are also excluded.

TABLE 1
ECONOMIC COSTS OF CRIMES

Type	Costs (Millions of Dollars)
Crimes against persons	815
Crimes against property	3,932
Illegal goods and services	8,075
Some other crimes	2,036
Total	14,858
Public expenditures on police, prosecution, and courts	3,178
Corrections	1,034
Some private costs of combatting crime	1,910
Over-all total	20,980

Source: President's Commission, (1967*d*, p. 44).

Table 1 also lists the Crime Commission's estimates of the direct costs of various crimes. The gross income from expenditures on various kinds of illegal consumption, including narcotics, prostitution, and mainly gambling, amounted to over $8 billion. The value of crimes against property, including fraud, vandalism, and theft, amounted to almost $4 billion,[3] while about $3 billion worth resulted from the loss of earnings due to homicide, assault, or other crimes. All the costs listed in the table total about $21 billion, which is almost 4 per cent of reported national

[2] Expenditures by the thirteen states with such legislation in 1959 totaled almost $2 million (see Landes, 1966).

[3] Superficially, frauds, thefts, etc., do not involve true social costs but are simply transfers, with the loss to victims being compensated by equal gains to criminals. While these are transfers, their market value is, nevertheless, a first approximation to the direct social cost. If the theft or fraud industry is "competitive," the sum of the value of the criminals' time input—including the time of "fences" and prospective time in prison—plus the value of capital input, compensation for risk, etc., would approximately equal the market value of the loss to victims. Consequently, aside from the input of intermediate products, losses can be taken as a measure of the value of the labor and capital input into these crimes, which are true social costs.

income in 1965. If the sizeable omissions were included, the percentage might be considerably higher.

Crime has probably become more important during the last forty years. The Crime Commission presents no evidence on trends in costs but does present evidence suggesting that the number of major felonies per capita has grown since the early thirties (President's Commission, 1967*a*, pp. 22–31). Moreover, with the large growth of tax and other legislation, tax evasion and other kinds of white-collar crime have presumably grown much more rapidly than felonies. One piece of indirect evidence on the growth of crime is the large increase in the amount of currency in circulation since 1929. For sixty years prior to that date, the ratio of currency either to all money or to consumer expenditures had declined very substantially. Since then, in spite of further urbanization and income growth and the spread of credit cards and other kinds of credit,[4] both ratios have increased sizeably.[5] This reversal can be explained by an unusual increase in illegal activity, since currency has obvious advantages over checks in illegal transactions (the opposite is true for legal transactions) because no record of a transaction remains.[6]

B. The Model

It is useful in determining how to combat crime in an optimal fashion to develop a model to incorporate the behavioral relations behind the costs listed in Table 1. These can be divided into five categories: the relations between (1) the number of crimes, called "offenses" in this essay, and the cost of offenses, (2) the number of offenses and the punishments meted out, (3) the number of offenses, arrests, and convictions and the public expenditures on police and courts, (4) the number of convictions and the costs of imprisonments or other kinds of punishments, and (5) the number of offenses and the private expenditures on protection and apprehension. The first four are discussed in turn, while the fifth is postponed until a later section.

1. Damages

Usually a belief that other members of society are harmed is the motivation behind outlawing or otherwise restricting an activity. The amount of harm

[4] For an analysis of the secular decline to 1929 that stresses urbanization and the growth in incomes, see Cagan (1965, chap. iv).

[5] In 1965, the ratio of currency outstanding to consumer expenditures was 0.08, compared to only 0.05 in 1929. In 1965, currency outstanding per family was a whopping $738.

[6] Cagan (1965, chap. iv) attributes much of the increase in currency holdings between 1929 and 1960 to increased tax evasion resulting from the increase in tax rates.

would tend to increase with the activity level, as in the relation

$$H_i = H_i(O_i),$$

with (1)

$$H_i' = \frac{dH_i}{dO_i} > 0,$$

where H_i is the harm from the ith activity and O_i is the activity level.[7] The concept of harm and the function relating its amount to the activity level are familiar to economists from their many discussions of activities causing external diseconomies. From this perspective, criminal activities are an important subset of the class of activities that cause diseconomies, with the level of criminal activities measured by the number of offenses.

The social value of the gain to offenders presumably also tends to increase with the number of offenses, as in

$$G = G(O),$$

with (2)

$$G' = \frac{dG}{dO} > 0.$$

The net cost or damage to society is simply the difference between the harm and gain and can be written as

$$D(O) = H(O) - G(O). \tag{3}$$

If, as seems plausible, offenders usually eventually receive diminishing marginal gains and cause increasing marginal harm from additional offenses, $G'' < 0$, $H'' > 0$, and

$$D'' = H'' - G'' > 0, \tag{4}$$

which is an important condition used later in the analysis of optimality positions (see, for example, the Mathematical Appendix). Since both H' and $G' > 0$, the sign of D' depends on their relative magnitudes. It follows from (4), however, that

$$D'(O) > 0 \text{ for all } O > O_a \text{ if } D'(O_a) \geq 0. \tag{5}$$

Until Section V the discussion is restricted to the region where $D' > 0$, the region providing the strongest justification for outlawing an activity. In that section the general problem of external diseconomies is reconsidered from our viewpoint, and there $D' < 0$ is also permitted.

The top part of Table 1 lists costs of various crimes, which have been interpreted by us as estimates of the value of resources used up in these

[7] The ith subscript will be suppressed whenever it is to be understood that only one activity is being discussed.

crimes. These values are important components of, but are not identical to, the net damages to society. For example, the cost of murder is measured by the loss in earnings of victims and excludes, among other things, the value placed by society on life itself; the cost of gambling excludes both the utility to those gambling and the "external" disutility to some clergy and others; the cost of "transfers" like burglary and embezzlement excludes social attitudes toward forced wealth redistributions and also the effects on capital accumulation of the possibility of theft. Consequently, the $15 billion estimate for the cost of crime in Table 1 may be a significant understatement of the net damages to society, not only because the costs of many white-collar crimes are omitted, but also because much of the damage is omitted even for the crimes covered.

2. The Cost of Apprehension and Conviction

The more that is spent on policemen, court personnel, and specialized equipment, the easier it is to discover offenses and convict offenders. One can postulate a relation between the output of police and court "activity" and various inputs of manpower, materials, and capital, as in $A = f(m, r, c)$, where f is a production function summarizing the "state of the arts." Given f and input prices, increased "activity" would be more costly, as summarized by the relation

$$C = C(A)$$

and (6)

$$C' = \frac{dC}{dA} > 0.$$

It would be cheaper to achieve any given level of activity the cheaper were policemen,[8] judges, counsel, and juries and the more highly developed the state of the arts, as determined by technologies like fingerprinting, wiretapping, computer control, and lie-detecting.[9]

One approximation to an empirical measure of "activity" is the number of offenses cleared by conviction. It can be written as

$$A \cong pO, \tag{7}$$

where p, the ratio of offenses cleared by convictions to all offenses, is the over-all probability that an offense is cleared by conviction. By substituting

[8] According to the Crime Commission, 85–90 per cent of all police costs consist of wages and salaries (President's Commission, 1967*a*, p. 35).

[9] A task-force report by the Crime Commission deals with suggestions for greater and more efficient usage of advanced technologies (President's Commission, 1967*e*).

(7) into (6) and differentiating, one has

$$C_p = \frac{\partial C(pO)}{\partial p} = C'O > 0$$

and (8)

$$C_o = C'p > 0$$

if $pO \neq 0$. An increase in either the probability of conviction or the number of offenses would increase total costs. If the marginal cost of increased "activity" were rising, further implications would be that

$$\begin{aligned} C_{pp} &= C''O^2 > 0, \\ C_{oo} &= C''p^2 > 0, \end{aligned} \tag{9}$$

and

$$C_{po} = C_{op} = C''pO + C' > 0.$$

A more sophisticated and realistic approach drops the implication of (7) that convictions alone measure "activity," or even that p and O have identical elasticities, and introduces the more general relation

$$A = h(p, O, a). \tag{10}$$

The variable a stands for arrests and other determinants of "activity," and there is no presumption that the elasticity of h with respect to p equals that with respect to O. Substitution yields the cost function $C = C(p, O, a)$. If, as is extremely likely, h_p, h_o, and h_a are all greater than zero, then clearly C_p, C_o, and C_a are all greater than zero.

In order to insure that optimality positions do not lie at "corners," it is necessary to place some restrictions on the second derivatives of the cost function. Combined with some other assumptions, it is *sufficient* that

$$\begin{aligned} C_{pp} &\geq 0, \\ C_{oo} &\geq 0, \end{aligned} \tag{11}$$

and

$$C_{po} \gtrless 0$$

(see the Mathematical Appendix). The first two restrictions are rather plausible, the third much less so.[10]

Table 1 indicates that in 1965 public expenditures in the United States on police and courts totaled more than $3 billion, by no means a minor

[10] Differentiating the cost function yields $C_{pp} = C''(h_p)^2 + C'h_{pp}$; $C_{oo} = C''(h_o)^2 + C'h_{oo}$; $C_{po} = C''h_oh_p + C'h_{po}$. If marginal costs were rising, C_{pp} or C_{oo} could be negative only if h_{pp} or h_{oo} were sufficiently negative, which is not very likely. However, C_{po} would be approximately zero only if h_{po} were sufficiently negative, which is also unlikely. Note that if "activity" is measured by convictions alone, $h_{pp} = h_{oo} = 0$, and $h_{po} > 0$.

item. Separate estimates were prepared for each of seven major felonies.[11] Expenditures on them averaged about $500 per offense (reported) and about $2,000 per person arrested, with almost $1,000 being spent per murder (President's Commission, 1967*a*, pp. 264–65); $500 is an estimate of the average cost

$$AC = \frac{C(p, O, a)}{O}$$

of these felonies and would presumably be a larger figure if the number of either arrests or convictions were greater. Marginal costs (C_o) would be at least $500 if condition (11), $C_{oo} \geq 0$, were assumed to hold throughout.

3. The Supply of Offenses

Theories about the determinants of the number of offenses differ greatly, from emphasis on skull types and biological inheritance to family upbringing and disenchantment with society. Practically all the diverse theories agree, however, that when other variables are held constant, an increase in a person's probability of conviction or punishment if convicted would generally decrease, perhaps substantially, perhaps negligibly, the number of offenses he commits. In addition, a common generalization by persons with judicial experience is that a change in the probability has a greater effect on the number of offenses than a change in the punishment,[12] although, as far as I can tell, none of the prominent theories shed any light on this relation.

The approach taken here follows the economists' usual analysis of choice and assumes that a person commits an offense if the expected utility to him exceeds the utility he could get by using his time and other resources at other activities. Some persons become "criminals," therefore, not because their basic motivation differs from that of other persons, but because their benefits and costs differ. I cannot pause to discuss the many general implications of this approach,[13] except to remark that criminal behavior becomes part of a much more general theory and does not require ad hoc concepts of differential association, anomie, and the like,[14] nor does it assume perfect knowledge, lightening-fast calculation, or any of the other caricatures of economic theory.

[11] They are willful homicide, forcible rape, robbery, aggravated assault, burglary, larceny, and auto theft.

[12] For example, Lord Shawness (1965) said, "Some judges preoccupy themselves with methods of punishment. This is their job. But in preventing crime it is of less significance than they like to think. Certainty of detection is far more important than severity of punishment." Also see the discussion of the ideas of C. B. Beccaria, an insightful eighteenth-century Italian economist and criminologist, in Radzinowicz (1948, I, 282).

[13] See, however, the discussions in Smigel (1965) and Ehrlich (1967).

[14] For a discussion of these concepts, see Sutherland (1960).

This approach implies that there is a function relating the number of offenses by any person to his probability of conviction, to his punishment if convicted, and to other variables, such as the income available to him in legal and other illegal activities, the frequency of nuisance arrests, and his willingness to commit an illegal act. This can be represented as

$$O_j = O_j(p_j, f_j, u_j), \tag{12}$$

where O_j is the number of offenses he would commit during a particular period, p_j his probability of conviction per offense, f_j his punishment per offense, and u_j a portmanteau variable representing all these other influences.[15]

Since only convicted offenders are punished, in effect there is "price discrimination" and uncertainty: if convicted, he pays f_j per convicted offense, while otherwise he does not. An increase in either p_j or f_j would reduce the utility expected from an offense and thus would tend to reduce the number of offenses because either the probability of "paying" the higher "price" or the "price" itself would increase.[16] That is,

$$O_{p_j} = \frac{\partial O_j}{\partial p_j} < 0$$

and (13)

$$O_{f_j} = \frac{\partial O_j}{\partial f_j} < 0,$$

which are the generally accepted restrictions mentioned above. The effect of changes in some components of u_j could also be anticipated. For example, a rise in the income available in legal activities or an increase in law-abidingness due, say, to "education" would reduce the incentive to enter illegal activities and thus would reduce the number of offenses. Or a shift in the form of the punishment, say, from a fine to imprisonment,

[15] Both p_j and f_j might be considered distributions that depend on the judge, jury, prosecutor, etc., that j happens to receive. Among other things, u_j depends on the p's and f's meted out for other competing offenses. For evidence indicating that offenders do substitute among offenses, see Smigel (1965).

[16] The utility expected from committing an offense is defined as

$$EU_j = p_j U_j(Y_j - f_j) + (1 - p_j)U_j(Y_j),$$

where Y_j is his income, monetary plus psychic, from an offense; U_j is his utility function; and f_j is to be interpreted as the monetary equivalent of the punishment. Then

$$\frac{\partial EU_j}{\partial p_j} = U_j(Y_j - f_j) - U_j(Y_j) < 0$$

and

$$\frac{\partial EU_j}{\partial f_j} = -p_j U_j'(Y_j - f_j) < 0$$

as long as the marginal utility of income is positive. One could expand the analysis by incorporating the costs and probabilities of arrests, detentions, and trials that do not result in conviction.

would tend to reduce the number of offenses, at least temporarily, because they cannot be committed while in prison.

This approach also has an interesting interpretation of the presumed greater response to a change in the probability than in the punishment. An increase in p_j "compensated" by an equal percentage reduction in f_j would not change the expected income from an offense[17] but could change the expected utility, because the amount of risk would change. It is easily shown that an increase in p_j would reduce the expected utility, and thus the number of offenses, more than an equal percentage increase in f_j[18] if j has preference for risk; the increase in f_j would have the greater effect if he has aversion to risk; and they would have the same effect if he is risk neutral.[19] The widespread generalization that offenders are more deterred by the probability of conviction than by the punishment when convicted turns out to imply in the expected-utility approach that offenders are risk preferrers, at least in the relevant region of punishments.

The total number of offenses is the sum of all the O_j and would depend on the set of p_j, f_j, and u_j. Although these variables are likely to differ significantly between persons because of differences in intelligence, age, education, previous offense history, wealth, family upbringing, etc., for simplicity I now consider only their average values, p, f, and u,[20] and write the market offense function as

$$O = O(p, f, u). \tag{14}$$

This function is assumed to have the same kinds of properties as the individual functions, in particular, to be negatively related to p and f and to be more responsive to the former than the latter if, and only if, offenders on balance have risk preference. Smigel (1965) and Ehrlich (1967) estimate

[17] $EY_j = p_j(Y_j - f_j) + (1 - p_j)Y_j = Y_j - p_j f_j$.

[18] This means that an increase in p_j "compensated" by a reduction in f_j would reduce utility and offenses.

[19] From n. 16

$$\frac{-\partial EU_j}{\partial p_j}\frac{p_j}{U_j} = [U_j(Y_j) - U_j(Y_j - f_j)]\frac{p_j}{U_j} \gtreqless \frac{-\partial EU_j}{\partial f_j}\frac{f_j}{U_j} = p_j U_j'(Y_j - f_j)\frac{f_j}{U_j}$$

as

$$\frac{U_j(Y_j) - U_j(Y_j - f_j)}{f_j} \gtreqless U_j'(Y_j - f_j).$$

The term on the left is the average change in utility between $Y_j - f_j$ and Y_j. It would be greater than, equal to, or less than $U_j'(Y_j - f_j)$ as $U_j'' \gtreqless 0$. But risk preference is defined by $U_j'' > 0$, neutrality by $U_j'' = 0$, and aversion by $U_j'' < 0$.

[20] p can be defined as a weighted average of the p_j, as

$$p = \sum_{j=1}^{n} \frac{O_j p_j}{\sum_{i=1}^{n} O_i},$$

and similar definitions hold for f and u.

functions like (14) for seven felonies reported by the Federal Bureau of Investigation using state data as the basic unit of observation. They find that the relations are quite stable, as evidenced by high correlation coefficients; that there are significant negative effects on O of p and f; and that usually the effect of p exceeds that of f, indicating preference for risk in the region of observation.

A well-known result states that, in equilibrium, the real incomes of persons in risky activities are, at the margin, relatively high or low as persons are generally risk avoiders or preferrers. If offenders were risk preferrers, this implies that the real income of offenders would be lower, at the margin, than the incomes they could receive in less risky legal activities, and conversely if they were risk avoiders. Whether "crime pays" is then an implication of the attitudes offenders have toward risk and is not directly related to the efficiency of the police or the amount spent on combatting crime. If, however, risk were preferred at some values of p and f and disliked at others, public policy could influence whether "crime pays" by its choice of p and f. Indeed, it is shown later that the social loss from illegal activities is usually minimized by selecting p and f in regions where risk is preferred, that is, in regions where "crime does not pay."

4. Punishments

Mankind has invented a variety of ingenious punishments to inflict on convicted offenders: death, torture, branding, fines, imprisonment, banishment, restrictions on movement and occupation, and loss of citizenship are just the more common ones. In the United States, less serious offenses are punished primarily by fines, supplemented occasionally by probation, petty restrictions like temporary suspension of one's driver's license, and imprisonment. The more serious offenses are punished by a combination of probation, imprisonment, parole, fines, and various restrictions on choice of occupation. A recent survey estimated for an average day in 1965 the number of persons who were either on probation, parole, or institutionalized in a jail or juvenile home (President's Commission 1967*b*). The total number of persons in one of these categories came to about 1,300,000, which is about 2 per cent of the labor force. About one-half were on probation, one-third were institutionalized, and the remaining one-sixth were on parole.

The cost of different punishments to an offender can be made comparable by converting them into their monetary equivalent or worth, which, of course, is directly measured only for fines. For example, the cost of an imprisonment is the discounted sum of the earnings foregone and the value placed on the restrictions in consumption and freedom. Since the earnings foregone and the value placed on prison restrictions vary from person to person, the cost even of a prison sentence of given duration is

not a unique quantity but is generally greater, for example, to offenders who could earn more outside of prison.[21] The cost to each offender would be greater the longer the prison sentence, since both foregone earnings and foregone consumption are positively related to the length of sentences.

Punishments affect not only offenders but also other members of society. Aside from collection costs, fines paid by offenders are received as revenue by others. Most punishments, however, hurt other members as well as offenders: for example, imprisonment requires expenditures on guards, supervisory personnel, buildings, food, etc. Currently about $1 billion is being spent each year in the United States on probation, parole, and institutionalization alone, with the daily cost per case varying tremendously from a low of $0.38 for adults on probation to a high of $11.00 for juveniles in detention institutions (President's Commission, 1967*b*, pp. 193–94).

The total social cost of punishments is the cost to offenders plus the cost or minus the gain to others. Fines produce a gain to the latter that equals the cost to offenders, aside from collection costs, and so the social cost of fines is about zero, as befits a transfer payment. The social cost of probation, imprisonment, and other punishments, however, generally exceeds that to offenders, because others are also hurt. The derivation of optimality conditions in the next section is made more convenient if social costs are written in terms of offender costs as

$$f' \equiv bf, \tag{15}$$

where f' is the social cost and b is a coefficient that transforms f into f'. The size of b varies greatly between different kinds of punishments: $b \cong 0$ for fines, while $b > 1$ for torture, probation, parole, imprisonment, and most other punishments. It is especially large for juveniles in detention homes or for adults in prisons and is rather close to unity for torture or for adults on parole.

III. Optimality Conditions

The relevant parameters and behavioral functions have been introduced, and the stage is set for a discussion of social policy. If the aim simply were deterrence, the probability of conviction, p, could be raised close to 1, and punishments, f, could be made to exceed the gain: in this way the number of offenses, O, could be reduced almost at will. However, an increase in p increases the social cost of offenses through its effect on the cost of combatting offenses, C, as does an increase in f if $b > 0$ through the effect on the cost of punishments, bf. At relatively modest values of p and f, these effects might outweigh the social gain from increased deterrence. Similarly,

[21] In this respect, imprisonment is a special case of "waiting time" pricing that is also exemplified by queuing (see Becker, 1965, esp. pp. 515–16, and Kleinman, 1967).

if the aim simply were to make "the punishment fit the crime," p could be set close to 1, and f could be equated to the harm imposed on the rest of society. Again, however, such a policy ignores the social cost of increases in p and f.

What is needed is a criterion that goes beyond catchy phrases and gives due weight to the damages from offenses, the costs of apprehending and convicting offenders, and the social cost of punishments. The social-welfare function of modern welfare economics is such a criterion, and one might assume that society has a function that measures the social loss from offenses. If

$$L = L(D, C, bf, O) \tag{16}$$

is the function measuring social loss, with presumably

$$\frac{\partial L}{\partial D} > 0, \quad \frac{\partial L}{\partial C} > 0, \quad \frac{\partial L}{\partial bf} > 0, \tag{17}$$

the aim would be to select values of f, C, and possibly b that minimize L.

