2015s-37

How Much Do Cartel Overcharge? (The "Working Paper" Version)

Marcel Boyer, Rachidi Kotchoni

Série Scientifique/Scientific Series

2015s-37

How Much Do Cartel Overcharge? (The "Working Paper" Version)

Marcel Boyer, Rachidi Kotchoni

Série Scientifique
Scientific Series

Montréal
Juillet 2015

CIRANO
Allier savoir et décision

Centre interuniversitaire de recherche en analyse des organisations

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du Ministère de l'Économie, de l'Innovation et des Exportations, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de l' l'Économie, de l'Innovation et des Exportations, and grants and research mandates obtained by its research teams.

Les partenaires du CIRANO

Partenaire majeur
Ministère de l'Économie, de l'Innovation et des Exportations

Partenaires corporatifs
Autorité des marchés financiers
Banque de développement du Canada
Banque du Canada
Banque Laurentienne du Canada
Banque Nationale du Canada
Bell Canada
BMO Groupe financier
Caisse de dépôt et placement du Québec
Fédération des caisses Desjardins du Québec
Financière Sun Life, Québec
Gaz Métro
Hydro-Québec
Industrie Canada
Intact
Investissements PSP
Ministère des Finances du Québec
Power Corporation du Canada
Rio Tinto Alcan
Ville de Montréal

Partenaires universitaires
École Polytechnique de Montréal
École de technologie supérieure (ÉTS)
HEC Montréal
Institut national de la recherche scientifique (INRS)
McGill University
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web.

ISSN 2292-0838 (en ligne)

Partenaire financier
Économie, Innovation et Exportations
Québec

How Much Do Cartel Overcharge? *
(The "Working Paper" Version) Ì

Marcel Boyer [†], *Rachidi Kotchoni* [‡]

¨ This document is an extended version of our article published in Review of Industrial Organisation (2015). It serves two purposes. First, it is intended for students and professionals who may not have the econometric background to go easily through our published article; hence it is more self-contained and pedagogical. Second, it develops the links with our previous working papers on the same topic, namely CIRANO 2011s-35 The Econometrics of Cartel Overcharges, CIRANO 2012s-15 How Much Do Cartels Typically Overcharge?, and TSE WP-462 (2014) How Much Do Cartels Overcharge? These working papers have now been withdrawn for circulation.

Résumé/abstract

L'estimation des surprix des cartels est au coeur de la politique de lutte aux cartels car elle est un élément clé de la détermination des pénalités. Connor et Lande (2008) survolent la littérature sur les majorations de prix des cartels et concluent à une augmentation moyenne variant entre 31% et 49%. Considérant un échantillon plus grand, Connor (2010) trouve une moyenne de 50,4% pour les cartels réussis. Cependant, les données utilisées dans ces études sont des estimations obtenues à partir de méthodologies, sources, et contextes différents plutôt que d'observations directes. De ce fait, ces données héritent potentiellement d'erreurs de modélisation et d'estimation, ainsi que de biais d'endogenéité et de publication. L'analyse directe des surprix dans l'échantillon de Connor révèle une distribution asymétrique, une importante hétérogénéité et la présence d'observations aberrantes. Au-delà du fait que les estimations des surprix sont potentiellement biaisées, l'estimation d'un modèle de régression linéaire avec de telles données sans un traitement adéquat des problèmes identifiés ci-dessus pourait produire des résultats fallacieux. Nous présentons une méta-analyse dans l'esprit de Connor and Bolotova (2006), mais qui tient compte adéquatement des problèmes mentionnés ci-dessus. Après correction des biais, nous obtenons une moyenne et une médiane de majorations de prix de l'ordre de 15,47% et 16,01%. Nos résultats débouchent sur des enjeux importants en matière de politique de la concurrence.

Mots-clés : Antitrust, Surprix de cartel, Heckman, Heckit, Divergence de Kullback-Leibler, Meta-analyse

*We are very grateful to the Editor (Lawrence White) of the Review of Industrial Organization for his patience as well as for his generous and challenging comments and suggestions. We also benefited from remarks by two anonymous referees and by Jimmy Royer and René Garcia. We remain solely responsible for its content.

[†] Emeritus Professor of Economics, Université de Montréal, Associate Member, Toulouse School of Economics
Fellow of CIRANO, www.cirano.qc.ca/~boyerm
[‡] Assistant Professor, African School of Economics, http://africanschoolofeconomics.com/faculty

The estimation of cartel overcharges lies at the heart of antitrust policy on cartel prosecution as it constitutes a key element in the determination of fines. Connor and Lande (2008) conducted a survey of cartels and found a mean overcharge estimate in the range of 31% to 49%. By examining more sources, Connor (2010) finds a mean of 50.4% for successful cartels. However, the data used in those studies are estimates obtained from different methodologies, sources and contexts rather than from direct observations. Therefore, these data are subject to model error, estimation error, endogeneity bias, and publication bias. An examination of the Connor database reveals that the universe of overcharge estimate is asymmetric, heterogenous and contains a number of influential observations. Beside the fact that overcharge estimates are potentially biased, fitting a linear regression model to the data without providing a carefull treatment of the problems raised above may produce distorted results. We conduct a meta-analysis of cartel overcharge estimates in the spirit of Connor and Bolotova (2006) while providing a sound treatment of these matters. We find bias-corrected mean and median overcharge estimates of 15.47% and 16.01%. Clearly, our results have significant antitrust policy implications.

Keywords : Antitrust, Cartel overcharges, Heckman, Heckit, Kullback-Leibler divergence, Meta-analysis

1. Introduction

Our aim in this paper is to determine the order of magnitude of the average cartel overcharge, based on a database of cartels used by Connor (2010).[1] This database contains overcharge estimates (OE) obtained from a survey of several studies on cartels as well as three types of variables. The first group (Y) consists of variables that describe the cartel episode (e.g. duration, scope, geography, etc.). The second group (Z) consists of factors that are posterior to the cartel episode (e.g. estimation method, publication source). The third group (W) consists of a single dummy variable which indicates whether the cartel was "found or pleaded guilty". While Y and W are likely related to the true overcharge, Z are likely to capture potential biases.

The raw OE data are themselves potentially biased and the variable W is likely endogenous. Hence, a naive OLS regression of the OE on Y, W and Z should be avoided. To verify whether the OE are biased or not, we perform a meta-regression analysis in the spirit of Connor and Bolotova (2006) who show that part of the variability of OE is indeed due to the bias factors.

We use a Kullback-Leibler divergence to compare the probability of an OE being larger than some value θ conditional on (Y,W) to the same probability conditional on (Y,W,Z). The two conditional probabilities are quite close for $\theta \in [0\%, 65\%]$ but diverge sharply for $\theta > 65\%$. This divergence is caused by the fact that the joint distribution of the variables involved in the probit models specified for the probability of (OE>θ) become degenerate as θ exceeds a certain threshold. Next, we regress the logarithm of OE on Y, W and Z on increasing subsamples of type $(0, \theta]$. The results allow us to identify the range $OE \in (0\%, 49\%]$ as the most reliable subsample on which a Heckit regression can be performed to obtain unbiased estimates of coefficients which can be used to infer bias-corrected OE for the whole sample.

Applying the methodology described above, we find mean and median bias-corrected OE of of 16.68% and 16.17% for the subsample of effective cartels (with stricly positive OE), and of 15.47% and 16.01% for the whole sample. These numbers are significantly lower than the means and medians of the raw OE data. Moreover, the comparison of bias-corrected mean and median OE reveals a fairly homogenous behavior of cartels across different types, geographical locations, and time (antitrust) periods.

The paper is organized as follows. In Section 2, we describe the context and the literature surrounding our research. Section 3 presents the raw OE and discusses data problems. Section 4 illustrates the danger of converting Lerner indices into OE while ignoring the competitive mark-

[1] The database actually describes cartel episodes. Each episode is treated as a different observation. There is no formal proof or admission of guilt for approximately one third of the observations. Hence, the data include convicted cartels as well as alleged ones. We sincerely thank Professor John Connor for generously making his database available to us.

ups. Sections 5 and 6 presents the methodology used to detect the presence of bias in the OE data. Section 7 presents the determinants of cartel overcharge unveiled by our meta-analysis. Section 8 presents the steps of our bias-correction methodology and the summary statistics for the bias-corrected OE. Section 9 presents an analysis of variance for the OE bias. Section 10 concludes. Table 2 and Appendix A contain summary statistics of the database. Appendix B summarises our previous attempts to model and bias-correct the raw OE. Appendices C and D discuss methodological aspects not covered in the main text.

2. The Context

The United States Sentencing Guidelines recommends a base fine of 10% of the affected volume of commerce to a firm convicted of cartel activity, plus another 10% for the harms "*inflicted upon consumers who are unable or for other reasons do not buy the product at the higher price*". This yields a fine of 20%, subject to further adjustments for aggravating and mitigating factors. The observed total financial fines generally fall in a range from 15% to 80% of affected sales. Moreover, there is a possibility of incarceration for the individuals involved in the collusion. For fiscal years 2010-2014 (5 years), U.S. antitrust prosecutions resulted in over US$5.14 billion in criminal fines and penalties, including the largest cartel fine in history namely $1.14 billion for the liquid crystal display (LCD) panel cartel, and more than 295 years of jail time.

In the European Union, the determination of fines accounts for the severity of the damages inflicted upon consumers, suppliers and clients as well as aggravating and mitigating factors. The basic fine is set in a range of 0% to 30% of affected commerce plus 15% to 25% as an additional dissuasive measure. However, the total fine must not exceed 10% of the "*worlwide group turnover in the financial year preceding the decision*."[2] For fiscal years 2010-2014 (5 years), European antitrust prosecutions resulted in over €8.93 billion in criminal fines and penalties. This amount includes the highest fine in history, namely €1.47 billion for the TV and computer monitor tubes cartel.

In Canada, penalties for "agreements between competitors to fix prices, restrict production, or allocate sales, customers or territories" (a *per se* offense since March 2010) may reach $25 million per count and up to 14 years in prison. But actual penalties levied have not been close to those maxima yet. As in the U.S., the Canadian Competition Bureau uses a proxy for economic harm of 20% of the volume of affected commerce to set fines, 10% for the basic overcharge and 10% for other harm including the deadweight loss.[3]

---

[2] Source: "*The EU competition Rules on Cartels*," document published the law firm *Slaughter and May* in 2012. Publically available on the website of the company.

[3] The harm caused to society is equal to the cartel overcharge, which being a transfer from buyers to sellers may not appear as an efficiency loss or harm, plus additional social costs, namely the deadweight losses

Connor and Lande (2008) examined a large number of OE studies and found an average in the range 31% to 49% and a median in the range 22% to 25%. Based on this, they concluded that *"the current Sentencing Commission presumption that cartels overcharge on average by 10% is much too low, and the current levels of cartel penalties should be increased significantly "*. A similar study conducted by Connor (2010) concludes that *"...penalty guidelines aimed at optimally deterring cartels ought to be increased"*.

Combe and Monnier (2011, 2013) performed an analysis of 64 cartels prosecuted by the European Commission and concluded that *"fines imposed against cartels by the European Commission are overall sub optimal."*

In criticizing the Canadian Competition Bureau approach, Kearney (2009) wrote "*The assumption of an average overcharge of 10 percent also has been put into question by economic survey evidence which suggests that the median long-run overcharge is much greater than 10 percent. Research conducted by Professor John Connor indicates 'that the median long-run overcharge for all types of cartels over all time periods is 25.0 percent …' Accordingly, an assessment of economic harm based on an estimated overcharge of 10 percent is not supported by the empirical evidence*."

However, Cohen and Scheffman (1989) argued that an increase of 1% of a price above its natural competition level usually results in a reduction of sales of more than 1%. Based on this, they concluded with respect to the USSG guidelines that *"at least in price-fixing cases involving a large volume of commerce, ten percent is almost certainly too high"*. Adler and Laing (1997) and Denger (2003) also judged that the fines imposed by the US Department of Justice are "*astronomical*" or "*excessive*" (Connor and Bolotova 2006, p. 1112).

Allain, Boyer and Ponssard (2011) develop a dynamic model of cartel stability and find that the cartel-level fines imposed by the European Commission in the 64 cartels analysed by Combe and Monnier are on averagre above the proper deterrence level.[4] Considering a more recent database at the firm level, Allain et alii (2015) conclude that the majority of firm-level fines imposed by the European Commission over the period 2005-2012 are above the deterrence level.

Hence, there is disagreement among specialists about the magnitude of cartel overcharges and thus, about optimal fines. Our paper contributes to this debate by providing an econometric method which appropriately deals with the limitations of the Connor database.

---

and the resources devoted to antitrust authorities for fighting cartels, plus other effects such as the impact on investment and employment, on entry and exit dynamics, on innovation and learning curve, etc.

[4] Boyer (2013) discusses this recent literature.

3. The Connor Database

As mentioned earlier, the database used for our study is an extended version of the one used in Connor (2010). The raw sample consists of 1178 cartels episodes, from which 59 are discarded for missing information. This leaves us with a sample of 1119 cartels with OE ranging from 0% to 1800%. The mean OE is 45.5% on the whole sample and 49% for the subsample of strictly positive OE. The mean is 20.5% for the cartels with OE lying strictly between 0% and 49%, representing 69.9% of the sample. OE that are larger than 49% represent 22.9% of the sample and the average OE for this subsample is 136.2%.

However, the sample means of 45.5% and 49% are influenced by a small number of outliers. Roughly 1% of OE are larger than 400% and when the 5% largest observations are left out of the sample, the average OE drops from 49% to 32%. These outliers should be treated carefully when using econometric methods that are sensitive to their presence (e.g., OLS regressions). The skewness of the distribution (Figure 1)[5] implies a significant difference between the means and medians.

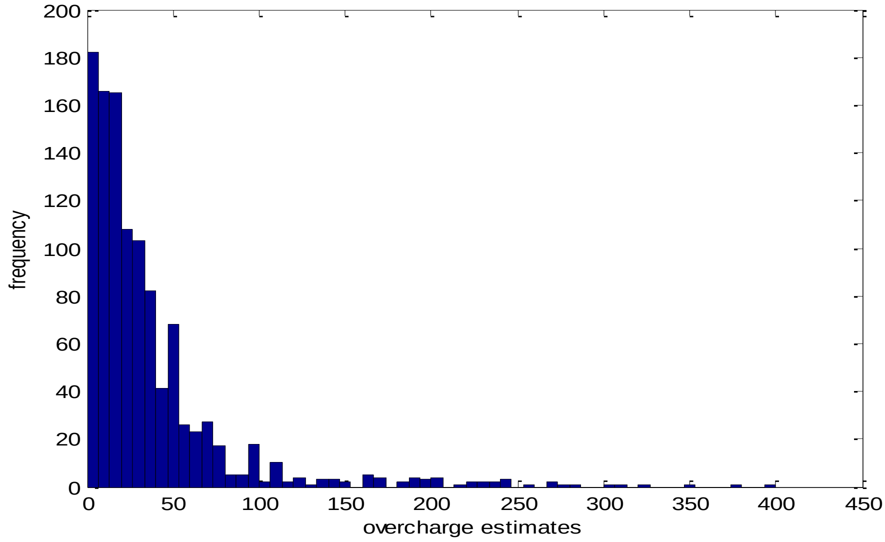
It should be emphasized that the overcharge data consists of estimates previously published by different experts and researchers. Therefore, they are potentially subject to model errors, estimation errors, endogeneity bias, and sample selection.[6]

The raw overcharge data are quite heterogenous across regions, scope (domestic vs international), and time periods (Table 1). This clearly raises aggregation problems. Indeed, the average overcharge obtained for the whole sample is meaningful only if the conditions that determine the but-for price are the same across time and markets. As noted by Levenstein and Suslow (2003), "*The reported price increases vary widely by industry and by source."*

[5] Connor (2014) finds a long run median overcharge of 23.0% and a mean of at least 49.0% for all cartels of all times. The skewness problem is pointed out by Connor and Lande (2008) and Connor (2010). It should be emphasized that Connor has been conservative in recording some of the OE, notably by tagging and excluding peak estimates from the sample (see Connor, 2010, pp. 48 and following).

[6] Ehmer and Rosati (2009) point out that many estimates in Connor's sample are obtained from "*a simple calculation of the difference between prices charged during the operation of the cartel and in other periods or other markets that are believed to be competitive. By completely neglecting all other factors that can cause prices to change, the authors of these estimates simplistically attribute the entire price variation to the effects of the cartel.*"

Figure 1
Overcharge Estimates: Distribution skewed to the right.



Note: Overcharges larger than 400% (1% of the sample) are not shown on this figure.

Table 1
Means and medians of raw OE per location and types of cartels.

		All Cartels	OE>0%	0%<OE≤49%	OE>49%	Cartels Before 1973	Cartels After 1973
All locations	Mean	45.46	49.01	20.46	136.21	61.98	38.88
	Median	23.00	25.00	18.30	73.00	29.00	21.20
	prop.	100.00	92.77	69.88	22.87	28.51	71.49
US	Mean	38.15	42.03	19.44	123.82	47.79	33.58
	Median	20.50	23.50	17.30	69.20	30.50	16.80
	prop.	30.02	27.25	21.36	5.90	9.65	20.37
EU	Mean	42.65	45.57	19.07	113.01	43.83	41.86
	Median	23.00	25.00	16.05	75.00	24.75	20.40
	prop.	33.51	31.36	22.52	8.85	13.40	20.11
Domestic	Mean	33.60	36.91	18.43	137.58	35.42	32.79
	Median	17.05	19.00	16.10	69.70	20.50	16.45
	prop.	46.82	42.62	36.01	6.61	14.47	32.35
International	Mean	55.91	59.29	22.62	135.66	89.38	43.91
	Median	30.00	31.75	22.00	74.45	37.00	27.50
	prop.	53.17	50.13	33.87	16.26	14.03	39.14

Note: The prop. are percentages of the total Connor sample (1119 cartels).

The following variables are listed in the Connor database:

$\hat{\theta}$ - the overcharge esimates (OE), which is described in Table 1.

Y_1 - Duration, discretized: 1 if duration is less than 5 years, 2 if duration is from 6 to 10 years, 3 if duration is from 11 to 15 years, and 4 if duration is 16 years or more.

Y_2- Scope: equals 1 if domestic and 0 if international.

Y_3- Bid rigging: equals 1 if Yes and 0 if No.

Y_4- Geographic market: five dummy variables for US, EU, ASIA, ROW including Latin America, and WORLD cartels that cannot be associated to a head region.

Y_5- Antitrust law regime in the US: six dummy variables for P1 (1770-1890), P2 (1891-1919), P3 (1920-1945), P4 (1946-1973), P5 (1974-1990), and P6 (1991-2004).

W- Found guilty or pleaded guilty: equals 1 if Yes and 0 if No.

Z_1- Overcharge estimation method: dummy variables for Price before conspiracy (PBEFOR), Price war (PWAR), Price after conspiracy (PAFTER), Yardstick (YARDST), Cost based (COST), Econometric modeling (ECON), Historical case study with no method specified (HISTOR), legal decisions (LEGAL) and other unspecified methods (OTHER).

Z_2- Type of publication: dummy variables for Peer reviewed journal (JOURNAL), Chapters in a book (EDBOOK), Monograph or book (MONOGR), Government report (GOVREP), Court or antitrust authority source (COURT), Newspapers (NEWSPAPER) Working paper (WORKP), and Speech or conference (SPEECH).

The Y variables describe the alleged cartel episode and are therefore objectively related to the true overcharge. The period dummy variables (Y_5) are used in Connor and Bolotova (2006) to capture the effect of changes in US antritrust law regimes over time. These dummies are closely related to eras identified and studied at length by Kovacic and Shapiro (2000). The early time periods (P1, P2 and P3) are likely to be more important for the US than for the rest of the world. This argues for interacting those time periods with the US geographical market dummy in our regressions.

The Z variables describes circumstances that are posterior to the occurrence of the cartel episode. They are rather subjective and may therefore generate an overcharge estimation bias. Regarding the estimation methods (Z_1), the traditional "yardstick" involves a cross-section comparison of firms, products or markets. The "before-and-after" and the "price war" methods might be considered as the time series version of the "yardstick". The "cost-based" and the "econometric" methods represent more sophisticated measurement efforts at implementing either version of the yardstick method. See Appendix B for a brief description of these overecharge estimation methods.

The variable W (Guilty or not) is alone in its category. It is potentially related to the true overcharge while open to subjectivity: a guilty plea or judgement is not a foolproof indicator of guilt, but an entity that chooses to plead guilty has likely been involved in an effective price-fixing conspiracy. This argues for treating W as a distinct category.

Our study uses the Y, W and Z variables described above to explain the OE.[7] Table 2 presents summary statistics of all variables. Additional summary statistics are presented in Appendix A.

Table 2
Summary statistics of the Connor database.

	mean	std	min	max
OE	45.46	102.90	0	1800
Duration	9.25	11.86	1	109
duration (discrete)	1.86	1.06	1	4
Domestic	0.47	0.50	0	1
BidRig	0.20	0.40	0	1
Guilty	0.66	0.48	0	1
US	0.30	0.46	0	1
EU	0.34	0.47	0	1
ASIA	0.09	0.28	0	1
ROW	0.04	0.19	0	1
WORLD	0.24	0.43	0	1
P1	0.01	0.08	0	1
P2	0.10	0.30	0	1
P3	0.11	0.31	0	1
P4	0.15	0.35	0	1
P5	0.57	0.50	0	1
P6	0.07	0.25	0	1
OTHER	0.06	0.24	0	1
HISTOR	0.02	0.13	0	1
PBEFOR	0.27	0.44	0	1
PWAR	0.02	0.14	0	1
PAFTER	0.13	0.34	0	1
COST	0.05	0.21	0	1
YARDST	0.14	0.35	0	1
ECON	0.14	0.34	0	1
LEGAL	0.18	0.38	0	1
JOURNAL	0.22	0.41	0	1
MONOGR	0.23	0.42	0	1
EDBOOK	0.06	0.24	0	1
GOVREP	0.23	0.42	0	1
COURT	0.01	0.09	0	1
NEWSPAPER	0.18	0.38	0	1
WORKP	0.01	0.07	0	1
SPEECH	0.06	0.24	0	1
Sample size		1119		

[7] In addition to Y_1 through Y_5, the vector Y contains the interactions of the US geographical market with the periods P_1, P_2 and P_3.

4. The Proper Characterization of the But-for Price

To illustrate the necessity to properly characterize the but-for price, let us consider the estimation error one can make by converting Lerner indexes directly into cartel overcharges.

Let $\tilde{p}$ be the price imposed by a cartel and p the but-for price, i.e. the price that would prevail absent the cartel. The cartel overcharge expressed as a percentage of the but-for price is given by $\delta = (\tilde{p} - p)/p$. While the cartel price $\tilde{p}$ is observed, the but-for price p needs to be estimated.

An important cause of bias in OE resides in the difficulties raised by the proper characterization of the but-for environment. Indeed, the observed time series of prices are resultant of several causes. For instance, an inelastic demand may grant a firm significant market power that translate into high mark-ups. Product differentiation can cause a previously pure-and-perfect competitive market to behave as a monopolistic competition one. However, oligopolistic markets have margins over *MC* that can be significantly larger than zero. As noted by Morrison (1990), *"The empirical results suggest that mark-ups in most U.S. manufacturing firms have increased over time, and tend to be countercyclical."* Hall (1988, Table 4) claims that the ratio *p/c* is in the range of 2 to 4 in US industries.

The proper but-for price is equal to the marginal cost plus a margin. In pure and perfect competition, this margin is low and even close to zero. The accurate assessment of this margin is quite important when converting Lerner indices into cartel overcharges.[8] The Lerner index of market power is defined as:

$$L = \frac{p - c}{p} \qquad (1)$$

where p is the market price and c is the marginal cost (MC). If the condition that would prevail in the absence of a cartel is pure and perfect competition, the but-for price is by $p = c$ so that $L = 0$. The corresponding overcharge in the cartelized market is given by $\delta = \frac{\tilde{p} - p}{p}$ and the Lerner index is $L = \frac{\tilde{p} - c}{\tilde{p}}$. The Lerner index is converted into overcharge via the formula $\delta = \frac{L}{1 - L}$.

In general, the competitive but-for price is equal to c plus a margin m: $p = c + m$. Likewise, the Lerner index in the cartelized market should be calculated as $L \equiv \frac{\tilde{m}}{c + \tilde{m}}$, where $\tilde{m}$ is the inflated mark-up due to the cartel. Therefore, the true cartel overcharge is given by $\delta \equiv \frac{\tilde{m} - m}{c + m}$. However, the overcharge that would be inferred from the Lerner index (wrongly) assuming pure and perfect competition as benchmark is:

[8] An overcharge calculation approach based on the Lerner index can fall within the family of econometric methods or cost-based methods depending on how this index is estimated.

$$\tilde{\delta} \equiv \frac{L}{1-L} = \frac{\tilde{m}}{c} = \delta + \frac{m}{c}(\delta + 1) \quad (2)$$

Equation (2) illustrates the danger of converting Lerner indices into OE by ignoring the existence of competitive mark-ups in the but-for price.[9] If the true overcharge is $\delta = 10\%$ and $\frac{m}{c} = 20\%$, the estimation bias implied by $\tilde{\delta}$ is 32%, i.e. more than three times the true value. Note that the bias is increasing in both the true δ and $\frac{m}{c}$ (see the following Table A) The other overcharge estimation methods are not necessarily exempt of biases either.[10]

Table A
Pitfall in the Conversion a Lerner Index Into an OE: Constant overcharge, constant margin, and bias increasing as the marginal cost decreases.

Parameters	Values				
δ	10%	10%	10%	10%	10%
m	0.02	0.02	0.02	0.02	0.02
c	0.20	0.1625	0.1250	0.0875	0.05
$\frac{m}{c}$	10%	12.3 %	16%	22.9%	40%
$\text{Bias} = \frac{m}{c}(\delta + 1)$	11%	14%	18%	25%	44%
$\tilde{\delta} = \delta + Bias$	21%	24%	28%	35%	54%
$\text{Bias}/\tilde{\delta}$	52.4%	58.3%	64.3%	71.4%	81.5%

5. A Formal Assessment of the Quality of the OE Data

We consider assessing the effectiveness profile of cartels from the OE data available to us, where the effectiveness profile is defined as the probability $\Pr(\theta_i > \theta | Y_i, W_i)$ of the true overcharge being strictly larger than a given thresold θ, for $\theta \geq 0$.

To this end, we assume the existence of a latent variable X^* such that: $\theta_i > \theta$ if and only if $X^* > 0$ and $X^* = a + Y_i b + W_i c + \varepsilon$, where $\varepsilon \sim N(0, \sigma)$. The variable X^* is specific to each θ although this dependence is not made explicit by the notation. Under these assumptions, we have:

$$\Pr(\theta_i > \theta | Y_i, W_i) = \Pr(X^* > 0 | Y_i, W_i) = \Pr(a + Y_i b + W_i c + \varepsilon > 0),$$

$$= \Pr(\varepsilon > -a - Y_i b - W_i c) = \Phi\,(a + Y_i b + W_i c) \quad (3)$$

[9] Table 2 of the appendix to Connor (2010) provides a "*Summary of Price-Fixing Damages, Social-Science Studies*". The table presents 280 cartels, their OE and a description of the estimation method. Of the 280 OE, 51 have been obtained by conversion of a Lerner index. Connor (2014) reports Lerner indices as OE without conversion, thereby subject to upward biases.

[10] For instance, see White (2001) for a critic of the before-and-after methods employed by Connor (1997) in the ADM-lysine case.

Unfortunately, we do not observe θ_i. Instead, we observe an estimate $\hat{\theta}_i$ which is potentially influenced by the exogenous (bias) factors Z_i. Thus, we can reasonably assume that:

$$\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right) = \Phi\left(\tilde{a} + Y_i\tilde{b} + W_i\tilde{c} + Z_i\tilde{d}\right) \tag{4}$$

If $\hat{\theta}_i$ were unaffected by the bias factors Z_i, then it would be an unbiased estimator of θ_i and $\tilde{d}$ should be equal to zero in Equation (4). We would then have:

$$\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right) = \Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right) \cong \Pr(\theta_i > \theta | Y_i, W_i)$$

Otherwise, the term $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right)$ can be quite different from both $\Pr(\theta_i > \theta | Y_i, W_i)$ and $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right)$. The impact of Z_i on OE can be detected by examining the following Kullback-Leibler divergence:

$$\Delta(\theta) = \frac{1}{n}\sum_{i=1}^{n} \Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right)\left[\log \Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right) - \log \Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right)\right] \tag{5}$$

The Kullback-Leibler divergence (see Kullback and Leibler, 1951) tells how dissimilar two distributions are. It is strictly positive if the two distributions are different and equals zero if and only if they coincide.