It is more convenient and transparent, however, to develop the discussion at this point in terms of a less general formulation, namely, to assume that the loss function is identical with the total social loss in real income from offenses, convictions, and punishments, as in

$$L = D(O) + C(p, O) + bpfO. \tag{18}$$

The term $bpfO$ is the total social loss from punishments, since bf is the loss per offense punished and pO is the number of offenses punished (if there are a fairly large number of independent offenses). The variables directly subject to social control are the amounts spent in combatting offenses, C; the punishment per offense for those convicted, f; and the form of punishments, summarized by b. Once chosen, these variables, via the D, C, and O functions, indirectly determine p, O, D, and ultimately the loss L.

Analytical convenience suggests that p rather than C be considered a decision variable. Also, the coefficient b is assumed in this section to be a given constant greater than zero. Then p and f are the only decision variables, and their optimal values are found by differentiating L to find the two first-order optimality conditions,[22]

$$\frac{\partial L}{\partial f} = D'O_f + C'O_f + bpfO_f + bpO = 0 \tag{19}$$

and

$$\frac{\partial L}{\partial p} = D'O_p + C'O_p + C_p + bpfO_p + bfO = 0. \tag{20}$$

[22] The Mathematical Appendix discusses second-order conditions.

If O_f and O_p are not equal to zero, one can divide through by them, and recombine terms, to get the more interesting expressions

$$D' + C' = -bpf\left(1 - \frac{1}{\varepsilon_f}\right) \tag{21}$$

and

$$D' + C' + C_p \frac{1}{O_p} = -bpf\left(1 - \frac{1}{\varepsilon_p}\right), \tag{22}$$

where

$$\varepsilon_f = -\frac{f}{O} O_f$$

and (23)

$$\varepsilon_p = -\frac{p}{O} O_p.$$

The term on the left side of each equation gives the marginal cost of increasing the number of offenses, O: in equation (21) through a reduction in f and in (22) through a reduction in p. Since $C' > 0$ and O is assumed to be in a region where $D' > 0$, the marginal cost of increasing O through f must be positive. A reduction in p partly reduces the cost of combatting offenses, and, therefore, the marginal cost of increasing O must be less when p rather than when f is reduced (see Fig. 1); the former could even be negative if C_p were sufficiently large. Average "revenue," given by $-bpf$, is negative, but marginal revenue, given by the right-hand side of

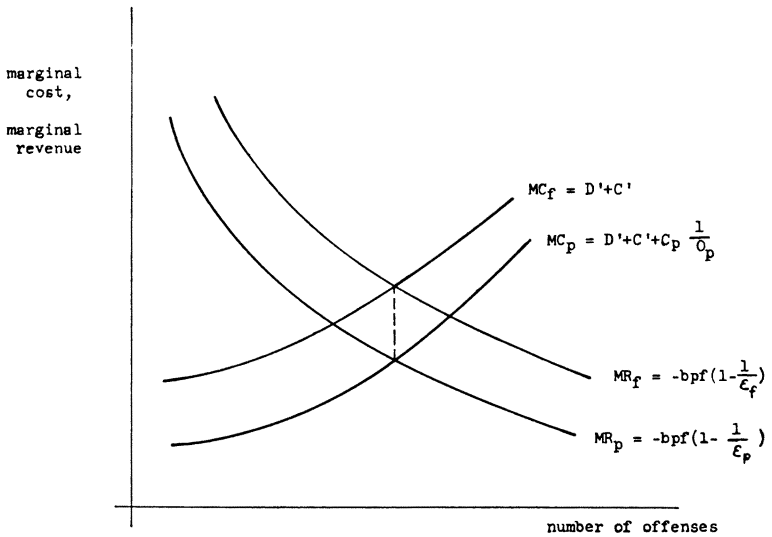


FIG. 1

equations (21) and (22), is not necessarily negative and would be positive if the elasticities ε_p and ε_f were less than unity. Since the loss is minimized when marginal revenue equals marginal cost (see Fig. 1), the optimal value of ε_f must be less than unity, and that of ε_p could only exceed unity if C_p were sufficiently large. This is a reversal of the usual equilibrium condition for an income-maximizing firm, which is that the elasticity of demand must exceed unity, because in the usual case average revenue is assumed to be positive.[23]

Since the marginal cost of changing O through a change in p is less than that of changing O through f, the equilibrium marginal revenue from p must also be less than that from f. But equations (21) and (22) indicate that the marginal revenue from p can be less if, and only if, $\varepsilon_p > \varepsilon_f$. As pointed out earlier, however, this is precisely the condition indicating that offenders have preference for risk and thus that "crime does not pay." Consequently, the loss from offenses is minimized if p and f are selected from those regions where offenders are, on balance, risk preferrers. Although only the attitudes offenders have toward risk can directly determine whether "crime pays," rational public policy indirectly insures that "crime does not pay" through its choice of p and f.[24]

I indicated earlier that the actual p's and f's for major felonies in the United States generally seem to be in regions where the effect (measured by elasticity) of p on offenses exceeds that of f, that is, where offenders are risk preferrers and "crime does not pay" (Smigel, 1965; Ehrlich, 1967). Moreover, both elasticities are generally less than unity. In both respects, therefore, actual public policy is consistent with the implications of the optimality analysis.

If the supply of offenses depended only on pf—offenders were risk neutral—a reduction in p "compensated" by an equal percentage increase in f would leave unchanged pf, O, $D(O)$, and $bpfO$ but would reduce the loss, because the costs of apprehension and conviction would be lowered by the reduction in p. The loss would be minimized, therefore, by lowering p arbitrarily close to zero and raising f sufficiently high so that the product pf would induce the optimal number of offenses.[25] A fortiori, if offenders

[23] Thus if $b < 0$, average revenue would be positive and the optimal value of ε_f would be greater than 1, and that of ε_p could be less than 1 only if C_p were sufficiently large.

[24] If $b < 0$, the optimality condition is that $\varepsilon_p < \varepsilon_f$, or that offenders are risk avoiders. Optimal social policy would then be to select p and f in regions where "crime does pay."

[25] Since $\varepsilon_f = \varepsilon_p = \varepsilon$ if O depends only on pf, and $C = 0$ if $p = 0$, the two equilibrium conditions given by eqs. (21) and (22) reduce to the single condition

$$D' = -bpf\left(1 - \frac{1}{\varepsilon}\right).$$

From this condition and the relation $O = O(pf)$, the equilibrium values of O and pf could be determined.

were risk avoiders, the loss would be minimized by setting p arbitrarily close to zero, for a "compensated" reduction in p reduces not only C but also O and thus D and $bpfO$.[26]

There was a tendency during the eighteenth and nineteenth centuries in Anglo-Saxon countries, and even today in many Communist and underdeveloped countries, to punish those convicted of criminal offenses rather severely, at the same time that the probability of capture and conviction was set at rather low values.[27] A promising explanation of this tendency is that an increased probability of conviction obviously absorbs public and private resources in the form of more policemen, judges, juries, and so forth. Consequently, a "compensated" reduction in this probability obviously reduces expenditures on combatting crime, and, since the expected punishment is unchanged, there is no "obvious" offsetting increase in either the amount of damages or the cost of punishments. The result can easily be continuous political pressure to keep police and other expenditures relatively low and to compensate by meting out strong punishments to those convicted.

Of course, if offenders are risk preferrers, the loss in income from offenses is generally minimized by selecting positive and finite values of p and f, even though there is no "obvious" offset to a compensated reduction in p. One possible offset already hinted at in footnote 27 is that judges or juries may be unwilling to convict offenders if punishments are set very high. Formally, this means that the cost of apprehension and conviction, C, would depend not only on p and O but also on f.[28] If C were more responsive to f than p, at least in some regions,[29] the loss in income could be minimized at finite values of p and f even if offenders were risk avoiders. For then a compensated reduction in p could raise, rather than lower, C and thus contribute to an increase in the loss.

Risk avoidance might also be consistent with optimal behavior if the loss function were not simply equal to the reduction in income. For example, suppose that the loss were increased by an increase in the ex post "price discrimination" between offenses that are not and those that are cleared by punishment. Then a "compensated" reduction in p would

[26] If $b < 0$, the optimal solution is p about zero and f arbitrarily high if offenders are either risk neutral or risk preferrers.

[27] For a discussion of English criminal law in the eighteenth and nineteenth centuries, see Radzinowicz (1948, Vol. I). Punishments were severe then, even though the death penalty, while legislated, was seldom implemented for less serious criminal offenses.

Recently South Vietnam executed a prominent businessman allegedly for "speculative" dealings in rice, while in recent years a number of persons in the Soviet Union have either been executed or given severe prison sentences for economic crimes.

[28] I owe the emphasis on this point to Evsey Domar.

[29] This is probably more likely for higher values of f and lower values of p.

increase the "price discrimination," and the increased loss from this could more than offset the reductions in C, D, and $bpfO$.[30]

IV. Shifts in the Behavioral Relations

This section analyzes the effects of shifts in the basic behavioral relations—the damage, cost, and supply-of-offenses functions—on the optimal values of p and f. Since rigorous proofs can be found in the Mathematical Appendix, here the implications are stressed, and only intuitive proofs are given. The results are used to explain, among other things, why more damaging offenses are punished more severely and more impulsive offenders less severely.

An increase in the marginal damages from a given number of offenses, D', increases the marginal cost of changing offenses by a change in either p or f (see Fig. 2*a* and *b*). The optimal number of offenses would necessarily decrease, because the optimal values of both p and f would increase. In this case (and, as shortly seen, in several others), the optimal values of p and f move in the same, rather than in opposite, directions.[31]

An interesting application of these conclusions is to different kinds of offenses. Although there are few objective measures of the damages done

[30] If p is the probability that an offense would be cleared with the punishment f, then $1 - p$ is the probability of no punishment. The expected punishment would be $\mu = pf$, the variance $\sigma^2 = p(1 - p)f^2$, and the coefficient of variation

$$v = \frac{\sigma}{\mu} = \sqrt{\frac{1 - p}{p}};$$

v increases monotonically from a low of zero when $p = 1$ to an infinitely high value when $p = 0$.

If the loss function equaled

$$L' = L + \psi(v), \qquad \psi' > 0,$$

the optimality conditions would become

$$D' + C' = -bpf\left(1 - \frac{1}{\varepsilon_f}\right) \tag{21}$$

and

$$D' + C' + C_p\frac{1}{O_p} + \psi'\frac{dv}{dp}\frac{1}{O_p} = -bpf\left(1 - \frac{1}{\varepsilon_p}\right). \tag{22}$$

Since the term $\psi'(dv/dp)(1/O_p)$ is positive, it could more than offset the negative term $C_p(1/O_p)$.

[31] I stress this primarily because of Bentham's famous and seemingly plausible dictum that "the more deficient in certainty a punishment is, the severer it should be" (1931, chap. ii of section entitled "Of Punishment," second rule). The dictum would be correct if p (or f) were exogenously determined and if L were minimized with respect to f (or p) alone, for then the optimal value of f (or p) would be inversely related to the given value of p (or f) (see the Mathematical Appendix). If, however, L is minimized with respect to both, then frequently they move in the same direction.

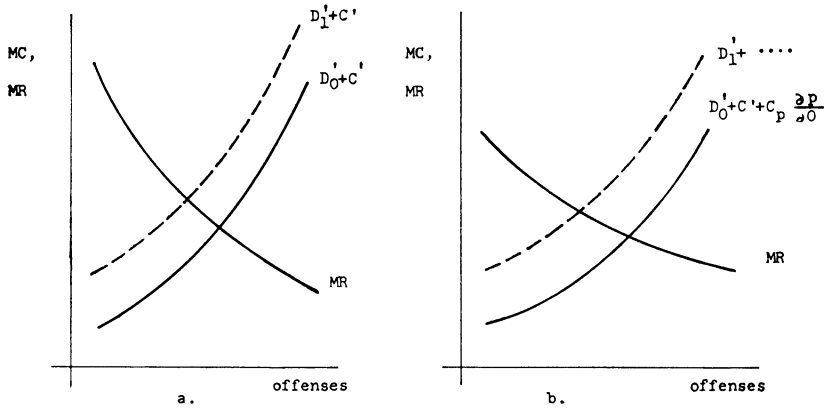


FIG. 2

by most offenses, it does not take much imagination to conclude that offenses like murder or rape generally do more damage than petty larceny or auto theft. If the other components of the loss in income were the same, the optimal probability of apprehension and conviction and the punishment when convicted would be greater for the more serious offenses.

Table 2 presents some evidence on the actual probabilities and punishments in the United States for seven felonies. The punishments are simply the average prison sentences served, while the probabilities are ratios of the estimated number of convictions to the estimated number of offenses and unquestionably contain a large error (see the discussions in Smigel, 1965, and Ehrlich, 1967). If other components of the loss function are ignored, and if actual and optimal probabilities and punishments are positively related, one should find that the more serious felonies have higher probabilities and longer prison terms. And one does: in the table, which lists the felonies in decreasing order of presumed seriousness, both the actual probabilities and the prison terms are positively related to seriousness.

Since an increase in the marginal cost of apprehension and conviction for a given number of offenses, C', has identical effects as an increase in marginal damages, it must also reduce the optimal number of offenses and increase the optimal values of p and f. On the other hand, an increase in the other component of the cost of apprehension and conviction, C_p, has no direct effect on the marginal cost of changing offenses with f and *reduces* the cost of changing offenses with p (see Fig. 3). It therefore reduces the optimal value of p and only partially compensates with an increase in f, so that the optimal number of offenses increases. Accordingly, an increase in both C' and C_p must increase the optimal f but can

TABLE 2

PROBABILITY OF CONVICTION AND AVERAGE PRISON TERM FOR SEVERAL MAJOR FELONIES, 1960

	Murder and Non-negligent Manslaughter	Forcible Rape	Robbery	Aggravated Assault	Burglary	Larceny	Auto Theft	All These Felonies Combined
1. Average time served (months) before first release:								
a) Federal civil institutions	111.0	63.6	56.1	27.1	26.2	16.2	20.6	18.8
b) State institutions	121.4	44.8	42.4	25.0	24.6	19.8	21.3	28.4
2. Probabilities of apprehension and conviction (per cent):								
a) Those found guilty of offenses known	57.9	37.7	25.1	27.3	13.0	10.7	13.7	15.1
b) Those found guilty of offenses charged	40.7	26.9	17.8	16.1	10.2	9.8	11.5	15.0
c) Those entering federal and state prisons (excludes many juveniles)	39.8	22.7	8.4	3.0	2.4	2.2	2.1	2.8

Source: 1, Bureau of Prisons (1960, Table 3); 2 (*a*) and (*b*), Federal Bureau of Investigation (1960, Table 10); 2 (*c*), Federal Bureau of Investigation (1961, Table 2), Bureau of Prisons (n.d., Table A1; 1961, Table 8).

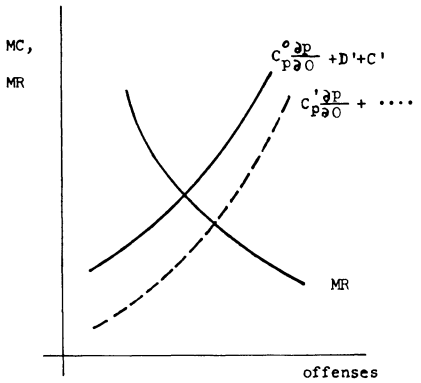


FIG. 3

either increase or decrease the optimal p and optimal number of offenses, depending on the relative importance of the changes in C' and C_p.

The cost of apprehending and convicting offenders is affected by a variety of forces. An increase in the salaries of policemen increases both C' and C_p, while improved police technology in the form of fingerprinting, ballistic techniques, computer control, and chemical analysis, or police and court "reform" with an emphasis on professionalism and merit, would tend to reduce both, not necessarily by the same extent. Our analysis implies, therefore, that although an improvement in technology and reform may or may not increase the optimal p and reduce the optimal number of offenses, it does reduce the optimal f and thus the need to rely on severe punishments for those convicted. Possibly this explains why the secular improvement in police technology and reform has gone hand in hand with a secular decline in punishments.

C_p, and to a lesser extent C', differ significantly between different kinds of offenses. It is easier, for example, to solve a rape or armed robbery than a burglary or auto theft, because the evidence of personal identification is often available in the former and not in the latter offenses.[32] This might tempt one to argue that the p's decline significantly as one moves across Table 2 (left to right) primarily because the C_p's are significantly lower for the "personal" felonies listed to the left than for the "impersonal" felonies listed to the right. But this implies that the f's would increase as one moved across the table, which is patently false. Consequently, the positive correlation between p, f, and the severity of offenses observed in

[32] "If a suspect is neither known to the victim nor arrested at the scene of the crime, the chances of ever arresting him are very slim" (President's Commission, 1967*e*, p. 8). This conclusion is based on a study of crimes in parts of Los Angeles during January, 1966.

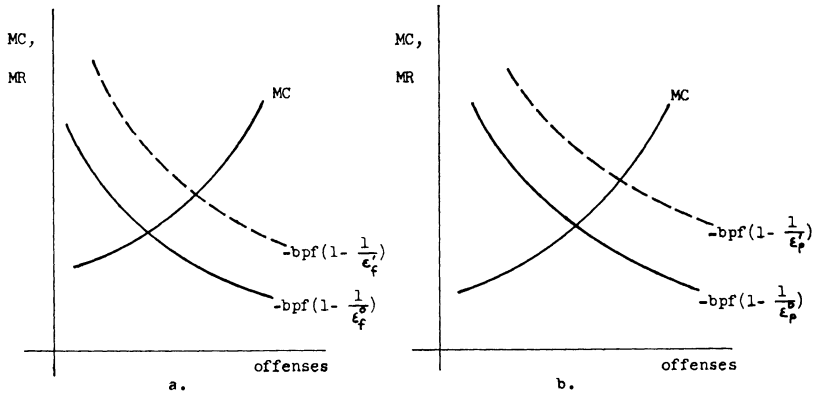


FIG. 4

the table cannot be explained by a negative correlation between C_p (or C') and severity.

If $b > 0$, a reduction in the elasticity of offenses with respect to f increases the marginal revenue of changing offenses by changing f (see Fig. 4*a*). The result is an increase in the optimal number of offenses and a decrease in the optimal f that is partially compensated by an increase in the optimal p. Similarly, a reduction in the elasticity of offenses with respect to p also increases the optimal number of offenses (see Fig. 4*b*), decreases the optimal p, and partially compensates by an increase in f. An equal percentage reduction in both elasticities a fortiori increases the optimal number of offenses and also tends to reduce both p and f. If $b = 0$, both marginal revenue functions lie along the horizontal axis, and changes in these elasticities have no effect on the optimal values of p and f.

The income of a firm would usually be larger if it could separate, at little cost, its total market into submarkets that have substantially different elasticities of demand: higher prices would be charged in the submarkets having lower elasticities. Similarly, if the total "market" for offenses could be separated into submarkets that differ significantly in the elasticities of supply of offenses, the results above imply that if $b > 0$ the total loss would be reduced by "charging" *lower* "prices"—that is, lower p's and f's—in markets with *lower* elasticities.

Sometimes it is possible to separate persons committing the same offense into groups that have different responses to punishments. For example, unpremeditated murderers or robbers are supposed to act impulsively and, therefore, to be relatively unresponsive to the size of punishments; likewise, the insane or the young are probably less affected

than other offenders by future consequences and, therefore,[33] probably less deterred by increases in the probability of conviction or in the punishment when convicted. The trend during the twentieth century toward relatively smaller prison terms and greater use of probation and therapy for such groups and, more generally, the trend away from the doctrine of "a given punishment for a given crime" is apparently at least broadly consistent with the implications of the optimality analysis.

An increase in b increases the marginal revenue from changing the number of offenses by changing p or f and thereby increases the optimal number of offenses, reduces the optimal value of f, and increases the optimal value of p. Some evidence presented in Section II indicates that b is especially large for juveniles in detention homes or adults in prison and is small for fines or adults on parole. The analysis implies, therefore, that other things the same, the optimal f's would be smaller and the optimal p's larger if punishment were by one of the former rather than one of the latter methods.

V. Fines

A. Welfare Theorems and Transferable Pricing

The usual optimality conditions in welfare economics depend only on the levels and not on the slopes of marginal cost and average revenue functions, as in the well-known condition that marginal costs equal prices. The social loss from offenses was explicitly introduced as an application of the approach used in welfare economics, and yet slopes as incorporated into elasticities of supply do significantly affect the optimality conditions. Why this difference? The primary explanation would appear to be that it is almost always implicitly assumed that prices paid by consumers are fully transferred to firms and governments, so that there is no social loss from payment.

If there were no social loss from punishments, as with fines, b would equal zero, and the elasticity of supply would drop out of the optimality condition given by equation (21).[34] If $b > 0$, as with imprisonment, some of the payment "by" offenders would not be received by the rest of society, and a net social loss would result. The elasticity of the supply of offenses then becomes an important determinant of the optimality conditions, because it determines the change in social costs caused by a change in punishments.

[33] But see Becker (1962) for an analysis indicating that impulsive and other "irrational" persons may be as deterred from purchasing a commodity whose price has risen as more "rational" persons.

[34] It remains in eq. (22), through the slope O_p, because ordinarily prices do not affect marginal costs, while they do here through the influence of p on C.

Although transferable monetary pricing is the most common kind today, the other is not unimportant, especially in underdeveloped and Communist countries. Examples in addition to imprisonment and many other punishments are the draft, payments in kind, and queues and other waiting-time forms of rationing that result from legal restrictions on pricing (see Becker, 1965) and from random variations in demand and supply conditions. It is interesting, and deserves further exploration, that the optimality conditions are so significantly affected by a change in the assumptions about the transferability of pricing.