Indeed, consider two probability distribution functions f(x) and g(x). Jensen's inequality implies:

$$\log E_g\left(\frac{f}{g}\right) \geq E_g \log\left(\frac{f}{g}\right)$$

where E_g is the expectations with respect to g(x). However, the LHS of the previous inequality is equal to zero since:

$$\log E_g\left(\frac{f}{g}\right) = \log \int \frac{f(x)}{g(x)} g(x)dx = \log \int f(x)dx = \log (1) = 0$$

Therefore:

$$E_g \log\left(\frac{f}{g}\right) = \int \log\left(\frac{f(x)}{g(x)}\right) g(x)dx \leq 0$$

Inverting the fraction inside the log yields:

$$\int \log\left(\frac{g(x)}{f(x)}\right) g(x)dx = \int \left(\log(g(x) - \log f(x)\right) g(x)dx \geq 0$$

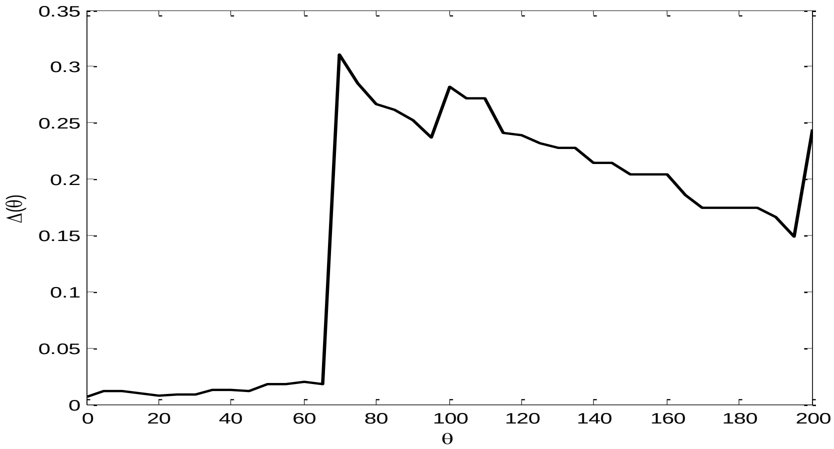
Finally letting $g(x) \equiv \Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right)$, $f(x) \equiv \Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right)$ and replacing the integral by a discrete summation lead to the expression provided for $\Delta(\theta)$ in equation (5).

Any discrepancy between $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right)$ and $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right)$ is attributable to Z_i. Figure 2 plots the values of the Kullback-Leibler divergence $\Delta(\theta)$ on the y-axis against θ on the x-axis. To obtain Figure 2, we estimate $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right)$ and $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right)$ for θ increasing by steps of 5% from 0% to 200%; hence, 40 Probit equations are estimated in total for each of the conditional probabilities $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right)$ and $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right)$.

We note that the predicted values of $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right)$ and $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right)$ agree to a large extent when θ lies between 0% and 65%. However, the two conditional probabilities diverge dramatically as soon as $\theta > 65\%$. This suggests that either OE lying above 65% are heavily biased or there is a quality of the data issue with some of the dummy regressors used in the Probit models.

Figure 2

Detecting the impact of the bias factors via the Kullback-Leibler divergence



Note: The Kullback-Leibler distance, Δ(θ), is on the y-axis and θ on the x-axis. The Probit models that predict $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i\right)$ and $\Pr\left(\hat{\theta}_i > \theta | Y_i, W_i, Z_i\right)$ are estimated by using a step of 5% for θ. The large jump observed at θ=65 is caused either by huge biases in OE or by other data problems involving the regressors.

In an attempt to understand the shape of Figure 2, we examine the averages of the explanatory variables on selected ranges of OE (See Table 3). We see for instance that there is no cartel from the ROW with overcharge above 60%. There is no US cartels of the period P1 (the USxP1 row) with more than 50% overcharge. Likewise, the number of OE collected from historical case studies and from speeches fall to zero after 70%. The latter fact deserves attention as the jump observed on Figure 2 occurs precisely between $\theta = 65\%$ and $\theta = 70\%$.

Some subtle data problems may remain unnoticed. For instance, the sample averages of HISTOR and SPEECH are the same for all headers of Table 3. A carefully examination of the data shows that 217 of the 219 OE lying above 50% are obtained via estimation methods other than "historical case studies" and released through publication media other than "speeches" (i.e., 217 of 219 cartels satisfy HISTOR=0 & SPEECH=0). In particular, there is no cartel such that HISTOR=1 & SPEECH=1. This kind of data problem will eventually translate into multi-collinearities.

Table 3

Summary statistics: Average values of the explanatory variables on selected ranges of OE

	OE>50	OE>60	OE>65	OE>70	OE>75
duration	1.927	1.910	1.950	1.985	2.009
domestic	0.288	0.270	0.275	0.267	0.274
BidRig	0.082	0.079	0.063	0.067	0.051
Guilty	0.543	0.551	0.525	0.519	0.521
US	0.260	0.270	0.263	0.237	0.222
EU	0.374	0.354	0.369	0.400	0.402
ASIA	0.091	0.101	0.094	0.096	0.103
ROW	0.005	**0.000**	**0.000**	**0.000**	**0.000**
P1	0.009	0.011	0.013	0.015	0.009
P3	0.146	0.152	0.156	0.170	0.179
P4	0.142	0.135	0.144	0.163	0.154
P5	0.137	0.146	0.138	0.126	0.128
P6	0.452	0.455	0.444	0.430	0.436
USxP1	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
USxP2	0.078	0.073	0.075	0.059	0.060
USxP3	0.046	0.056	0.063	0.059	0.060
OTHER	0.027	0.034	0.031	0.037	0.034
HISTOR	0.005	0.006	0.006	**0.000**	**0.000**
PBEFOR	0.292	0.258	0.281	0.274	0.248
PWAR	0.037	0.022	0.019	0.015	0.009
PAFTER	0.155	0.169	0.150	0.141	0.128
COST	0.068	0.084	0.088	0.074	0.077
YARDST	0.215	0.236	0.244	0.267	0.299
LEGAL	0.078	0.073	0.069	0.074	0.085
JOURNAL	0.169	0.169	0.150	0.133	0.094
MONOGR	0.269	0.264	0.281	0.296	0.291
EDBOOK	0.105	0.107	0.113	0.111	0.111
COURT	0.237	0.247	0.244	0.252	0.282
NEWSPAPER	0.023	0.028	0.025	0.015	0.017
WORKP	0.096	0.073	0.075	0.074	0.068
SPEECH	0.005	0.006	0.006	**0.000**	**0.000**
Sample size	219	178	160	135	117

Note: There is no cartel from the ROW with overcharge above 60%. Also, the number of OE collected from historical case studies and from speeches fall to zero after 70%.

To support the claim that biases in OE are attributable to the Z variables, it is necessary to assess to which extent our results are impacted by the data problems identified above. For robustness check, we repeat the exercise of Figure 2 after some data transformations. First, we remove the dummy variable ROW, thereby assuming that the reference group for the geographical market is "WORLD+ROW". Second, we merge the interaction variables USxP1 and USxP2. Third, we merge the estimation methods HISTOR and OTHER. Finally, we merge the publication sources NEWSPAPER and SPEECH. Table 4 shows the estimated Probit models for the binary variables defined by the headers of Table 3 after applying the data transformations. There is no visible identification problem based on the magnitudes of the estimated coefficients. Indeed, explosive coefficient estimates would be suggestive of the presence of identification issues.[11]

Table 4
Probit estimation results for selected ranges of OE > θ

	OE>50	OE>60	OE>65	OE>70	OE>75
Duration	0.05	0.05	0.06	0.06	0.07
domestic	-0.82***	-0.89***	-0.78***	-0.75***	-0.64***
BidRig	-0.38***	-0.41**	-0.50***	-0.39**	-0.57***
Guilty	-0.06	-0.01	-0.07	-0.05	-0.12
US	0.38**	0.45***	0.36**	0.31*	0.21
EU	0.19	0.14	0.15	0.19	0.15
ASIA	0.64***	0.72***	0.60***	0.65***	0.55**
P1	-0.11	0.31	0.45	0.72	0.61
P3	-0.52*	-0.45	-0.45	-0.42	-0.16
P4	-0.46	-0.40	-0.32	-0.27	-0.09
P5	-0.68**	-0.43	-0.43	-0.57	-0.26
P6	-0.74**	-0.53	-0.42	-0.42	-0.06
US x (P1+P2)	-0.09	-0.15	-0.08	-0.27	0.03
US x P3	0.50	0.70**	0.81**	0.70*	0.73**
OTHER+ HISTOR	-0.47*	-0.27	-0.30	-0.43	-0.53*
PBEFOR	0.17	0.11	0.16	0.02	-0.16
PWAR	0.61*	0.14	-0.02	-0.20	-0.58
PAFTER	0.28	0.37*	0.26	0.09	-0.08
COST	0.53**	0.77***	0.75***	0.48*	0.41
YARDST	0.39**	0.52***	0.54***	0.45**	0.48**
LEGAL	-0.19	-0.18	-0.18	-0.22	-0.24
JOURNAL	-0.34	-0.46*	-0.55**	-0.73***	-1.02***
MONOGR	-0.41**	-0.46**	-0.42*	-0.45**	-0.47**
EDBOOK	0.02	-0.16	-0.10	-0.19	-0.37
COURT	-0.12	-0.25	-0.24	-0.38	-0.40
NEWSPAPER+SPEECH	0.31	0.36	0.30	-0.39	-0.37
WORKP	-0.52**	-0.80***	-0.75***	-0.85***	-0.98***
Sample size	219	178	160	135	117

Note: The Probit models are estimated using transformed data. The table shows the estimated coefficients when ROW is removed so that the reference group for the geographical market is "WORLD+ROW", USxP1 and USxP2 are merged into USx(P1+P2),the estimation method HISTOR is merged with OTHER, and the publication sources NEWSPAPER and SPEECH are merged. Throughout the paper, ***, **, and * indicate that a coefficent is significant at 1%, 5% and 10% level, respectively.

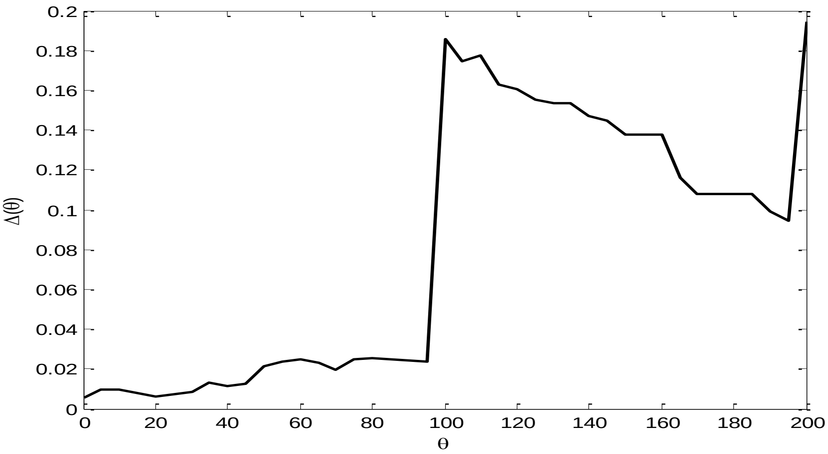
Figure 3 shows the curve of the Kullback-Leibler divergence based on the transformed data.[12] This curve has the same shape as on Figure 2. However, the data transformations have moved the jump from 65% on Figure 2 to around θ=95% on Figure 3. This suggests that the jumps seen on

[11] As a foretaste of our analysis of the determinants of cartel overcharge, we see in Table 4 that domestic cartel are less conducive to price increases than international cartels for all subgroups parametrized by θ. Similarly for bid rigging cartels compared to other cartels and for cartels reproted in monographies, court documents, and wortking papers compared with those reported in government reports (the reference group). We observe also that Asian cartels seem more conducive to price increases than cartels in the WORLD+ROW reference group for all subgroups parametrized by θ.

[12] Here too, we estimate 40 Probit equations for $\Pr(\hat{\theta}_i > \theta | Y_i, W_i)$ and for $\Pr(\hat{\theta}_i > \theta | Y_i, W_i, Z_i)$ to obtain Figure 3.

both figures are caused by the fact that the joint distribution of the variables involved in the probit models of $\Pr(\hat{\theta}_i > \theta | Y_i, W_i)$ and $\Pr(\hat{\theta}_i > \theta | Y_i, W_i, Z_i)$ become degenerate as the cursor θ is moved above certain levels.[13]

Figure 3
The Kullback-Leibler divergence after correcting identification issues



Note: Δ(θ) is on the y-axis and θ on the x-axis. The Probit models that predict $\Pr(\hat{\theta}_i > \theta | Y_i, W_i)$ and $\Pr(\hat{\theta}_i > \theta | Y_i, W_i, Z_i)$ are estimated by using a step of 5% for θ. The data transformations have moved the jump from θ=65 on Figure 2 to θ=95 on Figure 3.

In an effort to better understand the nature of the jumps seen on Figures 2 and 3, we examine separately the two components of the Kullback-Leibler divergence given by the following, where $\Delta(\theta) = \Delta_0(\theta) - \Delta_1(\theta)$:[14]

$$\Delta_0(\theta) = \frac{1}{n}\sum_{i=1}^{n} \Pr(\hat{\theta}_i > \theta | Y_i, W_i) \log \Pr(\hat{\theta}_i > \theta | Y_i, W_i) \tag{6}$$

$$\Delta_1(\theta) = \frac{1}{n}\sum_{i=1}^{n} \Pr(\hat{\theta}_i > \theta | Y_i, W_i) \log \Pr(\hat{\theta}_i > \theta | Y_i, W_i, Z_i), \tag{7}$$

Figure 4 shows the curves of $\Delta_0(\theta)$ and $\Delta_1(\theta)$ based on the transformed data. We see that $\Delta_1(\theta)$ diverges abruptly from $\Delta_0(\theta)$ at θ=95%, which is exactly where the large jump occurs on Figure 3. The fact that the curve of $\Delta_0(\theta)$ is smooth everywhere indicates that only the Z variables are causing the jump.

We have seen previously that the Kullback-Leibler distance $\Delta(\theta) = \Delta_0(\theta) - \Delta_1(\theta)$ is positive. Hence, $\Delta_0(\theta)$ is the upper bound of $\Delta_1(\theta)$ as the probabilities $\Pr(\hat{\theta}_i > \theta | Y_i, W_i, Z_i)$ approach $\Pr(\hat{\theta}_i > \theta | Y_i, W_i)$ uniformly over the sample. Any significant improvement in the model fit induced by Z_i will translate into a visible divergence between $\Delta_0(\theta)$ and $\Delta_1(\theta)$. We see on

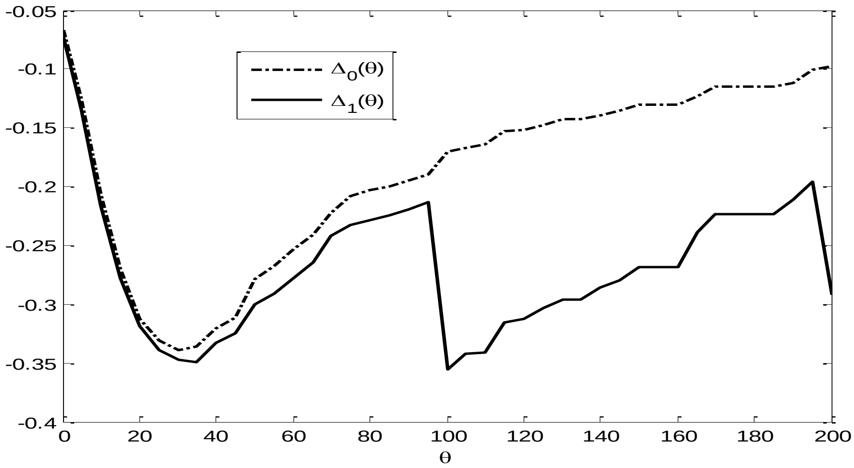
[13] By degeneracy, it is meant that some of our dummy regressors has no variation above a certain threshold of OE, as shown by Table 3.

[14] Note that $\Delta_0(\theta)$ and $\Delta_1(\theta)$ are negative since they involve the log of probabilities.

Figure 3 that the Kullback-Leibler divergence increases slowly as one moves from θ=0% to θ=95%.

In summary, the magnitude of the overcharge estimation bias is increasing with the raw OE, but the large jumps seen on Figures 2 and 3 are caused by data problems. This suggests to treat the subsample of cartels with OE>65% (14.29% of the sample) with caution. Note that nowhere is claimed that OE lying above 65% are all biased upward, nor that OE lying below 65% are all exempt of bias. Indeed, the Kullback-Leibler distance is not necessarily robust to positive biases that leave the proportion of OE larger than θ unaltered. The fact that the probabilities $\Pr(\hat{\theta}_i > \theta|Y_i, W_i)$ and $\Pr(\hat{\theta}_i > \theta|Y_i, W_i, Z_i)$ agree much at $\theta = 0$ means that Z_i does not affect the proportion of zeros OE, but it might influence the size of positive OE. The next step of our analysis deals with the latter aspect, i.e., the effect of Z_i on the conditional mean $E(\hat{\theta}_i|Y_i, W_i, Z_i)$.

Figure 4
Decomposition of $\Delta(\theta)$ into two components



Note: Any significant improvement in the model fit induced by Z_i will translate into a visible difference between $\Delta_0(\theta)$ and $\Delta_1(\theta)$. The difference between $\Delta_0(\theta)$ and $\Delta_1(\theta)$ becomes visible around θ=20 and increases slowly up to θ=95. However, $\Delta_1(\theta)$ diverges suddenly from $\Delta_0(\theta)$ after θ=95. The jump seen of Figure 4 occurs precisely at θ=95.

6. A Meta-Analysis of the Cartel OE

Meta-analyses are used in experimental fields to summarize the findings of studies on a particular topic. They may also be used to verify if the conditions of experiments impact their results. Hunter and Schmidt (2004) write: "*[...] In our view, the purpose of meta-analysis is to estimate what the results would have been had all the studies been conducted without methodological limitations and flaws*". The meta-analysis conducted here is consistent with this statement. Our goal is to understand what causes the bias in the raw OE, unveil the determinants of cartel overcharge, and predict bias-corrected OE.

It is reasonable to expect that the true overcharge depends on the conspiracy period, the duration of the cartel, the characteristics of the firms involved in the collusion, and factors alike. However, we do not observe the true overcharge. Instead, we observe an estimate of it that is equal to the actual overcharge plus a bias. This bias can be positive, null or negative. Hence in addition to factors that affect the true overcharge, we can expect the OE to be sensitive to subjective factors that may cause the bias, namely the estimation method or the publication source which are "posterior" to the occurrence of the conspiracy. Formally, the bias is defined as the influence of factors that affect the OE, but not the true overcharges.

If the true overcharge θ_i were observable, our objective would reduce to understanding what causes the bias in $\hat{\theta}_i$. To this end, we would simply regress $\log\hat{\theta}_i - \log\theta_i$ on Z_i. The bias-correction of $\hat{\theta}_i$ and the prediction of θ_i then become minor issues. As θ_i is not available to us, we specify a model where $\log\hat{\theta}_i$ is on the LHS and potential determinants of the true overcharge and of the estimation bias are on the RHS. In this approach, endogeneity issues regarding the determinants of the true overcharge need to be addressed.

We specify a log-linear meta-analysis model on truncated subsamples of type $\mathrm{OE} \in (0, \theta]$, where θ varies from 25% to 70% by steps of 1%.[15] The model estimated is:

$$\log\hat{\theta}_i = \beta_0 + Y_i\beta_1 + W_i\beta_2 + Z_i\beta_3 + \widehat{\imath mr}_i\beta_4 + e_i \qquad (8)$$

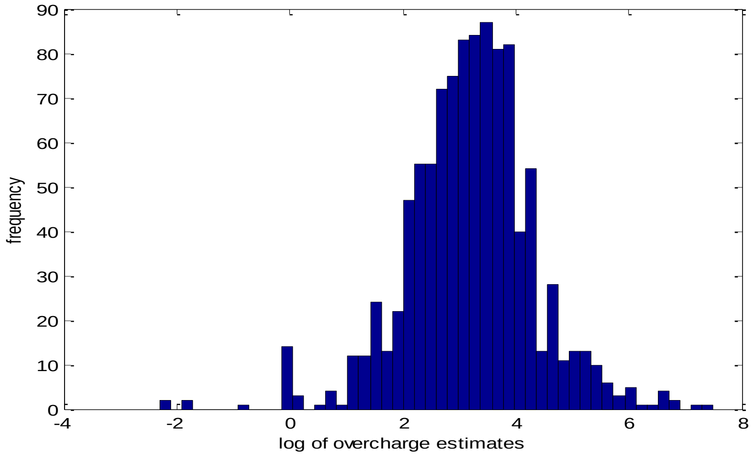
where $\hat{\theta}_i \in (0, \theta]$, with $\theta \in$ [25%, 70%], and $\widehat{\imath mr}_i$ is the inverse Mills ratio (IMR). The latter variable is calculated for each observation using the output of a preliminary Probit model fitted to the indicator of $\hat{\theta}_i \in (0, \theta]$ (i.e., the selection variable). Hence, we first estimate a Probit for each value of $\theta \in$ [25%, 70%][16] using the whole sample; next, we compute $\widehat{\imath mr}_i$ for each observation lying in the range $(0, \theta]$; finally, we estimate Equation (8) using the subsample $(0,\theta]$. This is the well-known Heckit procedure of Heckman (1979), where an IMR variable computed from a first step Probit is included in the second step estimation in order to control for selection biases that would arise from the right truncation of the sample at θ. Details on the Heckit procedure and the calculation of the IMR are given in Appendix D.

We prefer the Heckit to a right censored Tobit because the former is less restrictive than the latter. Moreover, the Tobit model assumes that the regressand (here, OE) has not been measured with systematic biases while our goal is precisely to estimate and remove the bias contaminating OE. We focus on modeling the log of OE because their distribution is more symmetric, as shown by Figure 5.

[15] We choose not to go beyond θ =70% as our previous analysis concluded that the overall data quality is low on the range OE>65%.
[16] That is, 45 different values of θ in total.

Figure 5

Logarithm of positive OE (92.8% of the sample)



The Heckit procedure requires that some regressors be included in the first step Probit from which the IMR is estimated and excluded from the second step regression. Such regressors are called exclusion variables and they ensure the identification of the parameters estimated at the second step. An ideal exclusion variable is a determinant of the probability of the OE belonging to the range $(0, \theta]$ that has no direct influence on the OE. Unfortunately, none of the regressors available to us is eligible for this role on theoretical grounds.

To circumvent this difficulty, we consider shrinking the information set repesented by the bias factors (Z) in the second step estimation. More precisely, we use Y and W along with the Z variables as regressors in the first step probit. In the second step regression, the estimation methods "historical case studies", "legal decisions" and "Other" are merged into a single category. The type of publications "journals" and "working papers" are merged under one group, "book chapters" and "monographs" under a second group, and "Newspapers” and "Speech" under a third group. All other categories are kept unchanged. This shrinkage of the information set plays as exclusion restrictions.

The endogeneity of W

The variable W_i which indicates whether the cartel case has been resolved with a guilty plea or decision is likely endogenous. Indeed, the decision to plead guilty is potentially related to the existence and size of the overcharge: A guilty plea may be suggestive that a firm has been involved in an effective price-fixing collusion. And similarly if the firm was found guilty.

We consider estimating two second step regressions, one in which W_i is included and another one in which it is instrumented. The first regression is predictive while the second is more structural. The structural regression is aimed at estimating the coefficients of the regressors without bias, while the predictive model targets point forecasts of the regressand.

In the structural approach, our instrumental regression consists of a Probit model where the probability of a guilty plea is conditioned by the Y-variables. The endogeneity problem is addressed by replacing W_i by the Probit prediction of Pr $(W_i = 1|Y_i)$. Further exclusion variables are in principle needed in the structural approach while none is available to us. Fortunately, singularity is avoided because the probability of a guilty plea is nonlinear in the included instruments.

The estimated coefficients obtained from Equation (8) are used to predict bias-corrected OE conditional on $\hat{\theta}_i \in (0, \theta]$:

$$\hat{\theta}_{bc1,i}(\theta) = \exp\left(\hat{\beta}_0 + Y_i\hat{\beta}_1 + W_i\hat{\beta}_2 + \widehat{\imath mr}_i\hat{\beta}_4 + \frac{\hat{\sigma}_e^2}{2}\right), \hat{\theta}_i \in (0, \theta] \quad (9)$$

We also compute bias-corrected estimates unconditionally[17] by removing the IMR and using the same parameter estimates as above:

$$\hat{\theta}_{bc2,i}(\theta) = \exp\left(\hat{\beta}_0 + Y_i\hat{\beta}_1 + W_i\hat{\beta}_2 + \frac{\hat{\sigma}_e^2}{2}\right), \hat{\theta}_i > 0 \quad (10)$$

The Probit from which the IMR is inferred is estimated with the subsample of cartels with strictly positive OE (1038 cartels), $\hat{\theta}_{bc1,i}(\theta)$ is computed for cartels with OE lying in the range (0,θ] while $\hat{\theta}_{bc2,i}(\theta)$ is computed for the subsample of successful cartels ($\hat{\theta} > 0$).

Given that the subsample used for estimation is truncated from the right, the Heckit should predict a larger average bias-corrected OE on the whole sample than in the subsample, that is:

$$\bar{\theta}_{bc1}(\theta) = \frac{1}{n(\theta)}\sum_{\hat{\theta}_i \in [0,\theta]} \hat{\theta}_{bc1,i}(\theta) \leq \frac{1}{n}\sum_{i=1}^{n} \hat{\theta}_{bc2,i}(\theta) = \bar{\theta}_{bc2}(\theta) \quad (11)$$

where $n(\theta)$ is the size of the subsample used for estimation. Furthermore, $\bar{\theta}_{bc1}(\theta)$ should be a non decreasing function of θ over the valid range of overcharges while $\bar{\theta}_{bc2}(\theta)$ should be approximately flat.

Violation of either of these rules is suggestive of the presence of biases in the OE data that our procedure failed to completely remove. Figures 6 and 7 plot $\bar{\theta}_{bc1}(\theta)$ and $\bar{\theta}_{bc2}(\theta)$ on the y-axis against θ on the x-axis.[18]

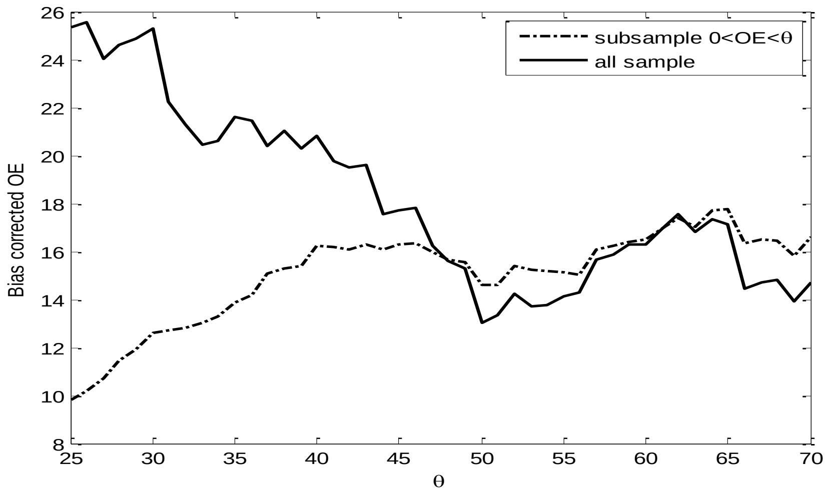
[17] That is, unconditionally on the truncation $\hat{\theta}_i \in (0, \theta]$. Indeed, the presence of the IMR makes the predictions of Equation (9) specific to the subsample used for estimation. This becomes obvious by noting that the logarithm of Equation (9) is the fitted part of Equation (8). Zero overcharges are excluded from the sample. Therefore, the predictions are conditional on $\hat{\theta}_i > 0$ as well.

[18] $\bar{\theta}_{bc2}(\theta)$ should be flat as it is an estimate of the average bias-corrected OE of all cartels based on the subsample of cartels with OE lying in the range (0,θ]. However, the representativeness of this subsample is not warranted when θ is small. This explains why the curve $\bar{\theta}_{bc2}(\theta)$ is decreasing at first. As for $\bar{\theta}_{bc1}(\theta)$, it is an estimate of the average bias-corrected OE for cartels with raw OE lying in an increasing range, namely (0,θ].

Figure 6 shows the case where W is used as regressor (predictive approach) while Figure 7 is for the case where W is instrumented (structural approach). On both figures, $\bar{\theta}_{bc2}(\theta)$ is overall decreasing on $\theta \leq 49\%$ and weakly increasing on $\theta > 49\%$. If the bias contaminating the raw OE had been completely removed, the curve of $\bar{\theta}_{bc2}(\theta)$ should be flat at least on the second half of the support of θ. This curve is flatter on Figure 7 than on Figure 6, which suggests that part of the problem is due to the endogeneity of W.