B. Optimality Conditions

If $b = 0$, say, because punishment was by fine, and if the cost of apprehending and convicting offenders were also zero, the two optimality conditions (21) and (22) would reduce to the same simple condition

$$D'(O) = 0. \tag{24}$$

Economists generally conclude that activities causing "external" harm, such as factories that pollute the air or lumber operations that strip the land, should be taxed or otherwise restricted in level until the marginal external harm equalled the marginal private gain, that is, until marginal net damages equalled zero, which is what equation (24) says. If marginal harm always exceeded marginal gain, the optimum level would be presumed to be zero, and that would also be the implication of (24) when suitable inequality conditions were brought in. In other words, if the costs of apprehending, convicting, and punishing offenders were nil and if each offense caused more external harm than private gain, the social loss from offenses would be minimized by setting punishments high enough to eliminate all offenses. Minimizing the social loss would become identical with the criterion of minimizing crime by setting penalties sufficiently high.[35]

Equation (24) determines the optimal number of offenses, $\hat{O}$, and the fine and probability of conviction must be set at levels that induce offenders to commit just $\hat{O}$ offenses. If the economists' usual theory of choice is applied to illegal activities (see Sec. II), the marginal value of these penalties has to equal the marginal private gain:

$$V = G'(\hat{O}), \tag{25}$$

where $G'(\hat{O})$ is the marginal private gain at $\hat{O}$ and V is the monetary value of the marginal penalties. Since by equations (3) and (24), $D'(\hat{O}) = H'(\hat{O}) - G'(\hat{O}) = 0$, one has by substitution in (25)

$$V = H'(\hat{O}). \tag{26}$$

[35] "The evil of the punishment must be made to exceed the advantage of the offense" (Bentham, 1931, first rule).

The monetary value of the penalties would equal the marginal harm caused by offenses.

Since the cost of apprehension and conviction is assumed equal to zero, the probability of apprehension and conviction could be set equal to unity without cost. The monetary value of penalties would then simply equal the fines imposed, and equation (26) would become

$$f = H'(\hat{O}). \tag{27}$$

Since fines are paid by offenders to the rest of society, a fine determined by (27) would exactly compensate the latter for the marginal harm suffered, and the criterion of minimizing the social loss would be identical, at the margin, with the criterion of compensating "victims."[36] If the harm to victims always exceeded the gain to offenders, both criteria would reduce in turn to eliminating all offenses.

If the cost of apprehension and conviction were not zero, the optimality condition would have to incorporate marginal costs as well as marginal damages and would become, if the probability of conviction were still assumed to equal unity,

$$D'(\hat{O}) + C'(\hat{O}, 1) = 0. \tag{28}$$

Since $C' > 0$, (28) requires that $D' < 0$ or that the marginal private gain exceed the marginal external harm, which generally means a smaller number of offenses than when $D' = 0$.[37] It is easy to show that equation (28) would be satisfied if the fine equalled the sum of marginal harm and marginal costs:

$$f = H'(\hat{O}) + C'(\hat{O}, 1).^{38} \tag{29}$$

In other words, offenders have to compensate for the cost of catching them as well as for the harm they directly do, which is a natural generalization of the usual externality analysis.

The optimality condition

$$D'(\hat{O}) + C'(\hat{O}, \hat{p}) + C_p(\hat{O}, \hat{p}) \frac{1}{O_p} = 0 \tag{30}$$

would replace equation (28) if the fine rather than the probability of

[36] By "victims" is meant the rest of society and not just the persons actually harmed.

[37] This result can also be derived as a special case of the results in the Mathematical Appendix on the effects of increases in C'.

[38] Since equilibrium requires that $f = G'(\hat{O})$, and since from (28)

$$D'(\hat{O}) = H'(\hat{O}) - G'(\hat{O}) = -C'(\hat{O}, 1),$$

then (29) follows directly by substitution.

conviction were fixed. Equation (30) would usually imply that $D'(\hat{O}) > 0$,[39] and thus that the number of offenses would exceed the optimal number when costs were zero. Whether costs of apprehension and conviction increase or decrease the optimal number of offenses largely depends, therefore, on whether penalties are changed by a change in the fine or in the probability of conviction. Of course, if both are subject to control, the optimal probability of conviction would be arbitrarily close to zero, unless the social loss function differed from equation (18) (see the discussion in Sec. III).

C. *The Case for Fines*

Just as the probability of conviction and the severity of punishment are subject to control by society, so too is the form of punishment: legislation usually specifies whether an offense is punishable by fines, probation, institutionalization, or some combination. Is it merely an accident, or have optimality considerations determined that today, in most countries, fines are the predominant form of punishment, with institutionalization reserved for the more serious offenses? This section presents several arguments which imply that social welfare is increased if fines are used *whenever feasible*.

In the first place, probation and institutionalization use up social resources, and fines do not, since the latter are basically just transfer payments, while the former use resources in the form of guards, supervisory personnel, probation officers, and the offenders' own time.[40] Table 1 indicates that the cost is not minor either: in the United States in 1965, about $1 billion was spent on "correction," and this estimate excludes, of course, the value of the loss in offenders' time.[41]

[39] That is, if, as seems plausible,

$$\frac{dC}{dp} = C' \frac{\partial O}{\partial p} + C_p > 0,$$

then

$$C' + C_p \frac{1}{\partial O/\partial p} < 0,$$

and

$$D'(\hat{O}) = -\left(C' + C_p \frac{1}{\partial O/\partial p}\right) > 0.$$

[40] Several early writers on criminology recognized this advantage of fines. For example, "Pecuniary punishments are highly economical, since all the evil felt by him who pays turns into an advantage for him who receives" (Bentham, 1931, chap. vi), and "Imprisonment would have been regarded in these old times [*ca.* tenth century] as a useless punishment; it does not satisfy revenge, it keeps the criminal idle, and do what we may, *it is costly*" (Pollock and Maitland, 1952, p. 516; my italics).

[41] On the other hand, some transfer payments in the form of food, clothing, and shelter are included.

Moreover, the determination of the optimal number of offenses and severity of punishments is somewhat simplified by the use of fines. A wise use of fines requires knowledge of marginal gains and harm and of marginal apprehension and conviction costs; admittedly, such knowledge is not easily acquired. A wise use of imprisonment and other punishments must know this too, however, and, in addition, must know about the elasticities of response of offenses to changes in punishments. As the bitter controversies over the abolition of capital punishment suggest, it has been difficult to learn about these elasticities.

I suggested earlier that premeditation, sanity, and age can enter into the determination of punishments as proxies for the elasticities of response. These characteristics may not have to be considered in levying fines, because the optimal fines, as determined, say, by equations (27) or (29), do not depend on elasticities. Perhaps this partly explains why economists discussing externalities almost never mention motivation or intent, while sociologists and lawyers discussing criminal behavior invariably do. The former assume that punishment is by a monetary tax or fine, while the latter assume that non-monetary punishments are used.

Fines provide compensation to victims, and optimal fines at the margin fully compensate victims and restore the status quo ante, so that they are no worse off than if offenses were not committed.[42] Not only do other punishments fail to compensate, but they also require "victims" to spend additional resources in carrying out the punishment. It is not surprising, therefore, that the anger and fear felt toward ex-convicts who in fact have *not* "paid their debt to society" have resulted in additional punishments,[43] including legal restrictions on their political and economic opportunities[44] and informal restrictions on their social acceptance. Moreover, the absence of compensation encourages efforts to change and otherwise "rehabilitate" offenders through psychiatric counseling, therapy, and other programs. Since fines do compensate and do not create much additional cost, anger toward and fear of appropriately fined persons do not easily develop. As a result, additional punishments are not usually levied against "ex-finees," nor are strong efforts made to "rehabilitate" them.

One argument made against fines is that they are immoral because, in effect, they permit offenses to be bought for a price in the same way that

[42] Bentham recognized this and said, "To furnish an indemnity to the injured party is another useful quality in a punishment. It is a means of accomplishing two objects at once—punishing an offense and repairing it: removing the evil of the first order, and putting a stop to alarm. This is a characteristic advantage of pecuniary punishments" (1931, chap. vi).

[43] In the same way, the guilt felt by society in using the draft, a forced transfer *to* society, has led to additional payments to veterans in the form of education benefits, bonuses, hospitalization rights, etc.

[44] See Sutherland (1960, pp. 267–68) for a list of some of these.

bread or other goods are bought for a price.[45] A fine *can* be considered the price of an offense, but so too can any other form of punishment; for example, the "price" of stealing a car might be six months in jail. The only difference is in the units of measurement: fines are prices measured in monetary units, imprisonments are prices measured in time units, etc. If anything, monetary units are to be preferred here as they are generally preferred in pricing and accounting.

Optimal fines determined from equation (29) depend only on the marginal harm and cost and not at all on the economic positions of offenders. This has been criticized as unfair, and fines proportional to the incomes of offenders have been suggested.[46] If the goal is to minimize the social loss in income from offenses, and not to take vengeance or to inflict harm on offenders, then fines should depend on the total harm done by offenders, and not directly on their income, race, sex, etc. In the same way, the monetary value of optimal prison sentences and other punishments depends on the harm, costs, and elasticities of response, but not directly on an offender's income. Indeed, if the monetary value of the punishment by, say, imprisonment were independent of income, the length of the sentence would be *inversely* related to income, because the value placed on a given sentence is positively related to income.

We might detour briefly to point out some interesting implications for the probability of conviction of the fact that the monetary value of a given fine is obviously the same for all offenders, while the monetary equivalent or "value" of a given prison sentence or probation period is generally positively related to an offender's income. The discussion in Section II suggested that actual probabilities of conviction are not fixed to all offenders but usually vary with their age, sex, race, and, in particular, income. Offenders with higher earnings have an incentive to spend more on planning their offenses, on good lawyers, on legal appeals, and even on bribery to reduce the probability of apprehension and conviction for offenses punishable by, say, a given prison term, because the cost to them of conviction is relatively large compared to the cost of these expenditures.

[45] The very early English law relied heavily on monetary fines, even for murder, and it has been said that "every kind of blow or wound given to every kind of person had its price, and much of the jurisprudence of the time must have consisted of a knowledge of these preappointed prices" (Pollock and Maitland, 1952, p. 451).

The same idea was put amusingly in a recent *Mutt and Jeff* cartoon which showed a police car carrying a sign that read: "Speed limit 30 M per H—$5 fine every mile over speed limit—pick out speed you can afford."

[46] For example, Bentham said, "A pecuniary punishment, if the sum is fixed, is in the highest degree unequal. . . . Fines have been determined without regard to the profit of the offense, to its evil, or to the wealth of the offender. . . . Pecuniary punishments should always be regulated by the fortune of the offender. The relative amount of the fine should be fixed, not its absolute amount; for such an offense, such a part of the offender's fortune" (1931, chap. ix). Note that optimal fines, as determined by eq. (29), do depend on "the profit of the offense" and on "its evil."

Similarly, however, poorer offenders have an incentive to use more of their time in planning their offenses, in court appearances, and the like to reduce the probability of conviction for offenses punishable by a given fine, because the cost to them of conviction is relatively large compared to the value of their time.[47] The implication is that the probability of conviction would be systematically related to the earnings of offenders: negatively for offenses punishable by imprisonment and positively for those punishable by fines. Although a negative relation for felonies and other offenses punishable by imprisonment has been frequently observed and deplored (see President's Commission, 1967*c*, pp. 139–53), I do not know of any studies of the relation for fines or of any recognition that the observed negative relation may be more a consequence of the nature of the punishment than of the influence of wealth.

Another argument made against fines is that certain crimes, like murder or rape, are so heinous that no amount of money could compensate for the harm inflicted. This argument has obvious merit and is a special case of the more general principle that fines cannot be relied on exclusively whenever the harm exceeds the resources of offenders. For then victims could not be fully compensated by offenders, and fines would have to be supplemented with prison terms or other punishments in order to discourage offenses optimally. This explains why imprisonments, probation, and parole are major punishments for the more serious felonies; considerable harm is inflicted, and felonious offenders lack sufficient resources to compensate. Since fines are preferable, it also suggests the need for a flexible system of instalment fines to enable offenders to pay fines more readily and thus avoid other punishments.

This analysis implies that if some offenders could pay the fine for a given offense and others could not,[48] the former should be punished solely by fine and the latter partly by other methods. In essence, therefore, these methods become a vehicle for punishing "debtors" to society. Before the cry is raised that the system is unfair, especially to poor offenders, consider the following.

Those punished would be debtors in "transactions" that were never agreed to by their "creditors," not in voluntary transactions, such as loans,[49] for which suitable precautions could be taken in advance by creditors. Moreover, punishment in any economic system based on

[47] Note that the incentive to use time to reduce the probability of a given prison sentence is unrelated to earnings, because the punishment is fixed in time, not monetary, units; likewise, the incentive to use money to reduce the probability of a given fine is also unrelated to earnings, because the punishment is fixed in monetary, not time, units.

[48] In one study, about half of those convicted of misdemeanors could not pay the fines (see President's Commission, 1967*c*, p. 148).

[49] The "debtor prisons" of earlier centuries generally housed persons who could not repay loans.

voluntary market transactions inevitably must distinguish between such "debtors" and others. If a rich man purchases a car and a poor man steals one, the former is congratulated, while the latter is often sent to prison when apprehended. Yet the rich man's purchase is equivalent to a "theft" subsequently compensated by a "fine" equal to the price of the car, while the poor man, in effect, goes to prison because he cannot pay this "fine."

Whether a punishment like imprisonment in lieu of a full fine for offenders lacking sufficient resources is "fair" depends, of course, on the length of the prison term compared to the fine.[50] For example, a prison term of one week in lieu of a $10,000 fine would, if anything, be "unfair" to wealthy offenders paying the fine. Since imprisonment is a more costly punishment to society than fines, the loss from offenses would be reduced by a policy of leniency toward persons who are imprisoned because they cannot pay fines. Consequently, optimal prison terms for "debtors" would not be "unfair" to them in the sense that the monetary equivalent to them of the prison terms would be less than the value of optimal fines, which in turn would equal the harm caused or the "debt."[51]

It appears, however, that "debtors" are often imprisoned at rates of exchange with fines that place a low value on time in prison. Although I have not seen systematic evidence on the different punishments actually offered convicted offenders, and the choices they made, many statutes in

[50] Yet without any discussion of the actual alternatives offered, the statement is made that "the money judgment assessed the punitive damages defendant hardly seems comparable in effect to the criminal sanctions of death, imprisonment, and stigmatization" ("Criminal Safeguards . . .," 1967).

[51] A formal proof is straightforward if for simplicity the probability of conviction is taken as equal to unity. For then the sole optimality condition is

$$D' + C' = -bf\left(1 - \frac{1}{\varepsilon_f}\right). \tag{1'}$$

Since $D' = H' - G'$, by substitution one has

$$G' = H' + C' + bf\left(1 - \frac{1}{\varepsilon_f}\right), \tag{2'}$$

and since equilibrium requires that $G' = f$,

$$f = H' + C' + bf\left(1 - \frac{1}{\varepsilon_f}\right), \tag{3'}$$

or

$$f = \frac{H' + C'}{1 - b(1 - 1/\varepsilon_f)}. \tag{4'}$$

If $b > 0$, $\varepsilon_f < 1$ (see Sec. III), and hence by eq. (4'),

$$f < H' + C', \tag{5'}$$

where the term on the right is the full marginal harm. If p as well as f is free to vary, the analysis becomes more complicated, but the conclusion about the relative monetary values of optimal imprisonments and fines remains the same (see the Mathematical Appendix).

the United States do permit fines and imprisonment that place a low value on time in prison. For example, in New York State, Class A Misdemeanors can be punished by a prison term as long as one year or a fine no larger than $1,000 and Class B Misdemeanors, by a term as long as three months or a fine no larger than $500 (*Laws of New York*, 1965, chap. 1030, Arts. 70 and 80).[52] According to my analysis, these statutes permit excessive prison sentences relative to the fines, which may explain why imprisonment in lieu of fines is considered unfair to poor offenders, who often must "choose" the prison alternative.

D. Compensation and the Criminal Law

Actual criminal proceedings in the United States appear to seek a mixture of deterrence, compensation, and vengeance. I have already indicated that these goals are somewhat contradictory and cannot generally be simultaneously achieved; for example, if punishment were by fine, minimizing the social loss from offenses would be equivalent to compensating "victims" fully, and deterrence or vengeance could only be partially pursued. Therefore, if the case for fines were accepted, and punishment by optimal fines became the norm, the traditional approach to criminal law would have to be significantly modified.

First and foremost, the primary aim of all legal proceedings would become the same: not punishment or deterrence, but simply the assessment of the "harm" done by defendants. Much of traditional criminal law would become a branch of the law of torts,[53] say "social torts," in which the public would collectively sue for "public" harm. A "criminal" action would be defined fundamentally not by the nature of the action[54] but by the inability of a person to compensate for the "harm" that he caused. Thus an action would be "criminal" precisely because it results in uncompensated "harm" to others. Criminal law would cover all such actions, while tort law would cover all other (civil) actions.

As a practical example of the fundamental changes that would be wrought, consider the antitrust field. Inspired in part by the economist's classic demonstration that monopolies distort the allocation of resources and reduce economic welfare, the United States has outlawed conspiracies

[52] "Violations," however, can only be punished by prison terms as long as fifteen days or fines no larger than $250. Since these are maximum punishments, the actual ones imposed by the courts can, and often are, considerably less. Note, too, that the courts can punish by imprisonment, by fine, or by *both* (*Laws of New York*, 1965, chap. 1030, Art. 60).

[53] "The cardinal principle of damages in Anglo-American law [of torts] is that of *compensation* for the injury caused to plaintiff by defendant's breach of duty" (Harper and James, 1956, p. 1299).

[54] Of course, many traditional criminal actions like murder or rape would still usually be criminal under this approach too.

and other constraints of trade. In practice, defendants are often simply required to cease the objectionable activity, although sometimes they are also fined, become subject to damage suits, or are jailed.

If compensation were stressed, the main purpose of legal proceedings would be to levy fines equal to[55] the harm inflicted on society by constraints of trade. There would be no point to cease and desist orders, imprisonment, ridicule, or dissolution of companies. If the economist's theory about monopoly is correct, and if optimal fines were levied, firms would automatically cease any constraints of trade, because the gain to them would be less than the harm they cause and thus less than the fines expected. On the other hand, if Schumpeter and other critics are correct, and certain constraints of trade raise the level of economic welfare, fines could fully compensate society for the harm done, and yet some constraints would not cease, because the gain to participants would exceed the harm to others.[56]

One unexpected advantage, therefore, from stressing compensation and fines rather than punishment and deterrence is that the validity of the classical position need not be judged a priori. If valid, compensating fines would discourage all constraints of trade and would achieve the classical aims. If not, such fines would permit the socially desirable constraints to continue and, at the same time, would compensate society for the harm done.

Of course, as participants in triple-damage suits are well aware, the harm done is not easily measured, and serious mistakes would be inevitable. However, it is also extremely difficult to measure the harm in many civil suits,[57] yet these continue to function, probably reasonably well on the whole. Moreover, as experience accumulated, the margin of error would decline, and rules of thumb would develop. Finally, one must realize that difficult judgments are also required by the present antitrust policy, such as deciding that certain industries are "workably" competitive or that certain mergers reduce competition. An emphasis on fines and compensation would at least help avoid irrelevant issues by focusing attention on the information most needed for intelligent social policy.

[55] Actually, fines should exceed the harm done if the probability of conviction were less than unity. The possibility of avoiding conviction is the intellectual justification for punitive, such as triple, damages against those convicted.

[56] The classical view is that $D'(M)$ always is greater than zero, where M measures the different constraints of trade and D' measures the marginal damage; the critic's view is that for some M, $D'(M) < 0$. It has been shown above that if D' always is greater than zero, compensating fines would discourage all offenses, in this case constraints of trade, while if D' sometimes is less than zero, some offenses would remain (unless $C'[M]$, the marginal cost of detecting and convicting offenders, were sufficiently large relative to D').

[57] Harper and James said, "Sometimes [compensation] can be accomplished with a fair degree of accuracy. But obviously it cannot be done in anything but a figurative and essentially speculative way for many of the consequences of personal injury. Yet it is the aim of the law to attain at least a rough correspondence between the amount awarded as damages and the extent of the suffering" (1956, p. 1301).

VI. Private Expenditures against Crime

A variety of private as well as public actions also attempt to reduce the number and incidence of crimes: guards, doormen, and accountants are employed, locks and alarms installed, insurance coverage extended, parks and neighborhoods avoided, taxis used in place of walking or subways, and so on. Table 1 lists close to $2 billion of such expenditures in 1965, and this undoubtedly is a gross underestimate of the total. The need for private action is especially great in highly interdependent modern economies, where frequently a person must trust his resources, including his person, to the "care" of employees, employers, customers, or sellers.

If each person tries to minimize his expected loss in income from crimes, optimal private decisions can be easily derived from the previous discussion of optimal public ones. For each person there is a loss function similar to that given by equation (18):

$$L_j = H_j(O_j) + C_j(p_j, O_j, C, C_k) + b_j p_j f_j O_j. \tag{31}$$

The term H_j represents the harm to j from the O_j offenses committed against j, while C_j represents his cost of achieving a probability of conviction of p_j for offenses committed against him. Note that C_j not only is positively related to O_j but also is negatively related to C, public expenditures on crime, and to C_k, the set of private expenditures by other persons.[58]

The term $b_j p_j f_j O_j$ measures the expected[59] loss to j from punishment of offenders committing any of the O_j. Whereas most punishments result in a net loss to society as a whole, they often produce a gain for the actual victims. For example, punishment by fines given to the actual victims is just a transfer payment for society but is a clear gain to victims; similarly, punishment by imprisonment is a net loss to society but is a negligible loss to victims, since they usually pay a negligible part of imprisonment costs. This is why b_j is often less than or equal to zero, at the same time that b, the coefficient of social loss, is greater than or equal to zero.