Figure 6

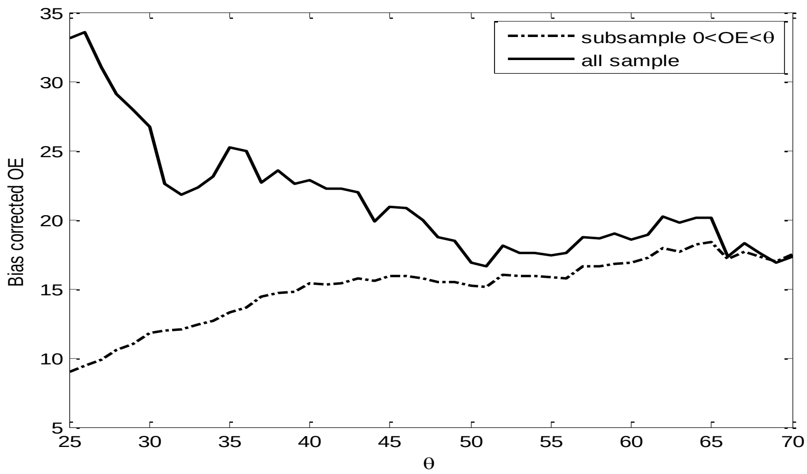
Average bias-corrected OE predicted for the subsample $\hat{\theta}_i \in [0, \theta]$ by the **predictive** model



Note: Dash-dotted line: average bias-corrected OE $\bar{\theta}_{bc1}(\theta)$ on the subsample used for estimation. Solid line: average bias-corrected OE $\bar{\theta}_{bc2}(\theta)$ on the whole sample

Figure 7

Average bias-corrected OE predicted for the subsample $\hat{\theta}_i \in [0, \theta]$ by the **structural** model



Note: Dash-dotted line: average bias-corrected OE $\bar{\theta}_{bc1}(\theta)$ on the subsample used for estimation. Solid line: average bias-corrected OE $\bar{\theta}_{bc2}(\theta)$ on the whole sample

In our empirical framework, the sample selection bias arising from the right truncation of the OE is controlled by the IMR, while the original bias contaminating the raw OE is corrected by Z. The fact that $\bar{\theta}_{bc2}(\theta)$ is increasing on $\theta > 49\%$ does not necessarily mean that the correction for sample selection is ineffective. Indeed, the same pattern would be observed if Z is less and less effective at capturing the initial bias contaminating the OE as the truncation threshold increases away from 49%.

As expected, $\bar{\theta}_{bc1}(\theta)$ is overall increasing in θ. On Figure 6 however, the curves of $\bar{\theta}_{bc1}(\theta)$ and $\bar{\theta}_{bc2}(\theta)$ are close and intertwined on the domain $\theta > 49\%$. This indicates that the Heckit procedure successfully corrects the sample selection bias.[19] This suggests that the least distorted Heckit model is the one estimated with $\theta = 49\%$.

On Figure 7, $\bar{\theta}_{bc1}(\theta)$ remains strictly below $\bar{\theta}_{bc2}(\theta)$. Furthermore, the curve of $\bar{\theta}_{bc1}(\theta)$ is smooth and monotonic to a greater extent on Figure 7 than on Figure 6. This again is suggestive that the endogeneity correction matters.

7. The Determinants of Cartel Overcharge

The results of Section 5 lead us to restrict the analysis of Section 6 to cartels with OE lower than 65%. The results of Section 6 further lead us to restrict the analysis to cartels with OE lower than 49%. The current section presents the estimation results when the sample is truncated at 49%. Appendix A presents summary statistics for the subsamples of alleged cartels with OE=0, 0<OE≤49% and OE>49%. Table 5 shows the Probit estimation results for the probability of a cartel being successful at raising its price above the competitive level.[20] This Probit is trustworthy given our previous finding that $\Pr(\hat{\theta}_i > \theta | Y_i, W_i)$ and $\Pr(\hat{\theta}_i > \theta | Y_i, W_i, Z_i)$ agree to a great extent at $\theta = 0$.

There is a positive link between the duration of a cartel and its probability of being successful. Domestic cartels tend to be less successful than international cartels while bid-rigging tends to be more successful than other forms of cartels. Cartels resolved with a guilty plea have a higher probability of being succesful than the other cartels. The geographical location and time period seem to have no effect on the probability of a cartel being successfull.

[19] As an estimator of the average overcharge of all cartels, $\bar{\theta}_{bc1}(\theta)$ is biased (due to sample selection) while $\bar{\theta}_{bc2}(\theta)$ is expected to be unbiased. However, the sample selection bias decreases naturally as the threshold θ increases. Therefore, the fact that $\bar{\theta}_{bc1}(\theta)$ converges to $\bar{\theta}_{bc2}(\theta)$ as θ increases suggests that the amount of sample selection bias removed by means of the IMR is reasonably accurate.

[20] Note that a cartel can fail at raising its price while being effective along other competitive dimensions (e.g., entry, capacity, innovation, advertizing, credit terms, etc.). The database available to us does not permit to address such issues.

Table 5

Probit model for the probability of a cartel being successful: $\Pr(\hat{\theta}_i > 0 | Y_i, W_i, Z_i)$

	Marginal Effects at the Median
Constant	0.14**
Duration	0.01**
Domestic	-0.03*
BidRig	0.08***
Guilty	0.04**
US	-0.03
EU	0.00
ASIA	-0.01
ROW	-0.03
P1	1.63
P3	-0.05
P4	0.01
P5	-0.04
P6	-0.02
US x P1	0.13
US x P2	0.06
US x P3	-0.03
OTHER	-0.09***
HISTOR	-0.18***
PBEFOR	0.02
PWAR	0.00
PAFTER	0.01
COST	0.00
YARDST	0.02
LEGAL	0.07*
JOURNAL	-0.04
MONOGR	0.05
EDBOOK	0.01
COURT	0.01
NEWSPAPER	1.65
WORKP	0.07*
SPEECH	1.67

Note: Estimation is done using the whole sample (1119 cartels). The coefficients of P1, NEWSPAPER and SPEECH have not converged. However, estimation results are similar when these dummies are removed or merged with other groups.

A higher proportion of zero OE is obtained via historical case studies with no method specified than in studies that specify an estimation method. The proportion of strictly positive OE obtained via legal decisions is higher than for other estimation methods. Finally, working papers contain a significantly higher proportion of strictly positive OE than other publication sources. The coefficients of the dummy variables P1, NEWSPAPER and SPEECH have not converged.[21] However, the estimation results are qualitatively similar when these dummies are removed.

Of all successful cartels, which ones are able to overcharge by more than 49%? In Table 6, we attempt to provide an answer by estimating a Probit model for $\Pr(\hat{\theta}_i > 49|Y_i, W_i, Z_i)$. The sample used for the estimation is restricted to the cartels with strictly positive OE. We note that domestic cartels and bid rigging cases are associated with a lower probability of an OE being more than 49%. The "Guilty" dummy (W) does not seem to be an important determinant of the probability of $\hat{\theta}_i > 49$, contrasting with what is found in Table 5 for $\Pr(\hat{\theta}_i > 0|Y_i, W_i, Z_i)$. Hence, pleading guilty is suggestive that the cartel has been effective, but not that the overcharge is above 49%.

The probability of overcharging by more than 49% is higher for US cartels (particularly, US cartels of Period P3) and ASIA cartels compared to the reference group WORLD. We also note that the Z-variables are significantly correlated with the indicator of $\hat{\theta}_i > 49$. Indeed, OE obtained via econometric methods and those published in working papers and academic journals have relatively lower probabilities of being more than 49%.

Table 7 presents the estimated coefficients of the meta-regressions. The dependent variable in these regressions is the log of OE. Therefore, if the coefficient of a RHS dummy variable is β_i, the average OE for the group where the dummy variable takes the value 1 is $exp(\beta_i)$ times the average OE of the reference group. This represents a percentage increase or decrease of $exp(\beta_i)-1$. As the OE is already expressed in percentage, we will read the results in terms of factor of increase or decrease in order to avoid confusions.

The predictive regression suggests that the ability of a cartel at raising its price increases with its duration. On average, the overcharge increases by a factor of *exp(0.06)=1.06* per quinquennium. Domestic cartels overcharge by a factor of *exp(-0.3)=0.74* less than international cartels. Cartels resolved with a guilty plea overcharge by a factor of *exp(0.27)=1.31* more than the other cartels. Cartels of the EU overcharge by *exp(-0.23)=0.79* less than cartels from other geographical location. OE obtained via a "Price before", "price war" or "yardstick" method are on average larger than those obtained by other estimation methods. Also, OE published in monographs, edited books, newspaper and speeches are on average higher than those published in other media.

In the structural model, the coefficient of the "Guilty" dummy variable is not significant. This suggest that a guilty plea has no causal effect on overcharges even though it has a predictive effect. In retrospect, this result is quite intuitive: a firm adopts a guilty plea strategy because it has

[21] Probit coefficients are expected to be reasonably small and not explosive (i.e., generally, but not necessarily, smaller than 1). In the Probit of Table 5, the estimated coefficients of NEWSPAPER and SPEECH are quite large (approx, =15) although the marginal effects are of reasonable magnitude.

been involved in a successful cartel (i.e., "positive overcharge" causes "guilty plea") and not the converse.

The R-square of the log-linear regression is slightly higher for the predictive model (0.09) than for the structural model (0.08). This is not surprising as the structural model is aimed at achieving an unbiased estimation of the parameters used to bias-correct the raw OE while the predictive model delivers the best fit of the OE in terms of in-sample mean square error.

Table 6

Probit model for $\Pr\left(\hat{\theta}_i > 49 | Y_i, W_i, Z_i\right)$

	Marginal Effects at the Median
Constant	-0.06
Duration	0.01
Domestic	-0.26***
BidRig	-0.11**
Guilty	-0.02
US	0.11**
EU	0.06
ASIA	0.22***
ROW	-0.14
P1	0.35
P3	-0.16
P4	-0.10
P5	-0.06
P6	-0.14
US x P1	-5.53
US x P2	0.04
US x P3	0.25**
OTHER	0.21***
HISTOR	0.20
PBEFOR	0.10*
PWAR	0.24**
PAFTER	0.13**
COST	0.16**
YARDST	0.18***
LEGAL	0.03
JOURNAL	-0.13*
MONOGR	-0.11
EDBOOK	-0.03
COURT	-0.11
NEWSPAPER	0.10
WORKP	-0.24***
SPEECH	-0.35

Note: Overcharge estimates that are equal to zero (81 observations) are excluded. This leaves us with 1038 cartels with strictly positive overcharge. To estimate the Probit models, we generate the dummy indicator of $\hat{\theta}_i > 49\%$. The length of that dummy is 1038; it contains 782 "zeros" and 256 "ones".

Table 7

Meta-Analysis of cartel OE

	Predictive regression	**Structural** regression
Constant	2.30***	2.16***
Duration	0.06*	0.06
Domestic	-0.30*	-0.44*
BidRig	-0.04	-0.03
Guilty	**0.27*** **	**-0.03**
US	-0.12	-0.06
EU	-0.23**	-0.21*
ASIA	0.10	0.15
ROW	-0.14	-0.24
P1	-0.25	0.07
P3	-0.22	-0.19
P4	-0.25	-0.20
P5	-0.02	0.07
P6	-0.18	-0.08
US x P1	-1.29	-4.59
US x P2	0.09	0.23
US x P3	0.42	0.41
IMR	**0.13**	**0.32**
OTHER+HISTOR+LEGAL	0.01	0.08
PBEFOR	0.25**	0.32***
PWAR	0.49*	0.60**
PAFTER	0.16	0.25*
COST	0.14	0.21
YARDST	0.32*	0.41***
JOURNAL+WORKP	0.16	0.11
MONOGR+EDBOOK	0.46***	0.41*
COURT	0.16	0.18
NEWSPAPER+SPEECH	0.64**	0.64**
R-square	0.09	0.08

Note: Estimation is done with the subsample $0\% < \hat{\theta}_i \leq 49\%$ (782 cartels). An IMR obtained from Table 6 is included to control for the right truncation of the sample. Therefore, the estimation results concern effective cartels only. The predictive model is a Heckit that ignores the endogeneity of "Guilty". In the structural model, "Guilty" is instrumented.

Connor and Bolotova (2006) performed a meta-analysis of cartel OE in which they modelled the OE as a linear function of Y and Z.

$$\hat{\theta}_i = \beta_0 + Y_i\beta_1 + W_i\beta_2 + Z_i\beta_3 + \varepsilon_i \quad (12)$$

They estimated different restrictions of the full model. For their full model (column [7] of Table 6 in their paper), they found that the OE is positively related to the duration, but does not depend on whether the firm is "guilty" or not; it is lower for domestic cartels and for cartels that have operated in the EU; it is neither higher nor lower for bid-rigging cases contrary to what is claimed by Cohen and Scheffman (1989); and that the size of overcharges has declined over time. Connor and Bolotova attributed the latter result to the increased severity of antitrust regulation.

Interestingly, they also found that the Z variables have significant impacts on OE. For example, they found the "yardstick" method to produce estimates that are at least 10% higher than the "after the conspiracy" method. For the publication sources, they found that "government reports" and "court reports" produce estimates that are respectively 22% lower and 15% higher than "monograph or book". The fact that the Z variables show significant effects in the regression suggests that the raw OE are indeed biased.

Our results differ in significant ways from theirs. First, we find that cartels who pleaded guilty have higher overcharge on average, but this effect is not causal. Second, we find that more recent cartels (periods P5 and P6) are not really different from cartels of previous periods regarding average pricing behavior. The antitrust law regime have no impact on the probability of an overcharge being positive. However, it has an impact on the distribution of positive overcharges. More precisely, US cartels of the period P3 have a higher probability of overcharging by more than 49% than other cartels in other periods.

8. Bias-correcting the OE

The coefficients estimated from the predictive regression (Table 7) are possibly distorted by the endogeneity of "Guilty". Therefore, these coefficients should not be used to predict bias-corrected OE. The coefficients estimated from the structural regression (Table 7) are expected to be unbiased and may therefore be used to predict bias-corrected OE. However, some information can still be gleaned from the residuals of the latter regression as it is potentially correlated with "Guilty".

Our ultimate objective, which is to obtain good predictions of average bias-corrected OE, requires that we first remove the causal effect of the Z-variables from the raw OE. Once this step is completed, a predictive regression of the ″cleaned″ OE onto Y can be used to estimate bias-corrected conditional means of OE. This strategy which combines the strenghs of the structural and reduced-form approaches is presented below.

To begin, we estimate the second step regression by excluding the "Guilty" dummy variable (W). The estimated equation on (0%, 49%] is:

$$\log\hat{\theta}_i = \hat{\beta}_0 + Y_i\hat{\beta}_1 + Z_i\hat{\beta}_3 + \widehat{\imath mr}_i\hat{\beta}_4 + e_i, \qquad (13)$$

This step is aimed at obtaining coefficient estimates that are not distorted by the endogeneity of W_i. The exclusion of W is justified by our previous finding that this variable has no causal effect on the overcharge. Next, we infer bias-corrected OE in (0%, 49%] by removing the contribution of the Z variables, that is:

$$\log\hat{\theta}_{bc,i} = \log\hat{\theta}_i - Z_i\hat{\beta}_3 \qquad (14)$$

Finally, we estimate a predictive regression of $\log\hat{\theta}_{bc,i}$ onto Y_i, W_i and $\widehat{\imath mr}_i$:

$$\log\widehat{\theta}_{bc,i} = \widehat{\beta}_0 + Y_i\widehat{\beta}_1 + W_i\widehat{\beta}_2 + \widehat{\imath mr}_i\widehat{\beta}_4 + e_i \qquad (15)$$

The variable W is included in regression (15) for the purpose of exploiting its preditive power. Equation (13) is a causal regression while Equation (15) is a predictive regression. Table 8 shows the estimated coefficients for Equation (15).

Table 8

Predictive Model of Bias-corrected OE

	Coefficients
Constant	2.57***
Duration	0.06**
Domestic	-0.28**
BidRig	-0.03
Guilty	0.24***
US	-0.13
EU	-0.25***
ASIA	0.05
ROW	-0.13
P1	-0.30
P3	-0.21
P4	-0.25
P5	-0.01
P6	-0.19
US x P1	0.62
US x P2	0.07
US x P3	0.42
IMR	-0.18
R-square	0.06

Note: Estimation output for Equation (15) using $0\% < \widehat{\theta}_i \leq 49\%$ (782 cartels). The estimation results concern effective cartels only.

We see that the "Guilty" dummy variable has significant predictive power on bias-corrected OE. This predictive power was missed by the structural regression shown in Table 7. The negative coefficient of the "domestic" dummy variable (-0.28, significant) is stronger than in the predictive model (-0.20, non significant) but weaker than in the structural model (-0.37, significant). Cartels that operate in the EU have lower overcharge than those from other geographical markets.

The bias-corrected OE of effective cartels are inferred by using the parameters estimates of Equation (15) into Formulas (10). First, one estimates Equation (13) on the subsample (0%, 49%] using the logarithm of the raw OE as regressand. Second, one uses Equation (14) to obtain bias-corrected OE for each cartel. Third, one estimates Equation (15) on the subsample (0%, 49%] using the log of bias-corrected OE as regressand. Finally, one uses Equation (10) to predict bias-corrected overcharges for all cartel with a positive overcharge conditional on Y and W.

Our previous analysis suggested that the raw OE data permits to identify an effective cartel from an ineffective one. Therefore, the initial 0% OE are assumed exempt of bias and left unchanged.

Table 9 replicates Table 1 with bias-corrected OE as input. For the subsample with initial estimates lying in the range (0%, 49%], we find a mean overcharge of 16.47% with a median of 16.17%. For the subsample of strictly positive OE, the mean is 16.68% while the median is 16.17%. For the whole sample (including the zeros), the predicted mean is 15.47% while the median is 16.01%.[22]

Table 9
Means and medians of Bias-corrected OE per location and types of cartels.

		All Cartels	OE > 0%	0%<OE≤49%	OE>49%	Cartels before 1973	Cartels after 1973
All locations	Mean	15.47	16.68	16.47	17.31	13.97	16.07
	Median	16.01	16.17	16.17	16.48	14.18	16.41
	prop.	100.00	92.77	69.88	22.87	28.51	71.49
US	Mean	14.36	15.82	15.69	16.30	15.13	14.00
	Median	14.48	15.19	15.04	16.13	16.37	14.20
	prop.	30.02	27.25	21.36	5.90	9.65	20.37
EU	Mean	13.51	14.43	14.05	15.39	12.54	14.15
	Median	14.08	14.20	13.67	15.32	13.18	15.86
	prop.	33.51	31.36	22.52	8.85	13.40	20.11
Domestic	Mean	12.93	14.21	14.09	14.86	13.08	12.87
	Median	13.68	13.81	13.79	14.09	13.39	13.81
	prop.	46.82	42.62	36.01	6.61	14.47	32.35
International	Mean	17.71	18.78	19.01	18.30	14.89	18.71
	Median	18.66	19.29	19.38	18.26	15.34	20.66
	prop.	53.17	50.13	33.87	16.26	14.03	39.14

Note: For comparison purposes, we use the same column headings as in Table 1 that presents the raw OE. The proportions (%) are fractions of the total Connor sample (1119 cartels)

For US cartels, we find a mean bias-corrected OE of 15,69% (with a median of 15.04%) for the subsample used for estimation and 14.36% (with a median of 14.48%) for the whole sample. For EU cartels, the corresponding figures are 14.05% (13.67%) and 13.51% (14.08%). Moreover, we find a mean bias-corrected OE of 17.71% (with a median of 18.66%) for international cartels and 12.93% (with a median of 13.68%) for domestic cartels. Finally, we find that post-1973 cartels achieved higher bias-corrected mean overcharge (16.07%) than pre-1973 cartels (13.97%). Overall, the means and medians bias-corrected OE shown in Table 9 suggest a more homogenous behaviour of cartels than the means and medians raw OE shown in Table 1.

[22] NOTE: these percentages are significantly lower than those reported at page 42 of our article Allain et alii (2015), published in *International Review of Law and Economics* 42, 38-47, which were due to a regrettable error in our MATLAB code.

Table 10 presents the *mean* bias-corrected OE for different categories of cartels according to whether they are domestic or international, in bid-rigging cases of not, and/or were found or pleaded guilty or not. Table 11 presents *median* bias-corrected OE for the same subgroups. The differences between the raw OE (Table 1 and the left-hand side of Tables 12 and 13) and the bias-corrected ones (Table 9 and the right-hand side of Tables 12 and 13) are quite striking. In several cases, the bias-corrected OE is at least twice smaller than the raw OE. Our results support that cartels are overall more similar than what the raw OE data suggest.

Table 10
Raw versus Bias-corrected *mean* OE.
All numbers are expressed in percentage (%)
"n.a." means that the corresponding category of cartels is not represented in the subsample of interest.

Cartel characteristics			Raw Average OE					Bias-Corrected Average OE				
domestic	bidrig	Guilty	US	EU	ASIA	ROW	WORLD	US	EU	ASIA	ROW	WORLD
Subsample with OE>0%												
Yes	Yes	Yes	27.61	18.41	29.47	17.40	18.70	15.04	12.91	16.64	13.77	17.04
No	Yes	No	n.a.	430.00	n.a.	n.a.	n.a.	n.a.	12.30	n.a.	n.a.	n.a.
Yes	No	No	50.62	25.79	89.33	18.67	n.a.	14.24	11.50	14.14	10.40	n.a.
No	No	Yes	29.51	53.30	37.60	28.57	72.27	19.57	17.06	21.94	19.17	21.72
Yes	Yes	No	21.18	16.67	4.88	14.50	n.a.	11.35	9.75	12.54	12.46	n.a.
No	Yes	Yes	38.05	34.27	35.40	28.73	20.18	18.16	16.43	21.20	19.06	23.82
Yes	No	Yes	52.54	20.92	57.64	11.41	n.a.	16.03	13.01	16.89	13.94	n.a.
No	No	No	59.59	76.71	29.00	50.00	43.48	17.86	14.31	17.11	14.31	18.20
Whole sample												
Yes	Yes	Yes	26.73	17.38	29.47	17.40	18.70	14.56	12.20	16.64	13.77	17.04
No	Yes	No	n.a.	430.00	n.a.	n.a.	n.a.	n.a.	12.30	n.a.	n.a.	n.a.
Yes	No	No	46.72	20.15	77.15	17.12	n.a.	13.14	8.99	12.21	9.54	n.a.
No	No	Yes	27.67	52.78	37.60	24.49	71.89	18.34	16.90	21.94	16.43	21.61
Yes	Yes	No	21.18	16.67	4.88	14.50	n.a.	11.35	9.75	12.54	12.46	n.a.
No	Yes	Yes	38.05	32.89	35.40	28.73	20.18	18.16	15.77	21.20	19.06	23.82
Yes	No	Yes	44.51	20.52	51.24	11.41	n.a.	13.58	12.76	15.01	13.94	n.a.
No	No	No	46.56	72.20	29.00	50.00	34.29	13.95	13.47	17.11	14.31	14.35

Table 11

Raw versus Bias-corrected *median* OE.

All numbers are expressed in percentage (%)

"n.a." means that the corresponding category of cartels is not represented in the subsample of interest.

Cartel characteristics			Raw Median Estimates					Bias-Corrected Median Estimates				
domestic	bidrig	Guilty	US	EU	ASIA	ROW	WORLD	US	EU	ASIA	ROW	WORLD
Subsample with OE>0%												
Yes	Yes	Yes	18.30	12.40	29.00	17.40	18.70	14.20	12.55	16.00	13.77	17.04
No	Yes	No	n.a.	430.00	n.a.	n.a.	n.a.	n.a.	12.30	n.a.	n.a.	n.a.
Yes	No	No	26.05	18.40	30.00	17.50	n.a.	13.98	11.41	14.49	10.18	n.a.
No	No	Yes	31.80	29.00	17.50	24.50	29.60	19.38	16.17	21.83	18.99	21.59
Yes	Yes	No	14.50	17.00	4.83	14.50	n.a.	10.82	9.47	12.54	12.46	n.a.
No	Yes	Yes	25.75	21.00	29.80	30.55	13.45	17.76	16.16	21.20	18.78	23.27
Yes	No	Yes	16.80	16.00	24.55	10.00	n.a.	15.49	12.20	16.48	13.78	n.a.
No	No	No	50.00	50.00	29.00	50.00	31.25	18.52	14.20	17.11	14.31	17.58
Whole sample												
Yes	Yes	Yes	18.05	12.15	29.00	17.40	18.70	14.20	12.55	16.00	13.77	17.04
No	Yes	No	n.a.	430.00	n.a.	n.a.	n.a.	n.a.	12.30	n.a.	n.a.	n.a.
Yes	No	No	24.70	13.50	24.50	13.60	n.a.	13.67	11.36	13.68	10.18	n.a.
No	No	Yes	27.80	29.00	17.50	21.65	29.48	19.38	16.17	21.83	18.99	21.59
Yes	Yes	No	14.50	17.00	4.83	14.50	n.a.	10.82	9.47	12.54	12.46	n.a.
No	Yes	Yes	25.75	17.00	29.80	30.55	13.45	17.76	15.70	21.20	18.78	23.27
Yes	No	Yes	14.90	15.95	23.25	10.00	n.a.	13.81	12.20	16.48	13.78	n.a.
No	No	No	36.50	50.00	29.00	50.00	25.00	15.17	14.20	17.11	14.31	16.25

9. Analysis of Variance for the OE Bias

In this section, we attempt to understand which of the Z variables causes more bias in the raw OE. For that purpose, we define the bias as the difference between the log of the raw OE and the log of the bias-corrected OE.

$$\hat{\Delta}_i = \log\hat{\theta}_i - \hat{\omega}_i,$$

where $\hat{\omega}_i$ is the fitted value of Equation (15). We regressed $\hat{\Delta}_i$ on a constant and all Z dummy variables, keeping "econometric method" and "government report" as reference groups.

We run separate regressions for effective cartels, cartels with OE≤49% and cartels with OE>49% (see Table 12). The percentage of explained variance is 8% for all effective cartels, 5% for cartels with OE≤49% and 16% for cartels with OE>49%. Thus, the R-square triples as we move from the subsample with OE≤49% to the one with OE>49%. This supports our previous finding that the raw OE lying above 49% are substantially more biased than the remainder of the sample.

Table 12

		Effective Cartels (OE>0)	Cartels with 0<OE<=49%	Cartels with OE>49%
Estimation methods	Constant	0.36**	-0.44***	2.18***
	OTHER	0.19	-0.07	-0.42*
	HISTOR	-0.19	-0.42	-0.28
	PBEFOR	0.33***	0.23**	-0.10
	PWAR	0.60**	0.42*	-0.19
	PAFTER	0.28**	0.15	-0.10
	COST	0.38**	0.12	0.20
	YARDST	0.65***	0.28***	0.35*
	LEGAL	0.05	0.05	0.03
Publication media	JOURNAL	-0.36**	0.23	-0.58***
	MONOGR	0.00	0.47***	-0.40**
	EDBOOK	0.23	0.52***	-0.55***
	COURT	-0.29*	0.21	-0.45***
	NEWSPAPER	0.19	0.69*	-0.83**
	WORKP	-0.47***	0.21	-0.56***
	SPEECH	-0.12	0.71*	-0.91
R-square		0.08	0.05	0.16
Sum of Squared Residuals		1194.55	543.29	110.28
Sample size		1038	782	256

Note: The dependent variable is the log-bias defined as $\hat{\Delta}_i = \log\hat{\theta}_i - \hat{\omega}_i$, where $\hat{\omega}_i$ is the fitted value of (15). The R-square doubles are we move from the subsample of all effective cartels to the subsample of cartels with OE>49%. Hence, the explanatory power of the Z variables is higher on the range OE>49%.

We perform a Chow test for the stability of the coefficients across the two subsamples. The test statistic is given by:

$$F = \frac{(1194.55 - 543.29 - 110.28)/16}{(543.29 + 110.28)/(1038 - 2 \times 16)} = 52.04$$

Under the null hypothesis that coefficients are the same on the two subsamples, the F statistic is a Fisher random variable with (16,1006) degrees of freedom. At level 1%, the critical value of the Fisher distribution with (16,1006) degrees of freedom is 1.9832. Hence, the null hypothesis is overwhelmingly rejected.

On the subsample of all effective cartels, all estimation methods (except HISTOR, LEGAL and OTHER) are biased upward relatively to the ECON method; the YARDST and PWAR methods seem to be the most important sources of positive bias on that subsample; OE published in government reports (GOVREP) are biased upward relatively to those published in academic journals, court decisions and working papers.

On the subsampe of cartels with OE>49%, OTHER is less biased than ECON while YARDST is more biased; all other methods entail similar bias as the ECON method; OE published in government reports are biased upward relatively to those published in all other medias.

On the subsample of cartels with OE≤49%, PBEFOR, PWAR and YARDST are more biased than ECON; OE published in government reports are less biased relatively to those published in all other medias.

YARDST is the only variables whose effect is significant and of the same sign on both subsamples. For all other variables, the effect is either significant on one subsample only or of opposite signs on the two subsamples.

10. Conclusion

Our study identifies the mean and median overcharges of cartels by performing a meta-analysis on an extended version of the database used in Connor (2010). Each observation in the sample is a potentially biased overcharge estimate (OE) obtained from a previous study. Three groups of variables describe the observations. The first group consists of variables (Y) that explain the true overcharge. The second group (Z) consists of factors that capture potential estimation biases. The third group contains a single variable (W) which indicates whether the alleged cartel pleaded or was found guilty. The latter variable is found to be endogenously related to the true overcharge. Our study bias-corrects the raw OE by cleaning them from the contribution of Z variables.