Since b_j and f_j are determined primarily by public policy on punishments, the main decision variable directly controlled by j is p_j. If he chooses a p_j that minimizes L_j, the optimality condition analogous to

[58] An increase in C_k—O_j and C held constant—presumably helps solve offenses against j, because more of those against k would be solved.

[59] The expected private loss, unlike the expected social loss, is apt to have considerable variance because of the small number of independent offenses committed against any single person. If j were not risk neutral, therefore, L would have to be modified to include a term that depended on the distribution of $b_j p_j f_j O_j$.

equation (22) is

$$H_j' + C_j' + C_{jp_j}\frac{\partial p_j}{\partial O_j} = -b_j p_j f_j\left(1 - \frac{1}{\varepsilon_{jp_j}}\right).^{60} \tag{32}$$

The elasticity ε_{jp_j} measures the effect of a change in p_j on the number of offenses committed against j. If $b_j < 0$, and if the left-hand side of equation (32), the marginal cost of changing O_j, were greater than zero, then (32) implies that $\varepsilon_{jp_j} > 1$. Since offenders can substitute among victims, ε_{jp_j} is probably much larger than ε_p, the response of the total number of offenses to a change in the average probability, p. There is no inconsistency, therefore, between a requirement from the optimality condition given by (22) that $\varepsilon_p < 1$ and a requirement from (32) that $\varepsilon_{jp_j} > 1$.

VII. Some Applications

A. Optimal Benefits

Our analysis of crime is a generalization of the economist's analysis of external harm or diseconomies. Analytically, the generalization consists in introducing costs of apprehension and conviction, which make the probability of apprehension and conviction an important decision variable, and in treating punishment by imprisonment and other methods as well as by monetary payments. A crime is apparently not so different analytically from any other activity that produces external harm and when crimes are punishable by fines, the analytical differences virtually vanish.

Discussions of external economies or advantages are usually perfectly symmetrical to those of diseconomies, yet one searches in vain for analogues to the law of torts and criminality. Generally, compensation cannot be collected for the external advantages as opposed to harm caused, and no public officials comparable to policemen and district attorneys apprehend and "convict" benefactors rather than offenders. Of course, there is

[60] I have assumed that

$$\frac{\partial C}{\partial p_j} = \frac{\partial C_k}{\partial p_j} = 0,$$

in other words, that j is too "unimportant" to influence other expenditures. Although usually reasonable, this does suggest a modification to the optimality conditions given by eqs. (21) and (22). Since the effects of public expenditures depend on the level of private ones, and since the public is sufficiently "important" to influence private actions, eq. (22) has to be modified to

$$D' + C' + C_p\frac{\partial p}{\partial O} + \sum_{i=1}^{n}\frac{dC}{dC_i}\frac{dC_i}{dp}\frac{\partial p}{\partial O} = -bpf\left(1 + \frac{1}{\varepsilon_p}\right), \tag{22'}$$

and similarly for eq. (21). "The" probability p is, of course, a weighted average of the p_j. Eq. (22′) incorporates the presumption that an increase in public expenditures would be partially thwarted by an induced decrease in private ones.

public interest in benefactors: medals, prizes, titles, and other privileges have been awarded to military heroes, government officials, scientists, scholars, artists, and businessmen by public and private bodies. Among the most famous are Nobel Prizes, Lenin Prizes, the Congressional Medal of Honor, knighthood, and patent rights. But these are piecemeal efforts that touch a tiny fraction of the population and lack the guidance of any body of law that codifies and analyzes different kinds of advantages.

Possibly the explanation for this lacuna is that criminal and tort law developed at the time when external harm was more common than advantages, or possibly the latter have been difficult to measure and thus considered too prone to favoritism. In any case, it is clear that the asymmetry in the law does not result from any analytical asymmetry, for a formal analysis of advantages, benefits, and benefactors can be developed that is quite symmetrical to the analysis of damages, offenses, and offenders. A function $A(B)$, for example, can give the net social advantages from B benefits in the same way that $D(O)$ gives the net damages from O offenses. Likewise, $K(B, p_1)$ can give the cost of apprehending and rewarding benefactors, where p_1 is the probability of so doing, with K' and $K_p > 0$; $B(p_1, a, v)$ can give the supply of benefits, where a is the award per benefit and v represents other determinants, with $\partial B/\partial p_1$ and $\partial B/\partial a > 0$; and b_1 can be the fraction of a that is a net loss to society. Instead of a loss function showing the decrease in social income from offenses, there can be a profit function showing the increase in income from benefits:

$$\Pi = A(B) - K(B, p_1) - b_1 p_1 a B. \tag{33}$$

If Π is maximized by choosing appropriate values of p_1 and a, the optimality conditions analogous to equations (21) and (22) are

$$A' - K' = b_1 p_1 a \left(1 + \frac{1}{e_a}\right) \tag{34}$$

and

$$A' - K' - K_p \frac{\partial p_1}{\partial B} = b_1 p_1 a \left(1 + \frac{1}{e_p}\right), \tag{35}$$

where

$$e_a = \frac{\partial B}{\partial a} \frac{a}{B}$$

and

$$e_p = \frac{\partial B}{\partial p_1} \frac{p_1}{B}$$

are both greater than zero. The implications of these equations are related to and yet differ in some important respects from those discussed earlier for (21) and (22).

For example, if $b_1 > 0$, which means that a is not a pure transfer but costs society resources, clearly (34) and (35) imply that $e_p > e_a$, since both $K_p > 0$ and $\partial p_1/\partial B > 0$. This is analogous to the implication of (21) and

(22) that $\varepsilon_p > \varepsilon_f$, but, while the latter implies that, at the margin, offenders are risk *preferrers*, the former implies that, at the margin, benefactors are risk *avoiders*.[61] Thus, while the optimal values of p and f would be in a region where "crime does not pay"—in the sense that the marginal income of criminals would be less than that available to them in less risky legal activities—the optimal values of p_1 and a would be where "benefits do pay"—in the same sense that the marginal income of benefactors would exceed that available to them in less risky activities. In this sense it "pays" to do "good" and does not "pay" to do "bad."

As an illustration of the analysis, consider the problem of rewarding inventors for their inventions. The function $A(B)$ gives the total social value of B inventions, and A' gives the marginal value of an additional one. The function $K(B, p_1)$ gives the cost of finding and rewarding inventors; if a patent system is used, it measures the cost of a patent office, of preparing applications, and of the lawyers, judges, and others involved in patent litigation.[62] The elasticities e_p and e_a measure the response of inventors to changes in the probability and magnitude of awards, while b_1 measures the social cost of the method used to award inventors. With a patent system, the cost consists in a less extensive use of an invention than would otherwise occur, and in any monopoly power so created.

Equations (34) and (35) imply that with any system having $b_1 > 0$, the smaller the elasticities of response of inventors, the smaller should be the probability and magnitude of awards. (The value of a patent can be changed, for example, by changing its life.) This shows the relevance of the controversy between those who maintain that most inventions stem from a basic desire "to know" and those who maintain that most stem from the prospects of financial awards, especially today with the emphasis on systematic investment in research and development. The former quite consistently usually advocate a weak patent system, while the latter equally consistently advocate its strengthening.

[61] The relation $e_p > e_a$ holds if, and only if,

$$\frac{\partial EU}{\partial p_1}\frac{p_1}{U} > \frac{\partial EU}{\partial a}\frac{a}{U}, \tag{1'}$$

where

$$EU = p_1 U(Y + a) + (1 - p_1)U(Y) \tag{2'}$$

(see the discussion on pp. 177–78). By differentiating eq. (2'), one can write (1') as

$$p_1[U(Y + a) - U(Y)] > p_1 a U'(Y + a), \tag{3'}$$

or

$$\frac{U(Y + a) - U(Y)}{a} > U'(Y + a). \tag{4'}$$

But (4') holds if everywhere $U'' < 0$ and does not hold if everywhere $U'' \geq 0$, which was to be proved.

[62] These costs are not entirely trivial: for example, in 1966 the U.S. Patent Office alone spent $34 million (see Bureau of the Budget, 1967), and much more was probably spent in preparing applications and in litigation.

Even if A', the marginal value of an invention, were "sizeable," the optimal decision would be to abolish property rights in an invention, that is, to set $p_1 = 0$, if b_1 and K[63] were sufficiently large and/or the elasticities e_p and e_a sufficiently small. Indeed, practically all arguments to eliminate or greatly alter the patent system have been based either on its alleged costliness, large K or b_1, or lack of effectiveness, low e_p or e_a (see, for example, Plant, 1934, or Arrow, 1962).

If a patent system were replaced by a system of cash prizes, the elasticities of response would become irrelevant for the determination of optimal policies, because b_1 would then be approximately zero.[64] A system of prizes would, moreover, have many of the same other advantages that fines have in punishing offenders (see the discussion in Sec. V). One significant advantage of a patent system, however, is that it automatically "meters" A', that is, provides an award that is automatically positively related to A', while a system of prizes (or of fines and imprisonment) has to estimate A' (or D') independently and often somewhat arbitrarily.

B. The Effectiveness of Public Policy

The anticipation of conviction and punishment reduces the loss from offenses and thus increases social welfare by discouraging some offenders. What determines the increase in welfare, that is "effectiveness," of public efforts to discourage offenses? The model developed in Section III can be used to answer this question if social welfare is measured by income and if "effectiveness" is defined as a ratio of the maximum feasible increase in income to the increase if all offenses causing net damages were abolished by fiat. The maximum feasible increase is achieved by choosing optimal values of the probability of apprehension and conviction, p, and the size of punishments, f (assuming that the coefficient of social loss from punishment, b, is given).[65]

[63] Presumably one reason patents are not permitted on basic research is the difficulty (that is, cost) of discovering the ownership of new concepts and theorems.

[64] The right side of both (34) and (35) would vanish, and the optimality conditions would be

$$A' - K' = 0 \tag{34'}$$

and

$$A' - K' - K_p \frac{\partial p_1}{\partial B} = 0. \tag{35'}$$

Since these equations are not satisfied by any finite values of p_1 and a, there is a difficulty in allocating the incentives between p_1 and a (see the similar discussion for fines in Sec. V).

[65] In symbols, effectiveness is defined as

$$E = \frac{D(O_1) - [D(\hat{O}) + C(\hat{p}, \hat{O}) + b\hat{p}\hat{f}\hat{O}]}{D(O_1) - D(O_2)},$$

where $\hat{p}$, $\hat{f}$, and $\hat{O}$ are optimal values, O_1 offenses would occur if $p = f = 0$, and O_2 is the value of O that minimizes D.

Effectiveness so defined can vary between zero and unity and depends essentially on two behavioral relations: the costs of apprehension and conviction and the elasticities of response of offenses to changes in p and f. The smaller these costs or the greater these elasticities, the smaller the cost of achieving any given reduction in offenses and thus the greater the effectiveness. The elasticities may well differ considerably among different kinds of offenses. For example, crimes of passion, like murder or rape, or crimes of youth, like auto theft, are often said to be less responsive to changes in p and f than are more calculating crimes by adults, like embezzlement, antitrust violation, or bank robbery. The elasticities estimated by Smigel (1965) and Ehrlich (1967) for seven major felonies do differ considerably but are not clearly smaller for murder, rape, auto theft, and assault than for robbery, burglary, and larceny.[66]

Probably effectiveness differs among offenses more because of differences in the costs of apprehension and conviction than in the elasticities of response. An important determinant of these costs, and one that varies greatly, is the time between commission and detection of an offense.[67] For the earlier an offense is detected, the earlier the police can be brought in and the more likely that the victim is able personally to identify the offender. This suggests that effectiveness is greater for robbery than for a related felony like burglary, or for minimum-wage and fair-employment legislation than for other white-collar legislation like antitrust and public-utility regulation.[68]

C. A Theory of Collusion

The theory developed in this essay can be applied to any effort to preclude certain kinds of behavior, regardless of whether the behavior is "unlawful." As an example, consider efforts by competing firms to collude in order to obtain monopoly profits. Economists lack a satisfactory theory of the determinants of price and output policies by firms in an industry, a theory that could predict under what conditions perfectly competitive, monopolistic, or various intermediate kinds of behavior would emerge. One by-product of our approach to crime and punishment is a theory of collusion that appears to fill a good part of this lacuna.[69]

[66] A theoretical argument that also casts doubt on the assertion that less "calculating" offenders are less responsive to changes in p and f can be found in Becker (1962).

[67] A study of crimes in parts of Los Angeles during January, 1966, found that "more than half the arrests were made within 8 hours of the crime, and almost two-thirds were made within the first week" (President's Commission 1967*e*, p. 8).

[68] Evidence relating to the effectiveness of actual, which are not necessarily optimal, penalties for these white-collar crimes can be found in Stigler (1962, 1966), Landes (1966), and Johnson (1967).

[69] Jacob Mincer first suggested this application to me.

The gain to firms from colluding is positively related to the elasticity of their marginal cost curves and is inversely related to the elasticity of their collective demand curve. A firm that violates a collusive arrangement by pricing below or producing more than is specified can be said to commit an "offense" against the collusion. The resulting harm to the collusion would depend on the number of violations and on the elasticities of demand and marginal cost curves, since the gain from colluding depends on these elasticities.

If violations could be eliminated without cost, the optimal solution would obviously be to eliminate all of them and to engage in pure monopoly pricing. In general, however, as with other kinds of offenses, there are two costs of eliminating violations. There is first of all the cost of discovering violations and of "apprehending" violators. This cost is greater the greater the desired probability of detection and the greater the number of violations. Other things the same, the latter is usually positively related to the number of firms in an industry, which partly explains why economists typically relate monopoly power to concentration. The cost of achieving a given probability of detection also depends on the number of firms, on the number of customers, on the stability of customer buying patterns, and on government policies toward collusive arrangements (see Stigler, 1964).

Second, there is the cost to the collusion of punishing violators. The most favorable situation is one in which fines could be levied against violators and collected by the collusion. If fines and other legal recourse are ruled out, methods like predatory price-cutting or violence have to be used, and they hurt the collusion as well as violators.

Firms in a collusion are assumed to choose probabilities of detection, punishments to violators, and prices and outputs that minimize their loss from violations, which would at the same time maximize their gain from colluding. Optimal prices and outputs would be closer to the competitive position the more elastic demand curves were, the greater the number of sellers and buyers, the less transferable punishments were, and the more hostile to collusion governments were. Note that misallocation of resources could not be measured simply by the deviation of actual from competitive outputs but would depend also on the cost of enforcing collusions. Note further, and more importantly, that this theory, unlike most theories of pricing, provides for continuous variation, from purely competitive through intermediate situations to purely monopolistic pricing. These situations differ primarily because of differences in the "optimal" number of violations, which in turn are related to differences in the elasticities, concentrations, legislation, etc., already mentioned.

These ideas appear to be helpful in understanding the relative success of collusions in illegal industries themselves! Just as firms in legal industries have an incentive to collude to raise prices and profits, so too do

firms producing illegal products, such as narcotics, gambling, prostitution, and abortion. The "syndicate" is an example of a presumably highly successful collusion that covers several illegal products.[70] In a country like the United States that prohibits collusions, those in illegal industries would seem to have an advantage, because force and other illegal methods could be used against violators without the latter having much legal recourse. On the other hand, in countries like prewar Germany that legalized collusions, those in legal industries would have an advantage, because violators could often be legally prosecuted. One would predict, therefore, from this consideration alone, relatively more successful collusions in illegal industries in the United States, and in legal ones in prewar Germany.

VIII. Summary and Concluding Remarks

This essay uses economic analysis to develop optimal public and private policies to combat illegal behavior. The public's decision variables are its expenditures on police, courts, etc., which help determine the probability (p) that an offense is discovered and the offender apprehended and convicted, the size of the punishment for those convicted (f), and the form of the punishment: imprisonment, probation, fine, etc. Optimal values of these variables can be chosen subject to, among other things, the constraints imposed by three behavioral relations. One shows the damages caused by a given number of illegal actions, called offenses (O), another the cost of achieving a given p, and the third the effect of changes in p and f on O.

"Optimal" decisions are interpreted to mean decisions that minimize the social loss in income from offenses. This loss is the sum of damages, costs of apprehension and conviction, and costs of carrying out the punishments imposed, and can be minimized simultaneously with respect to p, f, and the form of f unless one or more of these variables is constrained by "outside" considerations. The optimality conditions derived from the minimization have numerous interesting implications that can be illustrated by a few examples.

If carrying out the punishment were costly, as it is with probation, imprisonment, or parole, the elasticity of response of offenses with respect to a change in p would generally, in equilibrium, have to exceed its response to a change in f. This implies, if entry into illegal activities can be explained by the same model of choice that economists use to explain entry into legal activities, that offenders are (at the margin) "risk preferrers." Consequently, illegal activities "would not pay" (at the margin)

[70] An interpretation of the syndicate along these lines is also found in Schilling (1967).

in the sense that the real income received would be less than what could be received in less risky legal activities. The conclusion that "crime would not pay" is an optimality condition and not an implication about the efficiency of the police or courts; indeed, it holds for any level of efficiency, as long as optimal values of p and f appropriate to each level are chosen.

If costs were the same, the optimal values of both p and f would be greater, the greater the damage caused by an offense. Therefore, offenses like murder and rape should be solved more frequently and punished more severely than milder offenses like auto theft and petty larceny. Evidence on actual probabilities and punishments in the United States is strongly consistent with this implication of the optimality analysis.

Fines have several advantages over other punishments: for example, they conserve resources, compensate society as well as punish offenders, and simplify the determination of optimal p's and f's. Not surprisingly, fines are the most common punishment and have grown in importance over time. Offenders who cannot pay fines have to be punished in other ways, but the optimality analysis implies that the monetary value to them of these punishments should generally be less than the fines.

Vengeance, deterrence, safety, rehabilitation, and compensation are perhaps the most important of the many desiderata proposed throughout history. Next to these, minimizing the social loss in income may seem narrow, bland, and even quaint. Unquestionably, the income criterion can be usefully generalized in several directions, and a few have already been suggested in the essay. Yet one should not lose sight of the fact that it is more general and powerful than it may seem and actually includes more dramatic desiderata as special cases. For example, if punishment were by an optimal fine, minimizing the loss in income would be equivalent to compensating "victims" fully and would eliminate the "alarm" that so worried Bentham; or it would be equivalent to deterring all offenses causing great damage if the cost of apprehending, convicting, and punishing these offenders were relatively small. Since the same could also be demonstrated for vengeance or rehabilitation, the moral should be clear: minimizing the loss in income is actually very general and thus is *more useful* than these catchy and dramatic but inflexible desiderata.

This essay concentrates almost entirely on determining optimal policies to combat illegal behavior and pays little attention to actual policies. The small amount of evidence on actual policies that I have examined certainly suggests a positive correspondence with optimal policies. For example, it is found for seven major felonies in the United States that more damaging ones are penalized more severely, that the elasticity of response of offenses to changes in p exceeds the response to f, and that both are usually less than unity, all as predicted by the optimality analysis. There are, however, some discrepancies too: for example, the actual tradeoff between imprisonment and fines in different statutes is frequently less, rather than the

predicted more, favorable to those imprisoned. Although many more studies of actual policies are needed, they are seriously hampered on the empirical side by grave limitations in the quantity and quality of data on offenses, convictions, costs, etc., and on the analytical side by the absence of a reliable theory of political decision-making.

Reasonable men will often differ on the amount of damages or benefits caused by different activities. To some, any wage rates set by competitive labor markets are permissible, while to others, rates below a certain minimum are violations of basic rights; to some, gambling, prostitution, and even abortion should be freely available to anyone willing to pay the market price, while to others, gambling is sinful and abortion is murder. These differences are basic to the development and implementation of public policy but have been excluded from my inquiry. I assume consensus on damages and benefits and simply try to work out rules for an optimal implementation of this consensus.

The main contribution of this essay, as I see it, is to demonstrate that optimal policies to combat illegal behavior are part of an optimal allocation of resources. Since economics has been developed to handle resource allocation, an "economic" framework becomes applicable to, and helps enrich, the analysis of illegal behavior. At the same time, certain unique aspects of the latter enrich economic analysis: some punishments, such as imprisonments, are necessarily non-monetary and are a cost to society as well as to offenders; the degree of uncertainty is a decision variable that enters both the revenue and cost functions; etc.

Lest the reader be repelled by the apparent novelty of an "economic" framework for illegal behavior, let him recall that two important contributors to criminology during the eighteenth and nineteenth centuries, Beccaria and Bentham, explicitly applied an economic calculus. Unfortunately, such an approach has lost favor during the last hundred years, and my efforts can be viewed as a resurrection, modernization, and thereby I hope improvement on these much earlier pioneering studies.

Mathematical Appendix

This Appendix derives the effects of changes in various parameters on the optimal values of p and f. It is assumed throughout that $b > 0$ and that equilibrium occurs where

$$\frac{\partial D}{\partial O} + \frac{\partial C}{\partial O} + \frac{\partial C}{\partial p}\frac{\partial p}{\partial O} = D' + C' + C_p \frac{\partial p}{\partial O} > 0;$$

the analysis could easily be extended to cover negative values of b and of this marginal cost term. The conclusion in the text (Sec. II) that $D'' + C'' > 0$ is relied on here. I take it to be a reasonable first approximation that the elasticities of O with respect to p or f are constant. At several places a sufficient

condition for the conclusions reached is that

$$C_{pO} = C_{Op} = \frac{\partial^2 C}{\partial p \partial O} = \frac{\partial^2 C}{\partial O \partial p}$$

is "small" relative to some other terms. This condition is utilized in the form of a strong assumption that $C_{pO} = 0$, although I cannot claim any supporting intuitive or other evidence.