In order to assess the quality of the data, we used a Kullback-Leibler divergence to compare the probability of an OE being larger than θ conditional on (Y,W) to the same probability conditional on (Y,W,Z). The divergence between the two probabilities increases slowly on the range $\theta \in [0\%, 65\%]$, but a large jump occurs at $\theta = 65\%$. Although the results suggest the presence of bias in the raw OEs, this jump appears to be driven by other data problems.

We pursued our empirical investigations by estimating Heckit models for the log of OE on subsamples of type $(0, \theta]$. If the OE data were unbiased, the average bias-corrected OE for the subsample of cartels with $OE \in (0, \theta]$ should be increasing in θ and lower than the average bias-corrected OE inferred for all successful cartels. Also, the curve of the average bias-corrected OE of all successful cartels should be flat in the upper range of θ. We find the latter condition to be violated. The curve of the average bias-corrected OE of all successful cartels is decreasing on $\theta \leq 49\%$ and increasing on $\theta > 49\%$.

Acting on these results, we estimate our final meta-analysis model on the subsample of $OE \in (0\%, 49\%]$. We employ a Heckit procedure to infer the bias-corrected OE of all effective cartels (strictly positive OE). The raw OE that are equal to zero are included back unaltered in the sample used for prediction. Our meta-analysis delivers mean and median bias-corrected OE of 16.47% and 16.17% for the subsample with initial estimates lying in the range (0%, 49%], of 16.68% and 16.17% for the subsample of effective cartels, and of 15.47% and 16.01% for the whole sample.

Our results have significant implications for antitrust policy. Indeed, a major element in the prosecution of cartel is their capacity to exert upward pressures on prices. Becker (1968) and Landes (1983) examined the link between the cartel overcharge and the fine in a static game framework. Both authors concur that the optimal fine is equal to the illegal profit of the cartel divided by the probability of detection.

Allain et alii (2011) argued that the Becker-Landes rule must be interpreted with caution in a dynamic framework. They show that the optimal fine can be computed as either the annual illegal profit divided by the annual probability of detection, or the cumulative illegal profit over the lifetime of the cartel divided by its lifetime probability of detection.

Allain et alii (2015) and Harrington (2014) considered infinitely repeated games where a threat of deviation by a cartel member exists. Antitrust authorities may make the deviation profitable by granting partial or full leniency to the whistleblower. In such dynamic games, these authors show that the amount of the optimal fine is much lower than the correctly interpreted Becker-Landes rule suggests. Katsoulacos and Ulph (2013) conducted an analysis that accounts for the timing of antitrust authorities' decision and found that the optimal cartel fine is approximately 75% of the amount implied by the conventional formula.[23]

The mean and median bias-corrected OE obtained from our analysis have little to say about any specific cartel case where an overcharge is used as a measure of antitrust damage. The *true* overcharge in a given case depends on the specific set of facts regarding the challenged conduct, the structure of the industry to which the cartel belongs (e.g., sector, concentration, elasticity of demand), etc. In addition to the previous factors, the *estimated* overcharge depends on the availability and quality of data, the method used to estimate the but-for price, etc. Hence, the analysis conducted in this paper could be improved if more data on cartels become available.

[23] The link between the overcharge and the optimal fine is also studied by Nieberding (2006), Houba, Motchenkova and Wen (2013), Ginsburgh and Wright (2010) and Harrington (2005, 2010, 2014).

Appendix A: Summary statistics of the database

Table A1. Cartels with Raw OE in the interval (0%, 49%]

	mean	std	min	Max
OE	20.46	12.20	0,1	49
Duration	9.35	12.19	1	98
duration (discrete)	1.84	1.05	1	4
Domestic	0.52	0.50	0	1
BidRig	0.24	0.43	0	1
Guilty	0.71	0.45	0	1
US	0.31	0.46	0	1
EU	0.32	0.47	0	1
ASIA	0.09	0.28	0	1
ROW	0.04	0.20	0	1
WORLD	0.24	0.43	0	1
P1	0.01	0.09	0	1
P2	0.09	0.28	0	1
P3	0.10	0.30	0	1
P4	0.13	0.34	0	1
P5	0.62	0.49	0	1
P6	0.05	0.23	0	1
OTHER	0.05	0.21	0	1
HISTOR	0.01	0.08	0	1
PBEFOR	0.27	0.44	0	1
PWAR	0.02	0.13	0	1
PAFTER	0.13	0.34	0	1
COST	0.04	0.20	0	1
YARDST	0.13	0.34	0	1
ECON	0.16	0.37	0	1
LEGAL	0.19	0.39	0	1
JOURNAL	0.21	0.41	0	1
MONOGR	0.21	0.41	0	1
EDBOOK	0.05	0.22	0	1
GOVREP	0.25	0.43	0	1
COURT	0.01	0.08	0	1
NEWSPAPER	0.22	0.41	0	1
WORKP	0.01	0.08	0	1
SPEECH	0.05	0.21	0	1
Sample size		782		

Table A2. Cartels with Raw OE= 0% and OE >49%

	Cartels with Raw OE equal to 0%				Cartels with Raw OE above 49%			
	mean	std	min	Max	mean	Std	min	Max
OE	0.00	0.00	0	0	136.22	187.43	49,2	1800
Duration	7.16	5.38	1	26	9.62	12.33	1	109
duration (discrete)	1.86	1.10	1	4	1.92	1.08	1	4
Domestic	0.58	0.50	0	1	0.29	0.45	0	1
BidRig	0.07	0.26	0	1	0.11	0.31	0	1
Guilty	0.37	0.49	0	1	0.57	0.50	0	1
US	0.38	0.49	0	1	0.26	0.44	0	1
EU	0.30	0.46	0	1	0.39	0.49	0	1
ASIA	0.09	0.28	0	1	0.09	0.28	0	1
ROW	0.04	0.19	0	1	0.02	0.12	0	1
WORLD	0.20	0.40	0	1	0.25	0.44	0	1
P1	0.00	0.00	0	0	0.01	0.09	0	1
P2	0.19	0.39	0	1	0.13	0.34	0	1
P3	0.05	0.22	0	1	0.14	0.35	0	1
P4	0.25	0.43	0	1	0.16	0.37	0	1
P5	0.47	0.50	0	1	0.45	0.50	0	1
P6	0.05	0.22	0	1	0.11	0.31	0	1
OTHER	0.16	0.37	0	1	0.08	0.27	0	1
HISTOR	0.14	0.34	0	1	0.01	0.09	0	1
PBEFOR	0.20	0.40	0	1	0.29	0.45	0	1
PWAR	0.01	0.11	0	1	0.03	0.17	0	1
PAFTER	0.11	0.32	0	1	0.14	0.35	0	1
COST	0.04	0.19	0	1	0.06	0.24	0	1
YARDST	0.09	0.28	0	1	0.20	0.40	0	1
ECON	0.02	0.16	0	1	0.09	0.28	0	1
LEGAL	0.23	0.43	0	1	0.11	0.31	0	1
JOURNAL	0.44	0.50	0	1	0.16	0.37	0	1
MONOGR	0.21	0.41	0	1	0.30	0.46	0	1
EDBOOK	0.10	0.30	0	1	0.10	0.30	0	1
GOVREP	0.09	0.28	0	1	0.23	0.42	0	1
COURT	0.00	0.00	0	0	0.02	0.14	0	1
NEWSPAPER	0.05	0.22	0	1	0.11	0.31	0	1
WORKP	0.00	0.00	0	0	0.00	0.06	0	1
SPEECH	0.11	0.32	0	1	0.09	0.28	0	1
Sample size	81				256			

Appendix B: Our Previous Attempts to Bias-Correct the OE

In Boyer and Kotchoni (2011),[24].we conduct a meta-analysis by introducing three refinements with respect to Connor and Bolotova (2006). In a first step, we remove the cartels with alleged overcharge estimates that are larger than or equal to 50% as well as zero overcharge estimates representing 7.2% of the sample because the proportion of such estimates is susceptible of being affected by the publication bias. Second, we use a K-means analysis to separate the sample of overcharges into four “homogenous” groups. And third, we model the logarithm of overcharges as a linear function of the explanatory variables mentioned above, while assuming that the coefficients of variables that capture the bias vary across the clusters identified in the K-means analysis. A Heckman correction is used to eliminate the sample selection bias due to the exclusion of zeros and outliers.

Our results show that the bias captured by the estimation method and the publication source is substantial and economically significant. The bias-corrected estimates obtained from the meta-analysis by adequately neutralizing the effects of those variables suggest that the representative average overcharge estimate for cartels with initial overcharge estimates above 0% and below 50% is approximately 13.62% (with a median of 13.63%), while the average for all types of cartels is approximately 17.52% (with a median of 14.05%).

Critics of the analyses conducted in Boyer and Kotchoni (2011) centered on the following issues: the trimming of the sample at 50% has not been well-motivated; the regressors used for the meta-analysis include the indicators of the clusters identified in a prior K-means analysis on the same data; the Heckit procedure assumed that the same latent variable drives the occurrence of zeros OE and OE above 50%; the variable indicating whether the cartel members pleaded or were found guilty is likely endogenous.

In Boyer and Kotchoni (2012),[25] we conduct further K-means analyses which support that the subsample of cartels with raw OE lying in the range [0,50%] (70% of the sample) is quite representative of the whole universe of cartels. We then argue that the average bias-corrected OE obtained in Boyer and Kotchoni (2011) is one of the typical cartel.

Boyer and Kotchoni (2014)[26] largely follows the same methodology as in the present paper, but it fails to address the endogeneity of Guilty decision or plea. In particular, it relies on Kullback-Leibler divergences (in lieu of K-means analyses) to underscore the presence of bias in the raw OE

[24] Circulated under the title "The Econometrics of Cartel Overcharges" (CIRANO 2011s-35)
[25] Circulated under the title "How Much Do Cartels Typically Overcharge?" (CIRANO 2012s-15)
[26] Circulated under the title "How Much Do Cartels Overcharge?" (TSE WP-462)

data and bring other data problems to light. Boyer and Kotchoni (2014) find an average bias-corrected OE of 16.99% for the subsample of effective cartels and of 15.76% for all cartels.[27]

In the present paper, we move one step beyond Boyer and Kotchoni (2014) by taking the endogeneity of the “guilty” dummy variable into account. We apply the Heckit procedure to effective cartels so that only the right truncation of the data needs to be controlled; and the zero raw OE are included unaltered in the sample used to predict the bias-corrected OE of all cartels. We find an average bias-corrected OE of 16.68% (median of 16.17%) for the subsample of effective cartels and of 15.47% (median 16.01%) for all cartels.

Appendix C: The Overcharge Estimation Methods

Many authors acknowledge that overcharges represent the bulk of the damages caused by cartels. This explains why the largest body of the economic literature on cartels has been devoted to their price effects. The methods often used by academic researchers and forensic economists to estimate the cartel overcharge can be summarized into 5 groups: price before/after the conspiracy, price during a price war, yardstick method, cost-based method, and econometric method. Each of these approaches is briefly explained below.

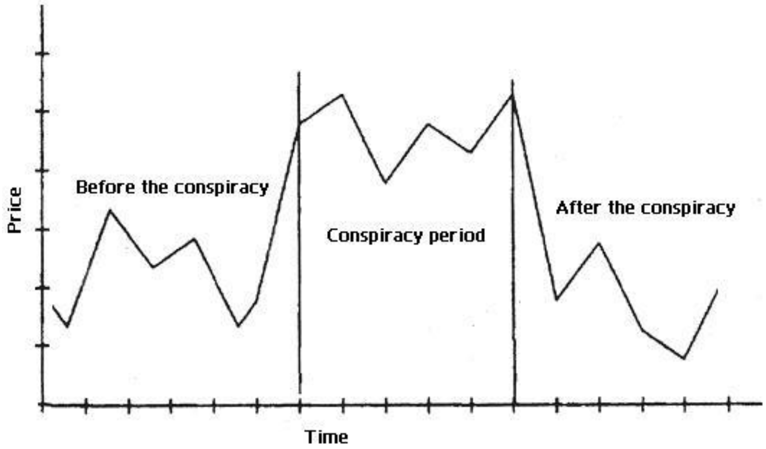
The " before-and-after" Methods

This method is based on the comparison of the price during the alleged cartel period with the price before and/or after the cartel. As pointed out by Connor (2007): *"this method should be called the " with-and-without collusion method" since the " before" period is really any nonconspiracy period---whether before, after, or during an intermediate pause in price-fixing"*

This method does not control for shifts in demand or cost functions. For that reason, Connor (2010b states:*"it is important that the ‘before’ period be one that is quite comparable to the conspiracy period with respect to demand and supply conditions. Shifts in buyer preferences, appearance or the disappearance of substitutes, or changes in the cost of production of the cartelized product during the affected period can cause overstatement or understatement of the overcharge."*

[27] The latter number is lower as it is a weighted average of 0% and 16.99%.

Figure C.1
Before-and-After Methods



The lack of robustness of the before-and-after method is illustrated by the following citation from Finkelstein and Levenback (1983): *"An obvious idea is to assume that competitive prices during the conspiracy period would have been the same as they were before or after the conspiracy or in interludes of competition within the conspiracy period [...] This estimate, however, meets the immediate objection that it is likely to be incorrect because changes in factors affecting price other than the conspiracy would have produced changes in competitive prices if there had been competition during the conspiracy period."*

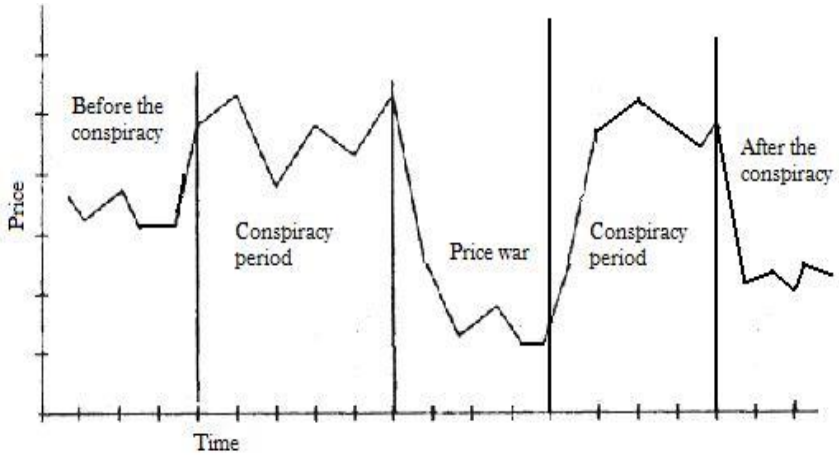
This lack of robustness is further enhanced by the difficulty of correctly specifying when the conspiracy begins and when it ends. Indeed, an incorrect specification of the conspiracy period may result in biased estimates and lead to wrong conclusions, as illustrated by the following citation from Levenstein, and Suslow (2002): *"Connor writes that there was 'disagreement about the dates of the conspiracy-effects period, the but-for price, and the type of industry conduct absent collusion.' Connor uses marginal cost (estimated from what he identifies as " highly competitive" periods) as the competitive price."*

The " Price during a price war" Method

This method uses the price during a price war of laps of collusion to proxy the but-for price. This method is basically an instance of the before-and-after method and thus suffers from the same limitations.