The social loss in income from offenses has been defined as

$$L = D(O) + C(O, p) + bpfO. \tag{A1}$$

If b and p were fixed, the value of f that minimized L would be found from the necessary condition

$$\frac{\partial L}{\partial f} = 0 = (D' + C')\frac{\partial O}{\partial f} + bpf(1 - E_f)\frac{\partial O}{\partial f}, \tag{A2}$$

or

$$0 = D' + C' + bpf(1 - E_f), \tag{A3}$$

if

$$\frac{\partial O}{\partial f} = O_f \neq 0,$$

where

$$E_f = \frac{-\partial f}{\partial O}\frac{O}{f}.$$

The sufficient condition would be that $\partial^2 L/\partial f^2 > 0$; using $\partial L/\partial f = 0$ and E_f is constant, this condition becomes

$$\frac{\partial^2 L}{\partial f^2} = (D'' + C'')O_f^2 + bp(1 - E_f)O_f > 0, \tag{A4}$$

or

$$\Delta \equiv D'' + C'' + bp(1 - E_f)\frac{1}{O_f} > 0. \tag{A5}$$

Since $D' + C' > 0$, and b is not less than zero, equation (A3) implies that $E_f > 1$. Therefore Δ would be greater than zero, since we are assuming that $D'' + C'' > 0$; and $\hat{f}$, the value of f satisfying (A3), would minimize (locally) the loss L.

Suppose that D' is positively related to an exogenous variable α. The effect of a change in α on $\hat{f}$ can be found by differentiating equation (A3):

$$D'_\alpha + (D'' + C'')O_f\frac{d\hat{f}}{d\alpha} + bp(1 - E_f)\frac{d\hat{f}}{d\alpha} = 0,$$

or

$$\frac{d\hat{f}}{d\alpha} = \frac{-D'_\alpha(1/O_f)}{\Delta}. \tag{A6}$$

Since $\Delta > 0$, $O_f < 0$, and by assumption $D'_\alpha > 0$, then

$$\frac{d\hat{f}}{d\alpha} = \frac{+}{+} > 0. \tag{A7}$$

In a similar way it can be shown that, if C' is positively related to an exogenous variable β,

$$\frac{d\hat{f}}{d\beta} = \frac{-C'_\beta(1/O_f)}{\Delta} = \frac{+}{+} > 0. \tag{A8}$$

If b is positively related to γ, then

$$(D'' + C'')O_f \frac{d\hat{f}}{d\gamma} + bp(1 - E_f)\frac{d\hat{f}}{d\gamma} + pf(1 - E_f)b\gamma = 0,$$

or

$$\frac{d\hat{f}}{d\gamma} = \frac{-b_\gamma pf(1 - E_f)(1/O_f)}{\Delta}. \tag{A9}$$

Since $1 - E_f < 0$, and by assumption $b_\gamma > 0$,

$$\frac{d\hat{f}}{d\gamma} = \frac{-}{+} < 0. \tag{A10}$$

Note that since $1/E_f < 1$,

$$\frac{d(p\hat{f}O)}{d\gamma} < 0. \tag{A11}$$

If E_f is positively related to δ, then

$$\frac{d\hat{f}}{d\delta} = \frac{E_{f\delta}bpf(1/O_f)}{\Delta} = \frac{-}{+} < 0. \tag{A12}$$

Since the elasticity of O with respect to f equals

$$\epsilon_f = -O_f \frac{f}{O} = \frac{1}{E_f},$$

by (A12), a reduction in ϵ_f would reduce $\hat{f}$.

Suppose that p is related to the exogeneous variable r. Then the effect of a shift in r on $\hat{f}$ can be found from

$$(D'' + C'')O_f \frac{d\hat{f}}{dr} + (D'' + C'')O_p p_r + C_{pO}p_r + bp(1 - E_f)\frac{\partial \hat{f}}{\partial r} + bf(1 - E_f)p_r = 0,$$

or

$$\frac{d\hat{f}}{dr} = \frac{-(D'' + C'')O_p(1/O_f)p_r - bf(1 - E_f)p_r(1/O_f)}{\Delta}, \tag{A13}$$

since by assumption $C_{pO} = 0$. Since $O_p < 0$, and $(D'' + C'') > 0$,

$$\frac{d\hat{f}}{dr} = \frac{(-) + (-)}{+} = \frac{-}{+} < 0. \tag{A14}$$

If f rather than p were fixed, the value of p that minimizes L, $\hat{p}$, could be found from

$$\frac{\partial L}{\partial p} = \left[D' + C' + C_p \frac{1}{O_p} + bpf(1 - E_p)\right]O_p = 0, \tag{A15}$$

as long as

$$\frac{\partial^2 L}{\partial p^2} = \left[(D'' + C'')O_p + C'_p + C_{pp}\frac{1}{O_p} + C_{pO} + C_p \frac{\partial^2 p}{\partial O \partial p} + bf(1 - E_p)\right]O_p > 0. \tag{A16}$$

Since $C_p' = C_{pO} = 0$, (A16) would hold if

$$\Delta' \equiv D'' + C'' + C_{pp}\frac{1}{O_p^2} + C_p\frac{1}{O_p}\frac{\partial^2 p}{\partial O \partial p} + bf(1 - E_p)\frac{1}{O_p} > 0. \tag{A17}$$

It is suggested in Section II that C_{pp} is generally greater than zero. If, as assumed,

$$D' + C' + C_p\frac{1}{O_p} > 0,$$

equation (A15) implies that $E_p > 1$ and thus that

$$bf(1 - E_p)\frac{1}{O_p} > 0.$$

If E_p were constant, $\partial^2 p/\partial O \partial p$ would be negative,[71] and, therefore, $C_p(1/O_p)(\partial^2 p/\partial O \partial p)$ would be positive. Hence, none of the terms of (A17) are negative, and a value of p satisfying equation (A15) would be a local minimum.

The effects of changes in different parameters on $\hat{p}$ are similar to those already derived for $\hat{f}$ and can be written without comment:

$$\frac{d\hat{p}}{d\alpha} = \frac{-D_\alpha'(1/O_p)}{\Delta'} > 0, \tag{A18}$$

$$\frac{d\hat{p}}{d\beta} = \frac{-C_\beta'(1/O_p)}{\Delta'} > 0, \tag{A19}$$

and

$$\frac{d\hat{p}}{d\gamma} = \frac{-b_\gamma pf(1 - E_p)(1/O_p)}{\Delta'} < 0. \tag{A20}$$

If E_p is positively related to δ',

$$\frac{d\hat{p}}{d\delta'} = \frac{E_{p\delta'}bpf(1/O_p)}{\Delta'} < 0. \tag{A21}$$

If C_p were positively related to the parameter s, the effect of a change in s on $\hat{p}$ would equal

$$\frac{d\hat{p}}{ds} = \frac{-C_{ps}(1/O_p^2)}{\Delta'} < 0. \tag{A22}$$

If f were related to the exogenous parameter t, the effect of a change in t on $\hat{p}$ would be given by

$$\frac{d\hat{p}}{dt} = \frac{-(D'' + C'')O_p f_t(1/O_p) - bf(1 - E_p)f_t(1/O_p) - C_p(\partial^2 p/\partial O \partial f)f_t(1/O_p)}{\Delta'} < 0 \tag{A23}$$

(with $C_{pO} = 0$), since all the terms in the numerator are negative.

If both p and f were subject to control, L would be minimized by choosing optimal values of both variables simultaneously. These would be given by the

[71] If E_p and E_f are constants, $O = kp^{-a}f^{-b}$, where $a = 1/E_p$ and $b = 1/E_f$. Then

$$\frac{\partial p}{\partial O} = -\frac{1}{ka}p^{a+1}f^b,$$

and

$$\frac{\partial^2 p}{\partial O \partial p} = \frac{-(a+1)}{ka}p^a f^b < 0.$$

solutions to the two first-order conditions, equations (A2) and (A15), assuming that certain more general second-order conditions were satisfied. The effects of changes in various parameters on these optimal values can be found by differentiating both first-order conditions and incorporating the restrictions of the second-order conditions.

The values of p and f satisfying (A2) and (A15), $\hat{p}$ and $\hat{f}$, minimize L if

$$L_{pp} > 0,\ L_{ff} > 0, \tag{A24}$$

and

$$L_{pp}L_{ff} > L_{fp}^2 = L_{pf}^2. \tag{A25}$$

But $L_{pp} = O_p^2\Delta'$, and $L_{ff} = O_f^2\Delta$, and since both Δ' and Δ have been shown to be greater than zero, (A24) is proved already, and only (A25) remains. By differentiating L_f with respect to p and utilizing the first-order condition that $L_f = 0$, one has

$$L_{fp} = O_fO_p[D'' + C'' + bf(1 - E_f)p_O] = O_fO_p\Sigma, \tag{A26}$$

where Σ equals the term in brackets. Clearly $\Sigma > 0$.

By substitution, (A25) becomes

$$\Delta\Delta' > \Sigma^2, \tag{A27}$$

and (A27) holds if Δ and Δ' are both greater than Σ. $\Delta > \Sigma$ means that

$$D'' + C'' + bp(1 - E_f)f_O > D'' + C'' + bf(1 - E_f)p_O, \tag{A28}$$

or

$$\frac{bfp}{O}(1 - E_f)E_f < \frac{bpf}{O}(1 - E_f)E_p. \tag{A29}$$

Since $1 - E_f < 0$, (A29) implies that

$$E_f > E_p, \tag{A30}$$

which necessarily holds given the assumption that $b > 0$; prove this by combining the two first-order conditions (A2) and (A15). $\Delta' > \Sigma$ means that

$$D'' + C'' + C_{pp}p_O^2 + C_pp_Op_{Op} + bf(1 - E_p)p_O > D'' + C'' + bf(1 - E_f)p_O. \tag{A31}$$

Since $C_{pp}p_O^2 > 0$, and $p_O < 0$, this necessarily holds if

$$C_pp_{Op} + bpf(1 - E_p) < bpf(1 - E_f). \tag{A32}$$

By eliminating $D' + C'$ from the first-order conditions (A2) and (A15) and by combining terms, one has

$$C_pp_O - bpf(E_p - E_f) = 0. \tag{A33}$$

By combining (A32) and (A33), one gets the condition

$$C_pp_{Op} < C_pp_O, \tag{A34}$$

or

$$E_{p_O,p} = \frac{p}{p_O}\frac{\partial p_O}{\partial p} > 1. \tag{A35}$$

It can be shown that

$$E_{p_O,p} = 1 + \frac{1}{E_p} > 1, \tag{A36}$$

and, therefore, (A35) is proven.

It has now been proved that the values of p and f that satisfy the first-order conditions (A2) and (A15) do indeed minimize (locally) L. Changes in different parameters change these optimal values, and the direction and magnitude can be found from the two linear equations

$$O_f\Delta\frac{\partial\tilde{f}}{\partial z} + O_p\Sigma\frac{\partial\tilde{p}}{\partial z} = C_1$$

and (A37)

$$O_f\Sigma\frac{\partial\tilde{f}}{\partial z} + O_p\Delta'\frac{\partial\tilde{p}}{\partial z} = C_2.$$

By Cramer's rule,

$$\frac{\partial\tilde{f}}{\partial z} = \frac{C_1O_p\Delta' - C_2O_p\Sigma}{O_pO_f(\Delta\Delta' - \Sigma^2)} = \frac{O_p(C_1\Delta' - C_2\Sigma)}{+}, \tag{A38}$$

$$\frac{\partial\tilde{p}}{\partial z} = \frac{C_2O_f\Delta - C_1O_f\Sigma}{O_pO_f(\Delta\Delta' - \Sigma^2)} = \frac{O_f(C_2\Delta - C_1\Sigma)}{+}, \tag{A39}$$

and the signs of both derivatives are the same as the signs of the numerators.

Consider the effect of a change in D' resulting from a change in the parameter α. It is apparent that $C_1 = C_2 = -D'_\alpha$, and by substitution

$$\frac{\partial\tilde{f}}{\partial\alpha} = \frac{-O_pD'_\alpha(\Delta' - \Sigma)}{+} = \frac{+}{+} > 0 \tag{A40}$$

and

$$\frac{\partial\tilde{p}}{\partial\alpha} = \frac{-O_pD'_\alpha(\Delta - \Sigma)}{+} = \frac{+}{+} > 0, \tag{A41}$$

since O_f and $O_p < 0$, $D'_\alpha > 0$, and Δ and $\Delta' > \Sigma$.

Similarly, if C' is changed by a change in β, $C_1 = C_2 = -C'_\beta$,

$$\frac{\partial\tilde{f}}{\partial\beta} = \frac{-O_pC'_\beta(\Delta' - \Sigma)}{+} = \frac{+}{+} > 0, \tag{A42}$$

and

$$\frac{\partial\tilde{p}}{\partial\beta} = \frac{-O_fC'_\beta(\Delta - \Sigma)}{+} = \frac{+}{+} > 0. \tag{A43}$$

If E_f is changed by a change in δ, $C_1 = E_{f\delta}bpf$, $C_2 = 0$,

$$\frac{\partial\tilde{f}}{\partial\delta} = \frac{O_pE_fbpf\Delta'}{+} = \frac{-}{+} < 0, \tag{A44}$$

and

$$\frac{\partial\tilde{p}}{\partial\delta} = \frac{-O_fE_fbpf\Sigma}{+} = \frac{+}{+} > 0. \tag{A45}$$

Similarly, if E_p is changed by a change in δ', $C_1 = 0$, $C_2 = E_{p\delta'}bpf$,

$$\frac{\partial\tilde{f}}{\partial\delta'} = -\frac{O_pE_{p\delta'}bpf\Sigma}{+} = \frac{+}{+} > 0, \tag{A46}$$

and

$$\frac{\partial\tilde{p}}{\partial\delta'} = \frac{O_fE_{p\delta'}bpf\Delta}{+} = \frac{-}{+} < 0. \tag{A47}$$

If b is changed by a change in γ, $C_1 = -b_\gamma pf(1 - E_f)$, $C_2 = -b_\gamma pf(1 - E_p)$, and

$$\frac{\partial \tilde{f}}{\partial \gamma} = \frac{-O_p b_\gamma pf[(1 - E_f)\Delta' - (1 - E_p)\Sigma]}{+} = \frac{-}{+} < 0, \tag{A48}$$

since $E_f > E_p > 1$ and $\Delta' > \Sigma$; also,

$$\frac{\partial \tilde{p}}{\partial \gamma} = \frac{-O_f b_\gamma pf[(1 - E_p)\Delta - (1 - E_f)\Sigma]}{+} = \frac{+}{+} > 0, \tag{A49}$$

for it can be shown that $(1 - E_p)\Delta > (1 - E_f)\Sigma$.[72] Note that when f is held constant the optimal value of p is decreased, not increased, by an increase in γ.

If C_p is changed by a change in s, $C_2 = -p_O C_{ps}$, $C_1 = 0$,

$$\frac{\partial \tilde{f}}{\partial s} = \frac{O_p p_O C_{ps}\Sigma}{+} = \frac{C_{ps}\Sigma}{+} = \frac{+}{+} > 0, \tag{A50}$$

and

$$\frac{\partial \tilde{p}}{\partial s} = \frac{-O_f p_O C_{ps}\Delta}{+} = \frac{-}{+} < 0. \tag{A51}$$

[72] The term $(1 - E_p)\Delta$ would be greater than $(1 - E_f)\Sigma$ if

$$(D'' + C'')(1 - E_p) + bp(1 - E_f)(1 - E_p)f_O > (D'' + C'')(1 - E_f) + bf(1 - E_f)^2 p_O,$$

or

$$(D'' + C'')(E_f - E_p) > -\frac{bpf}{O}(1 - E_f)\left[(1 - E_p)\frac{f_o O}{f} - (1 - E_f)\frac{p_o O}{p}\right],$$

$$(D'' + C'')(E_f - E_p) > -\frac{bpf}{O}(1 - E_f)[(1 - E_p)(E_f) - (1 - E_f)E_p],$$

$$(D'' + C'')(E_f - E_p) > -\frac{bpf}{O}(1 - E_f)(E_f - E_p).$$

Since the left-hand side is greater than zero, and the right-hand side is less than zero, the inequality must hold.

References

Arrow, Kenneth J. "Economic Welfare and Allocation of Resources for Invention," in National Bureau Committee for Economic Research. *The Rate and Direction of Inventive Activity: Economic and Social Factors.* Princeton, N.J.: Princeton Univ. Press (for the Nat. Bureau of Econ. Res.), 1962.

Becker, Gary S. "Irrational Behavior and Economic Theory," *J.P.E.*, Vol. LXX (February, 1962).

———. "A Theory of the Allocation of Time," *Econ. J.*, Vol. LXXV (September, 1965).

Bentham, Jeremy. *Theory of Legislation.* New York: Harcourt Brace Co., 1931.

Bureau of the Budget. *The Budget of United States Government, 1968, Appendix.* Washington: U.S. Government Printing Office, 1967.

Bureau of Prisons. *Prisoners Released from State and Federal Institutions.* ("National Prisoner Statistics.") Washington: U.S. Dept. of Justice, 1960.
———. *Characteristics of State Prisoners, 1960.* ("National Prisoner Statistics.") U.S. Dept. of Justice, n.d.
———. *Federal Prisons, 1960.* Washington: U.S. Dept. of Justice, 1961.
Cagan, Phillip. *Determinants and Effects of Changes in the Stock of Money, 1875–1960.* New York: Columbia Univ. Press (for the Nat. Bureau of Econ. Res.), 1965.
"Criminal Safeguards and the Punitive Damages Defendant," *Univ. Chicago Law Rev.*, Vol. XXXIV (Winter, 1967).
Ehrlich, Isaac. "The Supply of Illegitimate Activities." Unpublished manuscript, Columbia Univ., New York, 1967.
Federal Bureau of Investigation. *Uniform Crime Reports for the United States.* Washington: U.S. Dept. of Justice, 1960.
———. *Ibid.*, 1961.
Harper, F. V., and James, F. *The Law of Torts*, Vol. II. Boston: Little-Brown & Co., 1956.
Johnson, Thomas. "The Effects of the Minimum Wage Law." Unpublished Ph.D. dissertation, Columbia Univ., New York, 1967.
Kleinman, E. "The Choice between Two 'Bads'—Some Economic Aspects of Criminal Sentencing." Unpublished manuscript, Hebrew Univ., Jerusalem, 1967.
Landes, William. "The Effect of State Fair Employment Legislation on the Economic Position of Nonwhite Males." Unpublished Ph.D. dissertation, Columbia Univ., New York, 1966.
Laws of New York, Vol. II (1965).
Marshall, Alfred. *Principles of Economics.* 8th ed. New York: Macmillan Co., 1961.
Plant, A. "The Economic Theory concerning Patents for Inventions," *Economica*, Vol. I (February, 1934).
Pollock, F., and Maitland, F. W. *The History of English Law.* Vol. II. 2d ed. Cambridge: Cambridge Univ. Press, 1952.
President's Commission on Law Enforcement and Administration of Justice. *The Challenge of Crime in a Free Society.* Washington: U.S. Government Printing Office, 1967(*a*).
———. *Corrections.* ("Task Force Reports.") Washington: U.S. Government Printing Office, 1967(*b*).
———. *The Courts.* ("Task Force Reports.") Washington: U.S. Government Printing Office, 1967(*c*).
———. *Crime and Its Impact—an Assessment.* ("Task Force Reports.") Washington: U.S. Government Printing Office, 1967(*d*).
———. *Science and Technology.* ("Task Force Reports.") Washington: U.S. Government Printing Office, 1967(*e*).
Radzinowicz, L. *A History of English Criminal Law and Its Administration from 1750.* Vol. I. London: Stevens & Sons, 1948.
Schilling, T. C. "Economic Analysis of Organized Crime," in President's Commission on Law Enforcement and Administration of Justice. *Organized Crime.* ("Task Force Reports.") Washington: U.S. Government Printing Office, 1967.
Shawness, Lord. "Crime *Does* Pay because We Do Not Back Up the Police," *New York Times Magazine*, June 13. 1965.
Smigel, Arleen. "Crime and Punishment: An Economic Analysis." Unpublished M.A. thesis, Columbia Univ., New York, 1965.

Stigler, George J. "What Can Regulators Regulate? The Case of Electricity," *J. Law and Econ.*, Vol. V (October, 1962).
———. "A Theory of Oligopoly," *J.P.E.*, Vol. LXXII (February, 1964).
———. "The Economic Effects of the Antitrust Laws," *J. Law and Econ.*, Vol. IX (October, 1966).
Sutherland, E. H. *Principles of Criminology*, 6th ed. Philadelphia: J. B. Lippincott Co., 1960.

Exhibit 12 to the Cross-Examination of Professor Tadelis

1983

Optimal Sanctions for Antitrust Violations

William M. Landes

Optimal Sanctions for Antitrust Violations

William M. Landes†

Although economic analysis of the common law, crime, and legal decision making are relatively recent areas of research in the field of law and economics, economic analysis of antitrust, particularly the analysis of business practices described in antitrust cases, has been widespread and uncontroversial for many years. What has received less attention is the use of economics to examine antitrust enforcement itself.[1] This involves analyzing, for example, what is an antitrust injury, the appropriate sanctions for such an injury, the choice between public and private enforcement of antitrust laws and related questions on standing to sue, and the relevance of the antitrust victim's conduct to his ability to recover damages. In this paper I apply economics to some of the above issues.

Economic analysis of antitrust enforcement builds on the pioneering papers of Gary Becker and Ronald Coase.[2] Becker's paper was the first formal analysis of optimal penalties and probabilities of apprehension and conviction for criminal offenses. He showed

---

† Clifton R. Musser Professor of Economics, University of Chicago Law School. I developed the basic analysis of this paper and some of the applications in my antitrust lectures at the Law School in a course that I have co-taught regularly with Frank Easterbrook and Richard Posner. Their contributions to my lectures and to this paper have been substantial. I would also like to thank Dennis Carlton, Elisabeth Landes, William Lynk, Mitchell Polinsky, Andrew Rosenfield, and Robert Sherwin for helpful comments on an earlier draft of this paper.

[1] There have been a number of recent contributions, however. These include K. ELZINGA & W. BREIT, THE ANTITRUST PENALTIES: A STUDY IN LAW AND ECONOMICS (1976); R. POSNER & F. EASTERBROOK, ANTITRUST: CASES, ECONOMIC NOTES AND OTHER MATERIALS 545-72 (2d ed. 1981); W. SCHWARTZ, PRIVATE ENFORCEMENT OF THE ANTITRUST LAWS (1981); Block, Nold & Sidak, *The Deterrent Effect of Antitrust Enforcement*, 89 J. POL. ECON. 429 (1981); Easterbrook, *Predatory Strategies and Counterstrategies*, 48 U. CHI. L. REV. 263 (1981); Easterbrook & Fischel, *Antitrust Suits by Targets of Tender Offers*, 80 MICH. L. REV. 1155 (1982); Easterbrook, Landes & Posner, *Contribution Among Antitrust Defendants: A Legal and Economic Analysis*, 23 J.L. & ECON. 331 (1980); Landes & Posner, *Should Indirect Purchasers Have Standing To Sue Under the Antitrust Laws? An Economic Analysis of the Rule of* Illinois Brick, 46 U. CHI. L. REV. 602 (1979); Page, *Antitrust Damages and Economic Efficiency: An Approach to Antitrust Injury*, 47 U. CHI. L. REV. 467 (1980); and Note, *Rethinking Antitrust Damages*, 33 STAN. L. REV. 329 (1981).

[2] Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169 (1968); Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960).

that when the costs of enforcement are positive, it is generally not optimal to reduce the number of violations to zero. More surprisingly, Becker also showed that even if enforcement costs are zero, it is still not desirable to deter all violations because some offenses—where the gain to the offender exceeds the harm to the victim—are efficient. The concept of an efficient violation is the key to determining the optimal antitrust penalty. Although Coase's paper is usually cited for the proposition (known as "Coase's theorem") that in the absence of transaction costs alternative liability rules do not affect resource allocation, it marked the beginning of systematic study by economists and academic lawyers of the effects of alternative liability rules on resource allocation in high transaction cost settings. Because an antitrust violation is equivalent to an intentional tort, one can analyze many antitrust enforcement issues by applying the economic analysis of tort liability.

This paper is divided into three parts. The first part presents an economic analysis of optimal antitrust penalties. The second applies the analysis to several topics in antitrust including joint ventures, the social cost of monopoly, cartels that face competition from fringe firms, and predatory pricing. The final section applies the economic analysis of intentional torts to antitrust enforcement.

I. The Theory of Optimal Sanctions

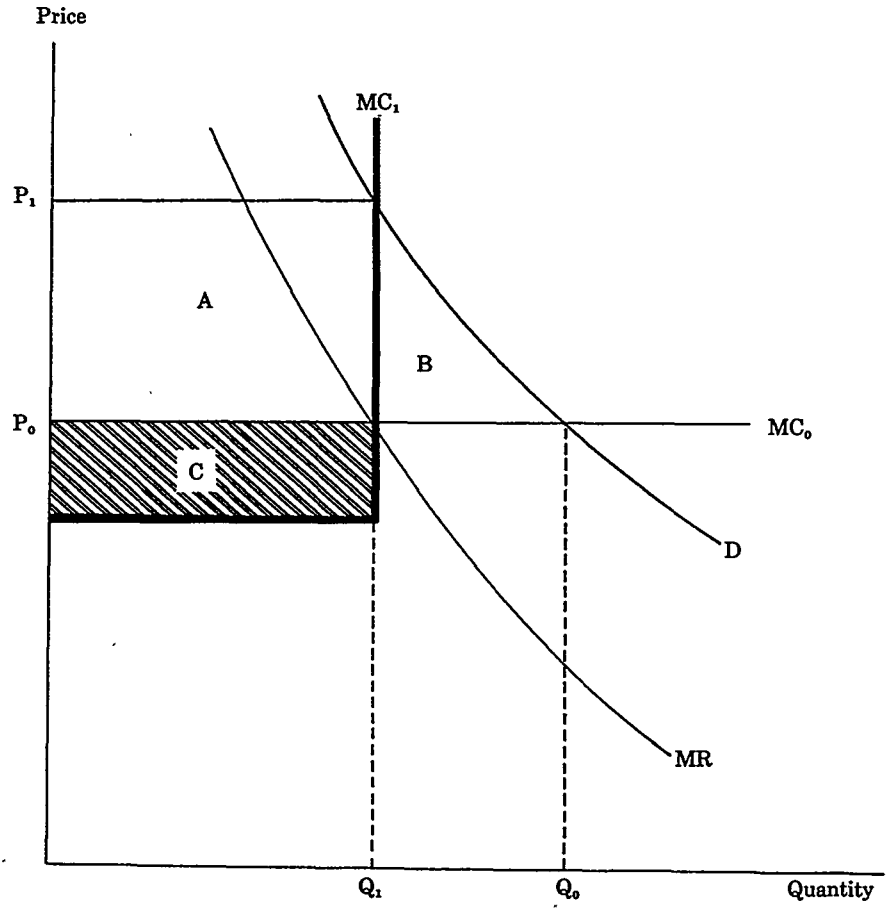
A. The Basic Model of Optimal Sanctions

I begin with a simple example that brings out the basic intuition of an an optimal sanction. In Figure 1, industry marginal cost, MC_0, is constant and equal to the supply curve under competition; the competitive output and price equal Q_0 and P_0; and a cartel would reduce output to Q_1 and raise price to P_1. I assume further that a cartel would impose a deadweight loss of \$50 (area B) and an aggregate overcharge of \$100 (area A). Total harm to consumers equals \$150, the sum of the aggregate overcharge and deadweight loss.

The standard economic rationale for making a cartel illegal is not that it charges too high a price or that it redistributes income from consumers to cartel members, but that it restricts output, causing a deadweight or efficiency loss (area B)—a loss to consumers without an offsetting gain to producers. To prevent this loss one can penalize cartel members by an amount sufficient to deter them from organizing a cartel in the first instance. To simplify the analysis of optimal penalties, I assume the following: all parties are risk neutral; enforcement costs, including legal fees, court costs,

and time costs, are zero; the penalty is a monetary fine; and, provisionally, the probability of apprehension and conviction is fixed and equal to one.[3] Given these assumptions, what is the optimal penalty?

FIGURE 1



One possible penalty is the social or deadweight loss of $50.

[3] I also assume that consumers do not take account of a possible damage recovery in deciding how many units of the good to purchase. If they did, this would complicate the analysis. Then, one would have to consider both consumer anticipations of a recovery and the response of the cartel, knowing that consumers adjust their purchases in light of the prospect of an antitrust recovery. This is considered briefly in Landes & Posner, *supra* note 1, at 606-08, and *infra* at notes 49-54 and accompanying text.

The argument might be that since enforcement costs are zero and the penalty is certain, an offender should be made to pay the deadweight or social cost of his offense. The difficulty with this rule is that despite the penalty, it still may be profitable to form the cartel. In our example, a $50 fine will be too low. Firms would not forgo cartel profits of $100 to avoid a $50 fine. Consider an analogy to theft. One would not deter a thief from stealing $100 cash by a penalty equal to the lower social cost of the theft, measured by the monetary equivalent of the services the victim forgoes by holding lower cash balances in response to the possibility of theft.

Alternatively, why not impose a fine many times greater than the social cost? A $10,000 fine surely would deter firms from forming a cartel to earn a $100 profit. Given our assumption that the cartel causes a $50 deadweight loss, a $10,000 fine or a $10 million fine would yield the correct outcome. In general, however, large fines will not yield the correct outcome.[4] A fine of $10,000 or even a fine of $151 in our example could be too large because deterrence alone is not the aim of penalties. The purpose of penalties, following Becker's model of crime and punishment, is to deter inefficient offenses, not efficient ones. Stated differently, the optimal level of offenses is generally greater than zero.

To explain, suppose the cartel is able to reduce production costs, but cost savings can be obtained only by restricting output.[5] Returning to Figure 1, let MC_1 denote the cartel's marginal cost. Notice that MC_1 is below MC_0 until the output Q_1, and at Q_1 the cartel's marginal cost curve becomes perfectly inelastic. Admittedly, the assumption of a fixed production capacity for the cartel at Q_1 is unrealistic. It has the advantage, however, of simplifying the explanation of an optimal fine. The aggregate cost saving from the cartel equals the shaded area C. If C is greater than $50, the cartel's offense is efficient: the offense produces greater cost savings than the deadweight loss it imposes by restricting output. For example, if C were $51, there is a net gain of $1 from the cartel.

---

[4] Fines of this magnitude raise two problems not considered in this paper. The first is the problem of marginal deterrence: if all fines are large and differences between them are small relative to differences in harm, offenders tend to commit the most harmful offenses. *See* Stigler, *The Optimum Enforcement of Laws*, 78 J. POL. ECON. 526, 527-28 (1970). The other is the possibility of legal error, which combined with large fines can deter socially valuable business behavior. For an application of legal error to antitrust enforcement, see Polinsky & Shavell, *Contribution and Claim Reduction Among Antitrust Defendants: An Economic Analysis*, 33 STAN. L. REV. 447 (1981).

[5] The question of how these savings could arise is discussed *infra* at notes 11-16 and accompanying text.

Because the cartel's total gain is $151 (equal to the $100 overcharge plus the $51 cost savings), a fine greater than $151 would deter its formation. This outcome would be inefficient, deterring a socially beneficial cartel. Under other circumstances, a fine less than $150 would be inefficient. For example, if C were $49 and the fine $148, the cartel would be formed ($149 profits minus $148 fine leaves the cartel with a net gain of $1), yet society would be worse off by $1 because the deadweight loss exceeds the cost savings made possible by the cartel.

The rule for determining the optimal fine or damage award is simple to state: the fine should equal the net harm to persons other than the offender.[6] In our example this harm is $150, the $100 overcharge plus the $50 deadweight loss. If we follow the net harm rule, the offense will only take place when the gain to the offender exceeds the net harm to others, and the cartel will be deterred when the gain is less than the harm.[7]

I note several additional points.

1. The imposition of a penalty in the example of Figure 1 does not depend on an actual increase in price above P_0. If it did, a cartel could avoid any antitrust liability by reducing output to Q_1 but maintaining price at P_0, for example, by rationing demand on a first-come-first-served basis.[8] Suppose the cartel's cost savings (area C) equaled $10. Without a penalty, the cartel could reap profits equal to its $10 cost savings, yet the arrangement would be inefficient; it would produce a net loss of $40, equal to a $50 deadweight loss minus a $10 cost savings. In contrast, the net harm rule yields an efficient outcome. Here the net harm is the $50 deadweight loss. Assuming no overcharge, a $50 penalty deters the inefficient cartel (whose cost saving is less than $50) but not the efficient one (whose cost saving is greater than $50).

2. Suppose there is no capacity constraint at Q_1 in Figure 1 and the cartel lowers marginal cost everywhere to MC_1. Using the net harm rule, the fine equals the aggregate overcharge plus dead-

---

[6] I use the term "net" harm to allow for possible benefits to third parties that result from offenses. This has important implications for optimal penalties which I explore *infra* in part II-B.

[7] Formally, the solution to the optimal line in Figure 1 is as follows: Let B=deadweight loss to consumers from reduced output; C=cost savings brought about by a reduction in output; F=fine; and A=monopoly profit (before taking account of C and F). We want to choose a value for F such that A + C — F (=net monopoly profit) is positive only if C>B. As described in text, F = A + B, then substituting (A + B) for F, A + C — (A + B)>0 only if C>B.

[8] I ignore in this example the monetary equivalent of the costs imposed on buyers by the rationing scheme.

weight loss (areas A + B) only when output falls below the initial competitive output Q_0. This would lead the cartel to produce at least Q_0 units of output. Assuming it produces Q_0, the net harm and hence fine would be zero, and net profits would equal the aggregate cost savings up to Q_0. Any lower output would lead to a fine and to lower net profits.[9] Still, one might view Q_0 as a "second-best" outcome compared to the greater output where MC_1 intersects the demand curve. But it is unclear how to use antitrust penalties to induce the cartel to produce this more efficient output. The required penalty (which would equal the overcharge and deadweight loss computed from the point where MC_1 intersects the demand curve) would remove any incentive for firms to form the cartel that creates the cost savings.[10]

3. The optimal fine or damage rule must be modified when the probability of apprehension and conviction is less than one and enforcement costs are positive. In such circumstances, the fine would equal net harm (which includes enforcement costs per case) divided by the probability of apprehension and conviction. In Figure 1, for example, if enforcement costs per case are $10 and the probability of conviction is one-third, the cartel causes harm of $160 ($100 overcharge plus $50 ordinary deadweight loss plus $10 enforcement costs), and the fine would equal $480 ($160 divided by ⅓). Since the expected value of the fine is $160 (480 x ⅓) firms would form a cartel only if their cost savings were greater than $60. This too is the efficient outcome because the offsetting deadweight loss, including enforcement costs, is $60.

---

[9] Let C equal the cost saving up to Q_0 and let k equal $Q_1/Q_0 < 1$ where Q_1 is any output less than Q_0. At Q_1 the cartel's profit equals A + kC - F. Under the net harm rule, the fine is (A + B). Substituting A + B for F yields a profit kC - B, which is less than C. Since the cartel's net profit is C at Q_0 (because the fine is zero) the cartel prefers Q_0 to any output Q_1 less than Q_0. Notice that if the reduction in marginal cost is sufficiently great and the demand curve sufficiently elastic at Q_0, the cartel's profit-maximizing output may be greater than Q_0.

[10] Patent pooling presents a comparable problem. Consider a group of firms that have valid but competing patents that enable the industry to lower marginal costs everywhere from MC_0 to MC_1. Absent cross-licensing and pooling (perhaps because of antitrust liability) the royalty rate would be close to zero, and firms would produce at the point where MC_1 intersects the demand curve. With pooling, the royalty rate would be positive, output lower, and price higher. The argument for pooling (analogous to the assumption in the text that the cartel "causes" the cost savings) is that its absence would sufficiently reduce the expected returns to innovation and, therefore, discourage firms from investing the resources necessary to develop the innovation.

B. Sanctions for a Cost-Saving Cartel

I have used the example of an efficient cartel to explain the meaning of an optimal antitrust penalty. I have not explained, however, why the cost savings shown in Figure 1 could not have been achieved without the cartel and without restricting output from Q_0 to Q_1. Consider the following example.

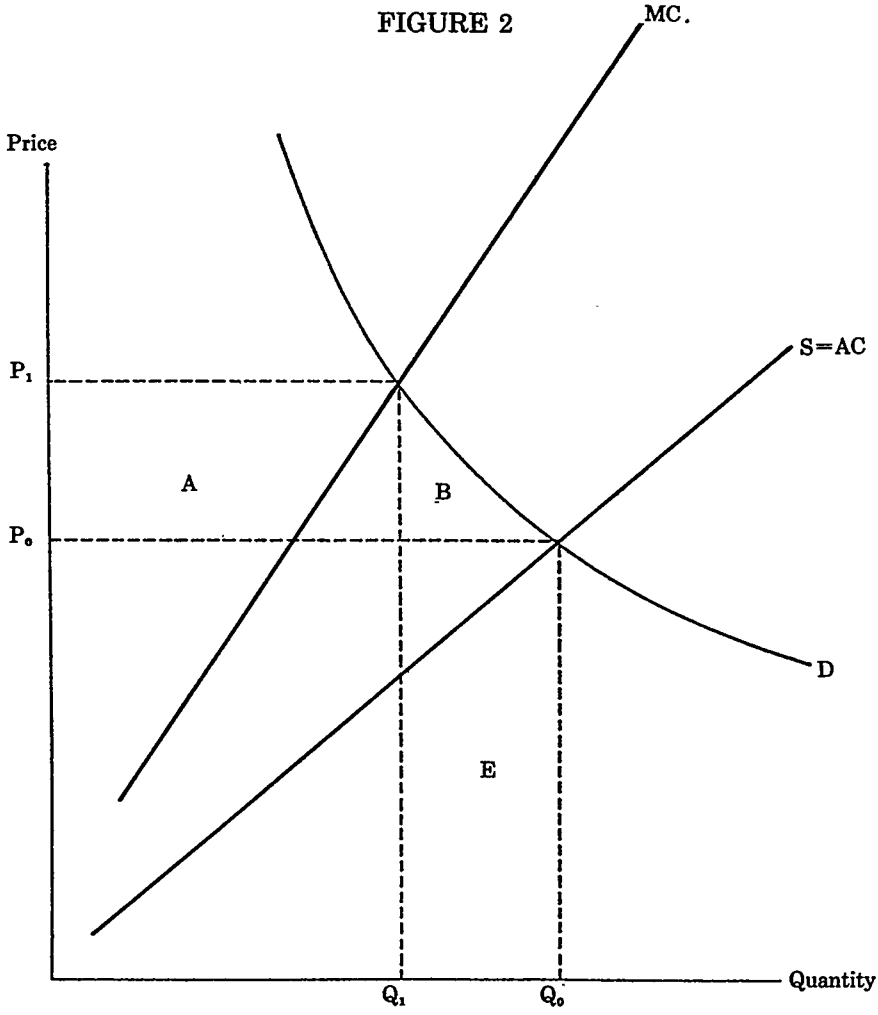
Suppose a group of firms pump oil from a common pool, and no firm has a property right to the oil until it is removed from the ground. This example illustrates a well-known economic problem: in the absence of property rights to a scarce resource, competition carries production beyond the efficient level because no firm takes account of the external technical diseconomies it imposes on other firms. Here an agreement among firms to restrict output and raise price can lead to a more efficient output level.[11] At the same time, if firms are free from antitrust liability, they may restrict output even further, possibly yielding a less efficient outcome than the precartel equilibrium. An antitrust penalty equal to the net harm will give firms an incentive to restrict output, but only to the efficient level.

Figure 2 illustrates the problem. To simplify, I assume that firms are equally efficient, each firm produces one unit of output, and there is free entry. Competition among equally efficient firms and free entry will eliminate any profits. Hence firms will enter the industry until each firm operates where its cost (equal to its average cost) equals price. This occurs at the output Q_0 and price P_0, where demand intersects the industry supply schedule, S. S is also the industry average cost curve because each point on it measures the firm's cost of producing one unit of output. As demand shifts to the right, new firms will enter and impose external technical diseconomies on the old firms, raising their cost. The entrant, however, only considers its cost and not the added cost it imposes on other firms. Thus, marginal cost of producing an additional unit of output (via a new entrant) is greater than average cost (the cost incurred by the new entrant). Therefore, at Q_0 the value of an additional unit of output, measured along the demand curve, is less

[11] The classic article on this subject is Knight, *Some Fallacies in the Interpretation of Social Cost*, 38 Q.J. Econ. 582 (1924), *reprinted in* Readings in Price Theory 160 (G. Stigler & K. Boulding eds. 1952). Knight analyzed the problem of transporting goods along a narrow but well-graded road. *Id.* at 584-88. The main point of Knight's article was to show that if someone owned the road, he would charge a price for access that led to its efficient use. An alternative mechanism to achieve efficiency is for firms to agree to limit their use of the road.

than the marginal cost of producing that output. Observe that Q_0 exceeds the efficient output level Q_1 which occurs when demand intersects marginal cost. Thus, an agreement among firms to restrict output to Q_1 would yield a more efficient allocation of resources. But in the absence of a fine, firms would have an incentive to restrict output below Q_1, to the point where marginal revenue equals marginal cost. This would produce the usual deadweight loss from a cartel or monopoly.

FIGURE 2



The solution is to set a penalty small enough to provide firms with an incentive to reduce output to Q_1 but not so small that they reduce output below Q_1. That fine equals the aggregate overcharge plus deadweight loss measured from the initial output Q_0. This is

the same net harm rule analyzed earlier. I will give an intuitive explanation for this result.[12] If firms continue to operate at Q_0, there is no fine and each firm would earn zero profits (price equals average cost by assumption). Firms would not operate at an output greater than Q_0 because average cost would exceed price and the firms would lose money. Now suppose the firms form a cartel and restrict output to Q_1, the point at which marginal cost equals price. Under the net harm rule presented in this paper, the cartel will be assessed a fine equal to the deadweight loss and net overcharge. After deducting the per-unit overcharge fine of $P_1 - P_0$,[13] each firm will receive the same net price as it did when output was Q_0. On the units no longer produced, Q_0-Q_1, firms forgo revenues equal to P_0 times Q_0-Q_1[14] and in addition incur penalties equal to the deadweight loss from restricting output.[15] In short, the revenue

---

[12] A formal proof of this result follows: Let the fine (F) equal the sum of the aggregate overcharge and deadweight loss to consumers measured from the initial zero profit equilibrium Q_0 and P_0. Note that at Q_0 and P_0 the fine equals zero, and it becomes positive and increases as the cartel restricts output and raises price. We can write F as

$$F = (P\text{-}P_0)Q + \int_{Q}^{Q_0} (P\text{-}P_0)\, dQ$$

where P and Q are the cartel's price ($>P_0$) and output ($< Q_0$) respectively. The cartel's profit (π), after deducting F, is $\pi = PQ - C(Q) - F$, where PQ is the cartel's total revenue and C(Q) is its cost function. Substituting for F yields

$$\pi = P_0Q - C(Q) - \int_{Q}^{Q_0} (P\text{-}P_0)\, dQ$$

The first-order condition for maximizing profits is obtained by taking the derivative of π with respect to Q and setting it equal to zero. This yields $\partial\pi/\partial Q = P_0 - C' + (P\text{-}P_0) = 0$ or $P = C'$, where C' is marginal cost and P is the price along the demand curve in Figure 2. The cartel maximizes profits by setting price equal to marginal cost—i.e., producing at the efficient output Q_1 in Figure 2.

I add that there is no single unique value for the optimal fine. A fine equal to the aggregate overcharge plus deadweight loss computed on *any* initial point between (P_0, Q_0) and (P_1, Q_1) will yield the efficient outcome. For example, imposing a fine on the cartel only if it reduces output below the efficient level, Q_1, will also yield the efficient solution. *See supra* text following note 8.

[13] At Q_1, price is equal to P_1, but after deducting the overcharge fine of $P_1 - P_0$ the net price will be P_0, the same net price as when output was Q_0.

[14] Represented by Area E in Figure 2.

[15] Represented by Area B in Figure 2.

loss to firms from restricting output is the area under the demand curve between Q_1 and Q_0 in Figure 2.[16] At the same time, however, costs fall by the area under the marginal cost curve between Q_1 and Q_0. Since marginal cost is above the demand curve everywhere in this interval, costs fall by more than revenues decline. Therefore, by restricting output the cartel earns a profit equal to the difference between the areas under the marginal cost and demand curve in the Q_0-Q_1 interval.

I have shown that firms earn positive profits at Q_1 compared to zero profits at Q_0, even though they incur an antitrust penalty at Q_1. Hence the penalty would not deter firms from forming the cost-saving cartel. To show that the cartel will produce at Q_1, I must show that Q_1 is the profit-maximizing output, given the penalty. If the cartel were to produce more than Q_1, its additional revenue would consist of two components: (1) a net price P_0 (after deducting the overcharge penalty) on each unit beyond Q_1, and (2) a reduction in the penalty for deadweight loss for the additional units produced. The sum of (1) and (2) is the area under the demand curve beyond Q_1. This is less than the costs of producing the additional units, the area under the marginal cost curve in Figure 2. Hence profits fall when output is greater than Q_1. Similarly, when firms lower output below Q_1, profits decline because costs fall by less than do revenues. In short, the cartel will produce at the efficient output when the penalty equals the aggregate overcharge and deadweight loss computed from the output Q_0 and price P_0.

II. Applications of the Model

Setting the expected fine equal to the net harm is a simple and, once explained, obvious rule for bringing about the efficient level of offenses.[17] Yet as applied it yields some surprising results, as the following examples show.

A. Joint Ventures

Assume competitors form a joint venture to sell their product. Suppose the joint venture has characteristics that one typically associates with a cartel; for example, competitors set up a common sales agency that makes all sales, or they agree not to sell in each

[16] This area under the demand curve is the sum of areas B and E.

[17] To simplify the exposition I continue to assume, unless stated otherwise, that the probability of apprehension and conviction is one and enforcement costs are zero.

other's territory. At the same time, assume that the cost savings associated with the joint venture are so large that the entire marginal cost curve shifts down, thereby increasing output. In terms of Figure 1, output now is greater than Q_0, and net harm, relative to the previous competitive output, is negative. Does this imply that the optimal fine should be zero and the joint venture does not violate the antitrust law? In general the answer is "yes," though two situations must be distinguished: the first is a hypothetical example; the second is the recent ASCAP litigation.[18]

First, suppose that at the time the automobile was introduced, automobile manufacturers formed a common sales agency enabling them to raise price and restrict output. Assume that the cost savings in transportation made possible by the substitution of the automobile for the horse and carriage were so large that in terms of Figure 1 (where the horizontal axis now denotes transportation services) cartel output is greater and price lower than the normal price and output of the pre-automobile period.

Second, consider ASCAP, an association of composers, lyricists, and publishers of music, which acts as a clearing house between its members and those who wish to perform copyrighted works from ASCAP's repertory. Under the typical arrangement between ASCAP and a user, such as a radio or television station, ASCAP issues a nonexclusive blanket license that permits the licensee to perform all of the songs in ASCAP's repertory for an annual fee without any limitation on the number of times each song is performed. Because ASCAP acquires only nonexclusive rights to its members' compositions, each member still retains the right to license performance rights for his compositions. Thus, a user can choose between an ASCAP license or a direct license of performance rights from the copyright holder of songs he wishes to perform.[19] The main advantage of ASCAP's blanket license is greatly reduced transaction and monitoring costs compared to the cost of negotiating and monitoring contracts with individual copyright holders to acquire performance rights to the thousands of songs used each year. In Figure 1 the horizontal axis would denote public performances, and the higher marginal cost curve would represent cost conditions when performance rights are licensed individually. With the blanket license, the marginal cost curve would be lower and coincident with the horizontal axis, resulting in a greater out-

[18] Broadcast Music, Inc. v. CBS, 441 U.S. 1 (1979) (commonly known as the *ASCAP* case).

[19] ASCAP's licensing system is described in *id.* at 4-5.

put and lower price (at the margin), compared to the case of individual licensing of performance rights.[20]

In both the automobile and ASCAP examples, output is greater and price lower than at the original competitive equilibrium position. Does it follow, therefore, that the antitrust law should treat both situations identically? And should the optimal fine be zero to free each group of any liability?

In the automobile example, the commercial development of the technology, not the common sales agency, is the source of the reduction in transportation costs. Comparing the level of output when both the technology and the sales agency exist to its level when both are absent would be relevant only if the sales agency, not the automobile, caused the reduction. The relevant comparison to make for determining whether there is an antitrust injury is between output with and without the common sales agency, *given* the innovation. Without the sales agency, price would have been even lower and output even greater. Therefore, the common sales agency of automobile manufacturers should be treated as an ordinary cartel. The penalty would equal the aggregate overcharge plus the deadweight loss computed on the assumption that the automobile would exist in the absence of the cartel.

In contrast, optimal damages are zero and there should be no antitrust liability in the *ASCAP* case. In *ASCAP* the source of the innovation was the blanket license that reduces transaction and monitoring costs. This makes the blanket license a more attractive option than the alternative of acquiring performance rights individually.[21] Without ASCAP and the blanket license's pooling of performance rights, costs would be higher and output of musical performances would be lower.[22]

---

[20] In effect, the blanket license is an example of two-part pricing: an access charge for performing songs in ASCAP's repertory, and a variable charge equal to zero for additional performances. Since the licensee is entitled to unlimited use, he will expand the number of performances until the added revenue from an additional performance is approximately zero. Because the social cost of using an existing composition is zero, this leads to efficient utilization of compositions (assuming the access charge is not so high that it discourages a potential user from acquiring licenses). This and other efficiency features of the blanket license, however, are not considered here.

[21] The fact that licensees choose the blanket license over the individual licenses means that the charge for the blanket license is less than the costs they would incur if they acquired performance rights individually.

[22] Consider the distinction between a patent pool of competing and complementary or blocking patents. A pool of competing patents is similar to a cartel among automobile manufacturers. By agreeing not to challenge competing patents, the pool enables patent holders to restrict output and increase royalties. In the case of blocking patents, the pool is what enables firms to take advantage of the full benefits of the innovation. Without a pool either

The treatment of the two examples under the antitrust laws is consistent with the economic analysis. The automobile manufacturers would be found guilty of a per se violation of section 1 of the Sherman Act. In the *ASCAP* case CBS argued that the blanket licenses of ASCAP and BMI[23] fixed prices and therefore were per se violations of the antitrust laws.[24] In rejecting this argument, the Supreme Court focused on the transaction cost savings of the blanket license and the alternative of CBS negotiating individual licenses with copyright holders. The Court held these considerations required that the case be judged under the rule of reason, a standard that balances the efficiency or output-enhancing features of a practice with possible restraints on competition.[25] To state this conclusion in terms of the analysis I have developed in this paper, the ASCAP combination causes no net harm, and it is essential to the cost savings. The automobile sales agency, however, is not essential to the cost savings, but it is the source of harm to consumers.

United States v. Sealy, Inc.,[26] illustrates another example of an output-enhancing joint venture. In *Sealy* a group of small mattress manufacturers (Sealy licensees) adopted a common trademark and allocated exclusive territories among themselves to manufacture and sell mattresses under the Sealy trademark. The licensees imposed no territorial restrictions on their sales of non-Sealy mattresses. The licensees accounted for about twenty percent of mattress sales in the United States.[27] The Supreme Court held the territorial allocation to be part of a horizontal scheme to fix prices and found it unlawful. Although territorial allocation

---

the innovation would not be developed commercially or potential licensees would incur large transaction costs to acquire rights to individual patents. This is similar to the added costs that radio and television stations would incur if they acquired performance rights from individuals instead of from ASCAP.

[23] BMI provides essentially the same service as ASCAP. *See* Broadcast Music, Inc. v. CBS, 441 U.S. at 5.

[24] *Id.* at 6.

[25] *Id.* at 2. On remand, ASCAP's blanket license policy was found lawful. *See* CBS v. ASCAP, 620 F.2d 930 (2d Cir. 1980), *cert. denied*, 450 U.S. 970 (1981). In subsequent litigation against ASCAP, a district court has held unlawful the blanket license acquired by television stations because it found that, unlike CBS, the local television stations did not have a commercially realistic alternative to the blanket license. Buffalo Broadcasting Co. v. ASCAP, 546 F. Supp. 274 (S.D.N.Y. 1982). In terms of my analysis this implies that the cost savings from the blanket license were even greater for television stations than for CBS. For this reason, and because the blanket license did not restrict output, the economically correct result would have been to hold lawful the blanket license in *Buffalo Broadcasting*.

[26] 388 U.S. 350 (1967).

[27] R. POSNER & F. EASTERBROOK, *supra* note 1, at 248.

among competitors is a classic antitrust device, its effects here were different. It provided an incentive for each licensee to promote and develop the Sealy trademark in its territory. Each licensee would be constrained from exercising market power within its territory both by non-Sealy manufacturers and other licensees. Viewed in this way the practice in *Sealy* is equivalent to a cost reduction (the trademark lowered the cost of providing information to consumers) that lowers price and increases output. In short, the Sealy arrangement enabled a small group of firms to expand output by overcoming free-rider problems associated with developing a trademark or brand name and therefore optimal damages are zero. Since the net harm was negative, the correct result would have been no liability.

B. The Social Cost of Monopoly

It has been argued recently that the social cost of monopoly is far greater than just its deadweight loss.[28] Because firms spend resources to monopolize a market and to maintain their monopoly position, part of the aggregate overcharge to consumers is transformed into a resource expenditure. In the limit, these expenditures could equal the present value of the monopoly profits. Because these expenditures produce nothing of value, they add to the social cost of a monopoly. For example, in Figure 1 social cost could be greater than $50 (area B) and possibly as high as $150 (area A + B). Will a social cost greater than deadweight loss have any effect on the optimal sanction analysis?

Recall that the optimal fine equals the net harm to the community excluding the violator. It is irrelevant, therefore, how cartel members spend their profits. They may spend $100 or $1 organizing and maintaining the cartel; either way, the optimal fine is $150. This implies further that one does not have to know the gain to the offender to set the optimal fine. Stated differently, the profitability of a violation is not a separate factor in setting the optimal fine. Information on profitability is relevant only insofar as it conveys information on net harm.

I can give a formal explanation of why the optimal fine remains at net harm. Imagine that real resources of $100 are spent on organizing and enforcing the cartel depicted in Figure 1. Al-

[28] *See* Posner, *The Social Costs of Monopoly and Regulation*, 83 J. POL. ECON. 807 (1975); Tullock, *The Welfare Costs of Tariffs, Monopolies, and Theft*, 5 W. ECON. J. 224, 231 (1967).

though the cartel's gross profits are $100 plus any cost saving, its net profits are decreased by the $100 expenditure. By assumption, net profits equal the cost savings from the reduction in output (area C in Figure 1). Suppose the cost saving is $149. In our initial analysis the cartel would have been efficient because a $149 cost saving offsets a $50 deadweight loss. A fine of $150, equal to $100 overcharge plus $50 deadweight loss, would leave cartel members with a $99 profit from setting up the cartel. When the cost of organizing and enforcing the cartel equals the overcharge, however, net profits would equal minus $1. The $150 penalty would exceed the $149 cost saving. Hence the cartel would not be formed. This too is the correct outcome. To obtain $149 cost savings by spending real resources valued at $100 and having consumers suffer a $50 deadweight loss would be inefficient. Suppose instead the cartel lowers cost by $151. Again a $150 fine yields the correct outcome. The cartel would be established because its members gain $1 ($251 gross profits minus $100 costs minus $150 fine). This is efficient because the net resource saving of $51 offsets the $50 deadweight loss to consumers.

C. Cartels with Less than One Hundred Percent Market Share

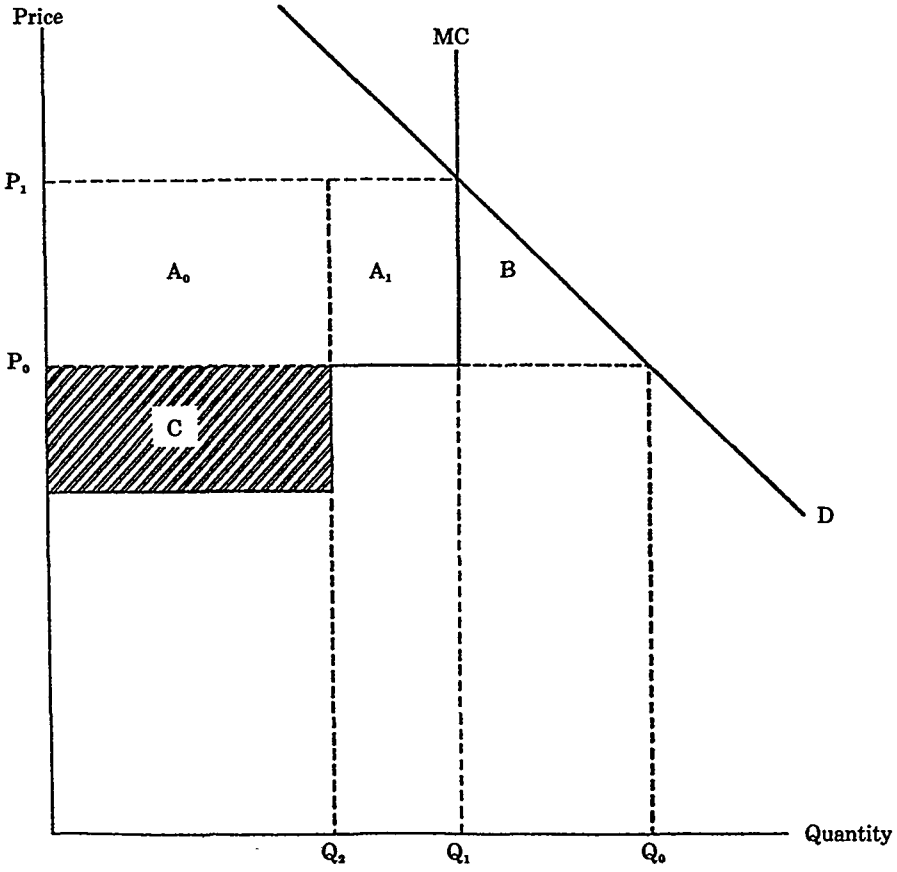
Often not all firms in an industry are cartel members. Assume that nonmembers are equally efficient, have constant marginal (average) cost equal to the prior competitive price (therefore they were not earning economic rents prior to formation of the cartel), and have a fixed capacity that limits their output. The last assumption is necessary to enable the cartel to restrict industry output and raise price above the competitive level.

Figure 3 illustrates the competitive and cartel equilibriums. Price increases from P_0 to P_1, output falls from Q_0 to Q_1, the cartel produces Q_2, and nonmembers, the competitive fringe, produce Q_1-Q_2.[29] Area B measures the deadweight loss (I ignore the possible additional element of deadweight loss analyzed in the previous subpart), areas A_0 and A_1 measure the aggregate overcharge, and area C the cost savings. As before, I assume that C would not occur without the cartel. The key question is should the optimal fine include the overcharge caused by the cartel but not received by its members (area A_1 in Figure 3). That is, should the fine equal A_0

[29] To derive the cartel's output, the cartel would maximize $P(Q)(Q - Q_f) - C(Q_c)$ where Q_f is the fixed fringe output, Q_c is the cartel's output, $C(Q_c)$ the cartel's cost function, and $Q = Q_f + Q_c$.

$+ A_1 + B$ or just $A_0 + B$?

FIGURE 3



Recall that the optimal fine equals the *net* harm to others. This requires, therefore, that any benefits to others be subtracted from the harm to consumers. Since the competitive fringe receives a benefit of A_1 from the higher than competitive price, net harm equals the harm to consumers, areas $A_0 + A_1 + B$, minus area A_1. To see why net not gross harm is the correct rule, consider the following example. Let $A_0 = \$75$, $A_1 = \$25$, $B = \$50$ and $C = \$51$. By assumption, this offense is efficient because the \$51 cost savings exceeds the \$50 deadweight loss. A fine equal to the gross harm of \$150 would yield the wrong result. Since the cartel's profit of \$126 is less than the proposed fine, the cartel would not form,

industry output would remain at the competitive level, and the cost savings would not be realized, resulting in an efficiency loss of $1. In contrast a fine equal to the net harm of $125 would leave the cartel with a profit of $1. The cartel would be formed and society would be better off by $1.

Although the net benefit rule is perfectly general, the conclusion that the cartel should not be liable for any overcharges on units sold by the competitive fringe holds only under the cost conditions illustrated in Figure 3. If the fringe's marginal cost were to exceed the previous competitive price, then their rents or benefits would be less than the harm to consumers on the units purchased from the fringe. In the limit, if the fringe's marginal cost were constant and equal to the cartel price, optimal damages would equal $A_0 + A_1 + B$, the entire overcharge plus the deadweight loss. The more typical case is one of rising marginal cost, implying an optimal fine between $A_0 + A_1 + B$ and $A_0 + B$, depending on the precise cost function.[30]

D. Predatory Pricing[31]

The economic argument against cartels is based on the deadweight loss cartels produce. This might suggest (particularly if there are no cost savings) that a deadweight loss resulting from a monopolizing activity is a sufficient reason to impose a penalty. I now show why this is false.

Consider a monopolist who responds to the entry of an equally efficient firm by reducing price below marginal cost to drive the

---

[30] The issue of allowing customers of the competitive fringe to recover from members of the cartel was recently considered in Mid-West Paper Prods. Co. v. Continental Group, Inc., 596 F.2d 573 (3d Cir. 1979), and *In re* Beef Indus. Antitrust Litig., 600 F.2d 1148 (5th Cir. 1979), *cert. denied*, 449 U.S. 905 (1980). In *Mid-West Paper* the court denied recovery to customers of the competitive fringe, finding that the customers of the fringe lacked standing to sue the cartel. The court gave two main reasons for denying recovery. First, since the plaintiff was not in a direct relationship with the cartel, it was the fringe and not the cartel that received a benefit on the sales to the plaintiff. Second, the court asserted that the plaintiff might have bought from the fringe at the overcharge price even in the absence of the cartel, making it difficult to determine whether the plaintiff was in fact damaged. 596 F.2d at 583-87. Although the first argument is consistent with the approach of this paper, the second seems to conflict with the basic economics of cartels. If the cartel was effective, it raised the price of paper bags (the product in the suit) and thus competitors of the cartel were able to sell at a higher price. If the cartel did not raise the price, and hence customers of the cartel's competitors did not pay a higher price, then neither the customers of the cartel nor its competitors should recover. In *Beef Industry* (a monopsony case) the court allowed ranchers selling beef to nonconspiring retailers to recover from the conspirators. 600 F.2d at 1166 n.24 (claim satisfies "target area" standing test).

[31] The analysis in this part is developed in more detail in Easterbrook, *supra* note 1.

entrant from the industry. The monopolist may have acquired his initial monopoly position lawfully, by being more efficient. Below-cost or predatory pricing to maintain monopoly, however, is unlawful.[32] Instead of monopoly, suppose the industry consisted initially of competitive firms of different sizes. Now let the largest firm attempt to monopolize the market by pricing below cost and driving out other firms that are equally efficient (at the margin). This too is unlawful.[33]

A number of studies of antitrust cases where predatory pricing has been alleged indicate that successful predatory pricing is rare or nonexistent.[34] The typical private case alleging predation is brought by competitors who are still in business but who have lost sales during a period in which predation is alleged or by competitors who have left the business for reasons other than predatory pricing.[35]

In order to analyze the correct damage measure or optimal fine in a predatory pricing case, I distinguish two forms of predation. In one form, predation is attempted but unsuccessful. Rivals may be temporarily driven from the industry during a period of below-cost pricing, but they return when the predator attempts to raise price to the monopoly level. The second form, successful predation, occurs when rivals are driven from the industry for a period sufficient to enable the predator to raise price to the monopoly level and to recoup more than his earlier losses.

Suppose a monopolist temporarily delays entry of equally efficient firms by setting price below cost and expanding output. Eventually, however, new firms enter, price rises to marginal cost, and the predation fails. Figure 4 illustrates the situation. During the period of predation price is P_1 and output Q_1; after predation

---

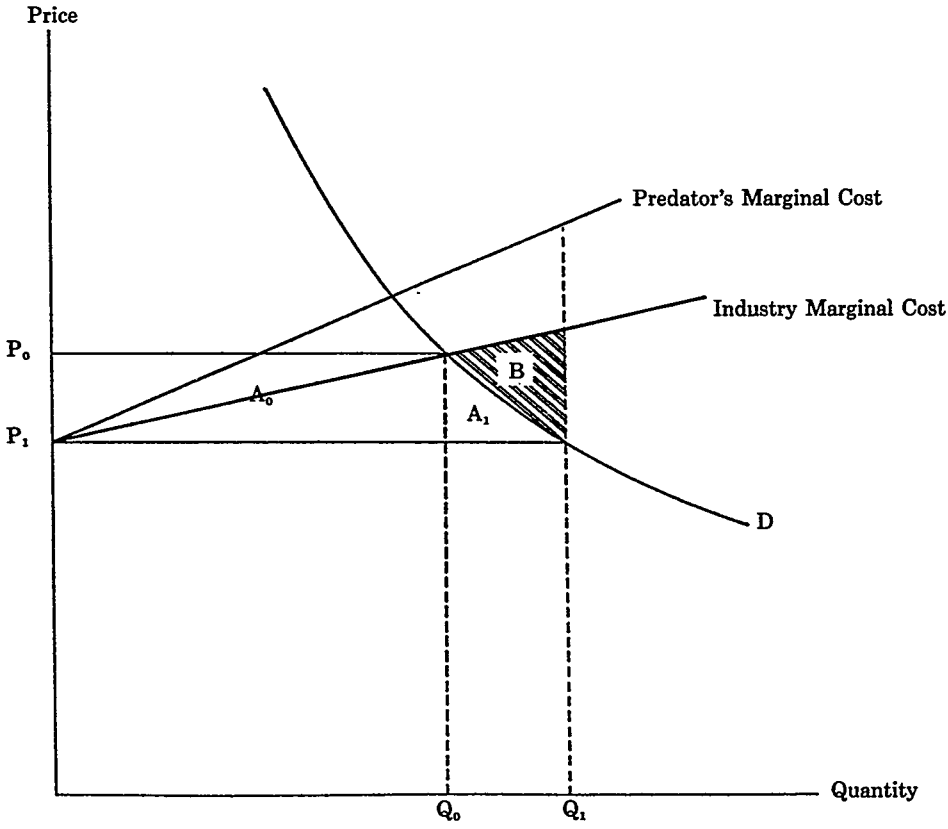
[32] *See* 3 P. AREEDA & D. TURNER, ANTITRUST LAW ¶ 711 (1978).

[33] *See id.*

[34] *See, e.g.*, Elzinga, *Predatory Pricing: The Case of the Gunpowder Trust*, 13 J.L. ECON. 223 (1970); Koller, *The Myth of Predatory Pricing: An Empirical Study*, ANTITRUST L. & ECON. REV., Summer 1971, at 105; McGee, *Predatory Price Cutting: The Standard Oil (N.J.) Case*, 1 J.L. & ECON. 137 (1958); McGee, *Ocean Freight Rate Conferences and the American Merchant Marine*, 27 U. CHI. L. REV. 191 (1960). *See generally* Easterbrook, *supra* note 1, at 312-18. In fact, most predatory pricing claims have failed for inability to show a dangerous probability of success.

[35] *See, e.g.*, United States v. Empire Gas Corp., 537 F.2d 296 (8th Cir. 1976), *cert. denied*, 429 U.S. 1122 (1977). For a discussion of this case, see Easterbrook, *supra* note 1, at 316-17.

FIGURE 4



ends and entry occurs, price returns to P_0 and output to Q_0.

To be sure, below-cost pricing causes a deadweight loss equal, in Figure 4, to the sum of (1) area B, the value to consumers of the additional output $Q_1 - Q_0$ in the predation period less the cheapest method of producing $Q_1 - Q_0$, and (2) the difference between the predator's and industry's marginal cost curves up to Q_1.[36] Deadweight loss is not a sufficient reason to impose a fine on the successful predator. Indeed, the optimal fine in this case is zero because there is no positive net harm to others. Figure 4 shows that buyers gain consumer surplus of A_0 on the units they would have purchased at the competitive price and A_1 on the added units they purchase because of predation, while competitors lose profits of

[36] Had the predator not lowered his prices to increase his market share, the output would have been produced at the lower industry marginal cost. The additional resources needlessly spent to produce the output are a deadweight loss.

kA_0, where k is the fraction of A_0 they would have received as producer surplus at the competitive price and output. Thus, during the period of below cost pricing there is a net gain of $(1 - k)A_0 + A_1$ to consumers and competitors. Figure 4 also shows that the predator's revenue is less than his costs by the difference between P_1Q_1 and the area under his marginal cost curve up to Q_1. Thus, the failed predator bears not only the full deadweight loss of his activity but much more. Because the deadweight loss is a private loss to the predator and not part of the net harm to parties other than the predator, there is no justification for imposing additional penalties. In short, the predator's activity is self-deterring because he bears its full social costs; the rest of us, on balance, are benefited. And provided this is so, the net harm, and therefore the optimal penalty, is zero.

Successful predation occurs when the monopolist is eventually able to raise price above marginal cost, more than recouping the losses incurred during the predation period. The preceding discussion demonstrates that net harm, and hence the optimal penalty, is the present value of the aggregate overcharge and deadweight loss during the recoupment period minus the net benefit to others during the period of predation.[37]

Two further points should be mentioned. First, at the time predation begins one may not know whether it will succeed. Yet the optimal penalty is zero in one case but positive in the other. Assuming we are able eventually to distinguish failed from successful predation, there should be no penalty or recovery until it is clear that the predation is successful. In effect, this rule would delay enforcement proceedings until the recoupment period begins. Moreover, it would probably eliminate most predatory pricing suits brought by competitors because the suits typically involve instances of failed predation.

Second, the measure of damages in private suits against successful predators should not be the competitor's lost profits but rather the present value of the ultimate overcharge to consumers and deadweight loss minus net benefits during the period of below-cost pricing.[38] The only justification for allowing competitors to sue

---

[37] Where predation is successful, this is of course a positive number. The predator is able to increase future profits such that the present value of the future profits is greater than the present cost of predation.

[38] If antitrust laws are a means to increase consumer welfare, then losses only to consumers, not to competitors, are relevant. For a discussion of the consumer welfare goals of antitrust, see R. BORK, THE ANTITRUST PARADOX 66 (1978); R. POSNER & F. EASTERBROOK, *supra* note 1, at 154.

is that they are more efficient private enforcers than are consumers, because the claim of any individual consumer is small relative to his costs of bringing suit. Thus, consumer enforcement may lead to too few cases being brought and hence to underdeterrence of antitrust violations.[39] Assuming competitors, in effect, are standing in the shoes of consumers, optimal deterrence requires that the damage award to a competitor should depend on net harm, not the competitor's lost profits.[40]

III. Antitrust as an Intentional Tort

In a recent paper, Richard Posner and I developed an economic model of intentional torts, such as assault, battery, conversion, and defamation, and showed that the model explained most important common law doctrines governing intentional torts.[41] Although price-fixing and monopolization are analytically equivalent to an intentional tort, and the Supreme Court sometimes has used tort principles to analyze antitrust violations,[42] we did not examine the implications of our model for antitrust policy. I do so now, after first summarizing our intentional torts model.

A. An Economic Model of Intentional Torts

From an economic standpoint, the way to distinguish intentional from ordinary torts is that in the former the injurer spends resources to increase the probability of harming the victim, but in the latter both parties spend resources to reduce this probability.[43] The intentional tortfeasor is by definition the party able

---

[39] Although there are methods for aggregating many small claims (e.g., class actions and parens patriae suits), these are considered costly enforcement methods relative to competitor suits. *See* Landes & Posner, *supra* note 1, at 607-08.

[40] If consumer enforcement is more efficient because, for example, there are several consumers with potentially large recoveries, then consumers, not competitors, should recover.

[41] *See* Landes & Posner, *An Economic Theory of Intentional Torts*, Int'l Rev. L. & Econ. 127 (1981).

[42] *See, e.g.*, Associated Gen. Contractors v. California State Council of Carpenters, 103 S. Ct. 897, 904-07 (1983); *id.* at 914 (Marshall, J., dissenting) (comparison of antitrust to intentional torts); Texas Indus., Inc. v. Radcliff Materials, Inc., 451 U.S. 630, 634 & n.5 (1981).

[43] This definition of intentional is actually too broad because it would define mistakes or self-defense as intentional torts. Our paper deals with this by including as costs the expense of avoiding mistakes or of not undertaking self-defense. Landes & Posner, *supra* note 41, at 137-38. Then the full costs of reducing the probability of injury due to a mistake or failure to undertake self-defense may be positive. In another sense our definition seems too restrictive because it excludes torts where the costs of avoidance are positive, but trivial relative to the expected harm. We deal with this under the category of reckless behavior. *Id.* at 130-32.

to avoid the tort at the lowest cost. He need only refrain from spending resources to avoid the tort; all other parties must spend resources to avoid injury. This simple difference leads to different implications concerning liability for intentional and ordinary torts. They are as follows.

First, provided the harm to the victim exceeds or equals the injurer's gain (before deducting his costs of committing the tort), the optimal or first-best solution (ignoring enforcement costs) is for the intentional tort not to take place. This saves the costs of committing the tort (i.e., the costs of avoidance are negative, while they are positive in the unintentional tort) and eliminates the harm which, by assumption, is not less than the injurer's gain. Holding the injurer liable for the victim's damages, adjusted upward if the probability of identifying and collecting damages from the injurer is less than one, will lead him not to inflict the injury.

Second, the injurer's liability should not be eliminated or reduced if the victim is contributorily negligent. For example, suppose the victim's expected harm is $75 if he spends nothing on self-protection but $50 if he spends $10. In an ordinary tort, the victim would be contributorily negligent for not spending $10 on self-protection and would not recover his damages under a negligence standard. Contributory negligence is not a defense to the injurer's liability in an intentional tort. Since expected costs are zero if the tort is not committed but positive even if the victim undertakes self-protection, the optimal outcome is for the tort not to occur. Making the victim's recovery depend on his not being contributorily negligent might lead him to spend $10 on self-protection. This sum is saved by a rule holding the injurer liable without allowing a contributory negligence defense.

Third, in ordinary torts, efficiency can be achieved without compensating the victims.[44] In contrast, efficiency usually requires compensating the victim of an intentional tort to prevent him from spending resources to avoid the tort. (Of course if the law perfectly deterred all intentional torts, there would be no occasions for compensation.) Suppose, however, certain activities are classified as intentional torts even though the expected harm is less than the injurer's net expected gain (after deducting his costs). Call these activities socially beneficial or efficient intentional torts. An example might be a hiker lost in a storm who breaks into an empty

[44] This is one of the basic results of the now standard economic analysis of a negligence system with a contributory negligence defense. *See* Landes & Posner, *The Positive Economic Theory of Tort Law*, 15 GA. L. REV. 851, 876-77 (1981).

cabin to obtain food and shelter. If the victim is not compensated (even if the offender pays a fine to the state), he would expend resources to reduce his expected harm. This, in turn, might deter the injurer from committing the tort because the victim's precautions have made it more costly to do so. To deter the tort, however, would be inefficient. In contrast, if the victim is fully compensated for his injury, he will not undertake expenditures for self-protection because he receives no benefit from them. As a result, the offender will commit the efficient tort.

B. Antitrust as an Intentional Tort

Nothing in the above analysis would be changed if we substituted "cartel" or "monopolization" for "intentional tort": the first-best, or optimal, outcome is to impose a penalty sufficient to deter the inefficient antitrust violation; there should be no defense of contributory negligence; and the victim should be compensated. I consider these conclusions in turn.

Part I examined the cartel that lowered costs by less than the deadweight loss. This is equivalent to the intentional tort that harms the victim by more than it benefits the injurer. The per se rule against price fixing, by always holding the cartel liable for the harm it does, deters firms from forming inefficient cartels. Similarly, the common law rule that holds the intentional tortfeasor liable for the victim's harm creates incentives for the efficient outcome.

As with intentional torts, the victim's conduct should not be relevant to the antitrust violator's liability. In short, there should be no defense of contributory negligence. Consider the following example. Let a manufacturer of nuclear reactors agree to provide its customers (utilities) with their future requirements of uranium at today's price plus an adjustment for future increases in mining costs. Assume the manufacturer holds a negligible inventory of uranium but plans to fulfill its contracts by making future purchases. Now suppose producers of uranium form a cartel and increase price. In response the manufacturer breaches its contracts with utilities, is sued by them, and in turn sues the cartel for treble damages.[45] The cartel argues that the manufacturer was contribu-

[45] This example is similar to the recent uranium litigation. Westinghouse, a supplier of nuclear reactors, breached its uranium contracts with utilities, was sued by the utilities, and in turn sued firms it claimed to be cartel members. The uranium dispute gave rise to a torrent of litigation. *See, e.g.*, Westinghouse Elec. Corp. v. Gulf Oil Corp., 588 F.2d 221 (7th Cir. 1978) (price-fixing suit against oil company); *In re* Westinghouse Elec. Corp. Uranium

torily negligent by failing to cover its contracts in the futures market and by signing contracts with utilities at low prices when it should have known about the cartel. Even if it is true that a well-managed or reasonably prudent firm would have covered in the futures market or obtained information on the likely formation of the cartel, however, our intentional torts analysis says the cartel's defense should not be allowed. If the defense were allowed, purchasers would avoid being contributorily negligent by incurring additional expenditures both to uncover possible antitrust violations among their suppliers and to hedge in futures markets against possible antitrust violations. Knowing this, suppliers would be deterred from forming inefficient cartels. This outcome, however, is less efficient than one where both cartels are deterred and expenditures on self-protection are avoided. We can achieve the latter outcome, as shown by our intentional torts analysis, by holding the cartel liable and rejecting a defense based on the victim's negligence.

The principal argument for awarding damages to private enforcers is to create powerful enforcement incentives. Awarding damages to victims, as we showed in our analysis of intentional torts, has the additional benefit of reducing socially inefficient expenditures on self-protection.[46] A component of the latter cost is the deadweight loss that arises when consumers substitute toward lower valued products. A possible benefit, therefore, of victim compensation is to reduce the incentive to substitute away from the monopolized product, thereby to reduce or eliminate the deadweight loss. This surprising result comes about in the following way.

To simplify, I make the following assumptions: the competitive price and output are $1 and 100 units; the cartel raises price by $1; the costs of discovering the cartel and prosecuting its members are zero; the probability of collecting damages is one; and consistent with my earlier analysis the penalty equals the sum of the aggregate overcharge ($1 x units purchased at the cartel price) and the deadweight loss. The latter two assumptions imply that firms will form a cartel only if the savings in costs made possible by their joint action more than offset the deadweight loss. What, however,

---

Contracts Litig., 563 F.2d 992 (10th Cir. 1977) (contract breach).

[46] When it is costly to apprehend and convict antitrust violators, victim enforcement may not be the preferred policy. For example, if the victims of a cartel are consumers who individually have insufficient stakes to pursue a violator, then enforcement by injured competitors of the violator or by public agencies may be the better solution.

is the size of the deadweight loss? Since customers of the cartel, by assumption, recover their $1 overcharge in a costless legal proceeding, the *net* price of the cartel's product is the same as the competitive price. No deadweight loss occurs because consumers continue to purchase at the cartel price the same number of units they purchased at the competitive price.

Before introducing some complications, note that the typical legal measure of damages in a price-fixing or monopolization case brought by a customer is treble the illegal overcharge times the number of units purchased.[47] Trebling in part reflects the belief that some price-fixing and monopolization offenses are not detected. Finally, attorney's fees and other costs may be assessed against the culpable defendant. One criticism of this damage computation is that it does not compensate for the deadweight loss. That is, customers who reduce their purchases or make none at all do not recover on the units they would have purchased if price had been lower.[48] Although one can support this exclusion because of difficulties in proving how much a consumer would have bought but did not, there is also a theoretical reason: anticipated compensation of the full overcharge can eliminate any deadweight loss. If consumers anticipate full recovery for any overcharge, they will not reduce their purchases below the original competitive level. Thus, for a given overcharge there will be no reduction in quantity demanded, and no deadweight loss.[49] I should mention several qualifications to this result.

---

[47] *See, e.g.,* Hanover Shoe, Inc. v. United Shoe Mach. Corp., 392 U.S. 481, 489-94 (1968). *See generally* R. POSNER & F. EASTERBROOK, *supra* note 1, at 549.

[48] There are other more important criticisms of damage awards. The most important is that treble damages are awarded even if the violation is not concealable and the probability of convicting the defendant is near unity. Furthermore, even if the offense is concealable and the probability less than one, there is no way of knowing whether three is the correct multiplier.

[49] Anticipated compensation for the overcharge will not always eliminate the deadweight loss. Consider first the example in Figure 1 where it was necessary to reduce output (from Q_0 to Q_1) to achieve cost savings (area C). Yet if demand is perfectly inelastic, how can the cartel reduce output? One way is to ration available supply among purchasers. Although consumers would be willing to pay any price above P_0 for the rationed output because their net price after deducting the overcharge recovery is still P_0, the net profits of the cartel is its cost saving. Since the latter can be less than the deadweight loss brought about by rationing, failure to compensate for the deadweight loss can lead to an inefficient outcome. Consider also the example in Figure 2. There we eliminated the deadweight loss caused by external technical diseconomies by reducing output from Q_0 to Q_1. But if output cannot be reduced, we cannot eliminate the deadweight loss. Alternatively, if rationing reduces output, the cartel will cut back too far, to the point where MC equals P_0 because the net demand curve faced by the cartel (after deducting the per unit overcharge) is horizontal at P_0.

(1) Customers may not anticipate any antitrust recovery. If they do not, they will substitute away from the monopolized product and a deadweight loss will occur. Yet an antitrust recovery, though uncertain, is a potential source of revenue and there is no strong reason to treat it differently than other uncertain revenue sources. (2) If customers are risk averse, uncertain damages, though compensating in terms of expected value, may not be sufficient to prevent substitution away from the monopolized product. (3) Even though the defendant pays the optimal penalty, only part may be received by the victim. For example, some of the victim's costs may not be recoverable,[50] and in a large class action suit a substantial share of the recovery may go not to victims but to lawyers.

Although points (2) and (3) imply that the deadweight loss is not eliminated, it is still reduced compared to the no compensation alternative.[51] A more fundamental objection is that I have ignored how the cartel responds to knowing that buyers anticipate an antitrust recovery. If the recovery is complete, the cartel, in effect, faces a perfectly inelastic demand curve at the competitive output. Although this leads to an indeterminate price, output remains at the competitive output and deadweight loss is zero.[52] However, if compensation is incomplete, the cartel faces, compared to the no-compensation case, a less elastic (but not perfectly inelastic) demand curve. In response, the cartel will set a higher price. If damages equal the net harm, the efficiency aspects of the rule that we analyzed earlier continue to hold.[53] Yet, if there is no compensation for the consumers' deadweight loss, then anticipated but partial recovery for the overcharge can lead to inefficient outcomes.[54]

---

[50] For example, the plaintiff's time costs are not recoverable.

[51] Partial compensation has the effect of rotating the original demand curve to the right around the competitive equilibrium because the consumer will treat a given overcharge as representing a smaller net increase in price. (The net increase equals the overcharge minus partial recovery, per unit purchased.) Therefore, for a given overcharge, the reduction in quantity demanded will decrease as the partial recovery increases. And the smaller the reduction in quantity demanded, the smaller the deadweight loss. Note that we still measure deadweight loss along the original demand curve because the difference between it and the demand curve that takes account of partial recovery denotes the expected recovery per unit purchased.

[52] The conclusion is subject to the qualifications presented *supra* at note 49.

[53] One exception is the analysis depicted in Figure 2. Since the demand curve is less elastic compared to the no compensation case, the cartel equilibrium occurs where the less elastic demand curve intersects marginal cost. This leads to an output greater than the efficient output Q_1.

[54] The analysis is similar to that presented *supra* at note 49 except that here the output reduction is brought about by a higher price.

Conclusion

In this paper I have developed the economic theory of the optimal antitrust penalty. The optimal penalty should equal the net harm to persons other than the offender, adjusted upward if the probability of apprehension and conviction is less than one. This sanction encourages efficient behavior. It will only deter those violations that impose deadweight losses greater than the cost savings brought about by the violation. It will not deter efficient violations, those where the cost savings exceed the deadweight loss.

I applied the analysis of optimal penalties to the following enforcement issues in antitrust. I showed that the optimal penalty is zero for joint ventures that increase output. Similarly, whether one characterizes cooperative behavior among firms as a joint venture or as a cartel depends generally on whether cooperation expands or restricts output. I showed also that the optimal penalty is unaffected when part or all of the monopoly overcharge is transformed into a social loss by firms spending resources to organize and maintain the cartel. I examined whether customers who bought from competitors of the cartel at a higher than competitive price should be able to recover from the cartel for the overcharge. If the competitors are no less efficient than cartel members, no recovery should be allowed. In the case of predatory pricing, I showed that the optimal penalty is zero when below-cost pricing is followed immediately by entry. When below-cost pricing delays entry enough for the monopolist to more than recoup the losses incurred during the predation period, then the optimal penalty is positive and equals the net harm in the recoupment period minus the gains to consumers during the period of below-cost pricing. In the final section of the paper I applied the economic analysis of intentional torts to an antitrust violation. This enabled me to show the efficiency gains from compensating victims of an antitrust violation and to explain why it is inefficient to allow a defense of contributory negligence to an antitrust violation.