Figure C.2

Price war interlude



In Yinne (2003), the *"overcharge is calculated as imports*price increase/(1+price increase), and the price increase is estimated from the observed price drop subsequent to the cartel's demise."* Unfortunately, the price drop that is referred to in this citation is probably not driven by normal economic forces. In general, prices during a price war can be significantly lower than in natural competition equilibrium. As acknowledged by Connor (2007): *"a predatory episode before cartel formation will strongly overestimate the overcharge, as happened in lysine."*

The Yardstick Method

This method compares the prices during the conspiracy period with comparable or yardstick, assumed competitive firms, product or geographic markets during the alleged cartel period. The yardstick method should be used with caution because an increase in price due to domestic market cartelization can cause a partial demand shift toward nearby markets. Similar domestic firms that are not participating in the collusion will tend to follow the cartel price (umbrella effect). Connor (2010b) states: *"The yardstick approach involves the identification of a market similar to the one in which prices were fixed but where prices were unaffected by the conspiracy. A yardstick market should have cost structures and demand characteristics highly comparable to the cartelized market, yet lie outside the orbit of the cartel's influence."*

Hence, the Yardstick method may not be appropriate in the case of certain international cartels (e.g. the ADM-lysine case).

The " Cost-Based" Method

This method is based on the observation that changes in price should reflect changes in costs. The direct way to apply this method is to estimate the production costs by using the (accounting) information on firms involved in the cartel. In the lysine cartel case, prosecutors have introduced the confidential production and sales records of ADM as exhibits, and these documents are now publicly available. (See Connor 2001).

But typically, academic researchers and economic experts do not have access to confidential documents of firms. In general, the overcharge is thus approximated by substracting a *" reasonable margin"* from the actual cartel profit and dividing by the production volume. The reasonable margin should include not only the marginal production cost, but also other factors that causes the natural competition price to be larger than the marginal production cost. In particular, it should include opportunity costs, the risk premium and oligopoly mark-up when relevant. Failure to account for these would lead to overestimating the overcharge.

"Econometric" Methods

Econometric methods are not tied to a particular economic theory. This denomination gathers all methods using more or less sophisticated econometric models to assess the but-for price. Econometric methods can be used to simulate an oligopolistic competition (Cournot, Bertrand), to predict the Lerner index of market power or to estimate a demand and cost functions that account for dynamic market conditions. A simple example is given in Froeb, Koyak and Werden (1993): *"To estimate conspiracy-free prices for the earlier periods, we first fit logarithms of frozen perch winning bid prices for the post-conspiracy period on the logarithms of fresh perch prices for the corresponding month and for the prior five months. These opportunity cost variables explain 77.0% of the variance in frozen perch bid prices. This regression is then used to `backcast' predicted, conspiracy-free prices for the earlier periods."*

The potential of this method is illustrated by the following citation from Connor (2007): *"Demand for animal feed rises in the winter months, which results in an increase in the derived demand for lysine in the fall of each year. Econometric methods are better equipped to handle seasonal shifters than the simple before-and-after method. Because collusion is best timed to begin when seasonal demand rises, ignoring this factor will lead to an overestimate of damages".* This citation suggests that an econometric model if correctly built is more robust than simpler methods. However, this does not necessarily mean that econometric methods are always more reliable than other methods. In fact, building a good econometric model requires a careful selection of the relevant control variables to include. Also, the estimation results must be interpreted in light of the theory underlying the model.

Appendix D: Controlling for the Sample Selection Bias

Leaving roughly 30% of cartels out of the sample[28] used for our estimation may cause some of our conclusions to loose generality. This would be true even if we know that including these cartels deteriorates the results of the analysis. The absence or deletion of a non negligible proportion of observations raises a sample selection problem well documented in Heckman (1979). In the present case, the deletion of observations (with original overcharge estimates of 0% and of ≥49%) for estimation purposes is done to avoid obtaining distorted results due to outliers.

In models that are linear in the parameters, Heckman (1979) shows that the bias induced by sample selection on regression coefficients is a missing regressor problem. To fix ideas, let us assume there exists a latent index X_i^* indicating the quality of the data.[29] In the present context, the quality of an observation is defined in relation with its contribution to the quality of the estimation results. In concrete terms, large positive overcharge estimates are considered poor quality data, not necessarily because they are measured with more bias than other observations, but because including them distorts the relevancy of the results of our analysis.[30]

Suppose that the quality indicator X_i^* takes the following value for cartel i:

$$X_i^* = A + Y_i B + W_i C + Z_i D + u_i^* \qquad \text{(D.1)}$$

where u_i^* follows a standard normal distribution. Assume that this latent variable is such that the cartel *i* is included in the meta-analysis if and only if $X_i^* > 0$, while the cartel is exluded otherwise.[31] By definition, the estimation of the model is done only on the portion of the sample where $X_i^* > 0$. The expectation of the overcharge estimates X_i on this subsample is:

$$E(X_i|Y, Z, X_i^* > 0) = \alpha + Y_i\beta + W_i\gamma + Z_i\delta + E(\varepsilon_i|X_i^* > 0) \qquad \text{(D.2)}$$

The sample selection bias is controlled by estimating the following equation:

$$X_i = \alpha + Y_i\beta + W_i\gamma + Z_i\delta + E(\varepsilon_i|X_i^* > 0) + e_i, \qquad \text{(D.3)}$$

[28] In total, 22.87% of overcharge estimates larger than 49% and 7.23% of zero overcharge estimates.

[29] A latent variable is an unobserved variable that may have observable implications. A latent variable may be inferred using some of its observable implications along with mild identification assumptions.

[30] In the same vein, the log of zero is equal to minus infinity, hence zeros may be considered outliers in a log-linear regression analysis

[31] If the population of interest is the whole universe of cartels, our assumption on the latent variable is consistent with a situations where $X_i^* \geq b$ for zero overcharge estimates, $X_i^* < b$ for estimates lying in (0%, 49%] and $X_i^* \geq b$ the for observations above 49%. This amounts to assuming that the latent factor is nonlinear in the OE. However, one might argue that the latent variable that drives the occurrence of zeros is different from the one that drives the occurrence of outliers. To address this critic, we assume in the present paper that zero OE are unbiased and pursue the meta-analysis using the subsample of cartels with OE>0. This leads us to find the bias-corrected OE of effective cartels, which is combines with the zero OE (left unaltered) to obtain the bias-corrected OE of all cartels.

where the error term $e_i \equiv \varepsilon_i - E(\varepsilon_i|X_i^* > 0)$ has zero expectation by construction. It can be shown that:

$$E(\varepsilon_i|X_i^* > 0) = \zeta \frac{\varphi(A+Y_iB+W_iC+Z_iD)}{\Phi(A+Y_iB+W_iC+Z_iD)} \qquad \text{(D.4)}$$

where the theoretical value of θ is $\mathrm{Corr}(\varepsilon_i, u_i)\hat{\sigma}_\varepsilon^2$ and φ and Φ are the standard normal density and cumulative distributions respectively.[32] Let us consider the inverse Mills ratio (IMR) defined as:

$$\mathrm{imr}_i = \frac{\varphi(A+Y_iB+W_iC+Z_iD)}{\Phi(A+Y_iB+W_iC+Z_iD)} \qquad \text{(D.5)}$$

When restricted to the subsample of cartels defined by $X_i^* > 0$, the estimating equation (D.3) is equivalent to:

$$X_i = \alpha + Y_i\beta + W_i\gamma + Z_i\delta + \mathrm{imr}_i\zeta + e_i, \qquad \text{(D.6)}$$

If the IMR is not included in the estimating equation, the coefficients β, γ and ζ are estimated with bias. In this case, the model can still be used to predict the mean overcharge conditional on $X_i^* > 0$, but not conditional on Y, W and Z. Given this, only the model that controls for sample selection can be used to bias-correct individual overcharge estimates.

A probit can be estimated to infer fitted values of X_i^* and the IMR. Such a probit analysis is interesting per se because it permits to see the categories of cartels that have been excluded the most. We do not observe the latent variable, but we do know if an observation is excluded (I_i=0) or not (I_i=1). If observation i is included, then it must be the case that $X_i^* > 0$, which in turn implies that $A + Y_iB + W_iC + Z_iD + u_i^* > 0$, or equivalently, $-u_i^* < A + Y_iB + W_iC + Z_iD$. Because u_i^* is standard normal, the likelihood of this observation is equal to:

$$L_i = \Phi(A + Y_iB + W_iC + Z_iD) \qquad \text{(D.7)}$$

where Φ is the cumulative distribution function of the standard normal. Likewise, if observation i is excluded, then it must be the case that $X_i^* < 0$, which implies that $-u_i^* > A + Y_iB + W_iC + Z_iD$. Hence the likelihood of an excluded observation is given by $1 - L_i$, where L_i is defined above. The sample log-likelihood of the probit is thus given by:

$$\mathcal{L}(A, B, C, D) = \sum_{i=1}^{T}\{I_i * \log L_i + (1 - I_i) * \log(1 - L_i)\} \qquad \text{(D.8)}$$

Maximizing this log-likelihood with respect to parameters A, B, C and D gives their Probit estimates $\hat{A}, \hat{B}, \hat{C}$, and $\hat{D}$.

---

[32] This result is obtained by exploiting the definition of $E(\varepsilon_i|X_i^* > 0)$, which is the expectation of ε with respect to its density conditional on $-u_i^* < A + Y_iB + W_iC + Z_iD$. This density is of truncated type and is obtained by exploiting the bivariate Gaussian distribution of (ε_i, u_i^*). See Heckman (1979) for details.

Let $\widehat{X}_i^{\,*} = \widehat{A} + Y_i\widehat{B} + W_i\widehat{C} + Z_i\widehat{D}$ be the fitted values of $X_i^{\,*}$. Empirically, the sample selection bias is controlled in the meta analysis by including the estimated inverse Mills ratio obtained from the Probit extimation as an additional regressor. For an included cartel, we have:

$$\widehat{\imath mr}_i = \frac{\varphi(\widehat{X}_i^{\,*})}{\Phi(\widehat{X}_i^{\,*})} \tag{D.9}$$

Heckman (1979) shows that including $\widehat{\imath mr}_i$ as an additional regressor allows to consistently estimate the coefficients of the model linking the dependent variable (original overcharge estimates) to the explanatory variables Y, W and Z, as if the whole sample were available and used. The estimating equation is thus given by:

$$X_i = \alpha + Y_i\beta + W_i\gamma + Z_i\delta + \widehat{\imath mr}_i\zeta + e_i. \tag{D.10}$$

The meta-analysis conducted in this paper can be summarized as follows: (D.8) is maximized using the subsample of cartels with OE>0; next, the estimated coefficients $\widehat{A}, \widehat{B}, \widehat{C}$, and $\widehat{D}$ are used to calculate the IMR given in (D.9) for the subsample of cartels with $OE \in (0\%, 49\%]$; next, the regression (D.10) is estimated using the subsample of cartels with $OE \in (0\%, 49\%]$; finally, the coefficients obtained from (D.10) are used to bias-correct all OE>0. Zero OE as assumed exempt of bias.

References

Adler, H. and D.J. Laing (1997). Explosion of international criminal antitrust enforcement. *The Corporate Counsellor 11, 1.*

Adler, H. and D.J. Laing (1999). As corporate fines skyrocket. *Business Crimes Bulletin 6, 1.*

Allain, Marie-Laure, Marcel Boyer, Rachidi Kotchoni and Jean-Pierre Ponssard (2015). Are Cartel Fines Optimal? Theory and Evidence from the European Union. *International Review of Law and Economics* 42, 38-47

Allain, Marie-Laure, Marcel Boyer and Jean-Pierre Ponssard (2011). The determination of optimal fines in cartel cases: Theory and practice. *Concurrences - Competition Law Journal* 4-2011, 32-40 (Selected as *Best Academic Economics Article - 2012 Antitrust Writing Awards*, Institute of Competition Law, New York, and George Washington University Law School, Washington).

Becker, Gary (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76: 169-217.

Boyer, Marcel (2013). The Fining of Cartels. *Concurrences – Competition Law Journal* 1-2013, 27-33.

Boyer, Marcel and Rachidi Kotchoni (2011). The Econometrics of Cartel Overcharge. *CIRANO WP* 2011s-35.

Boyer, Marcel and Rachidi Kotchoni (2012). How Much Do Cartel Typically Overcharge? *CIRANO WP* 2012s-15.

Boyer, Marcel and Rachidi Kotchoni (2014). How Much Do Cartel Overcharge? *TSE WP*-462.

Cohen, Mark A. and David T. Scheffman (1989). The Antitrust Sentencing Guideline: Is the Punishment Worth the Costs? *American Criminal Law Review* 27: 331-366.

Combe, Emmanuel and Constance Monnier (2011). Fines against hard core cartels in Europe: The myth of over-enforcement. *Antitrust Bulletin 56, 235-275*.

Combe, Emmanuel and Constance Monnier (2013). Quelle est l'ampleur de la sous-dissuasion des cartels en Europe ? Compléments sur nos résultats. *Concurrences - Competition Law Journal* 1-2013, 16-26.

Connor, John M. (1997). The Global Lysine Price-Fixing Conspiracy of 1992-1995. *Review of Agricultural Economics*, Vol. 19, No. 2, pp. 412-427

Connor, John M. (2010). *Price-fixing Overcharges, 2nd Edition* (available online at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1610262)

Connor, John M. (2014). Price-fixing Overcharges. *The Law and Economics of Class Actions: Research in Law and Economics*, Volume 26, 249-387

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

Connor, John M. and Robert H. Lande (2008). Cartel Overcharges and Optimal Cartel fines. *Issues in Competition Law and Policy*, S.W. Waller (ed.), Volume 3, AMA Section of Antitrust Law, chapter 88, pp. 2203-2218.

Denger, Michael L. (2003). Too Much or Too Little, A Summary of Discussion. *American Bar Association's Antitrust Remedies Forum*, Washington, DC. April.

Ehmer, C. and F. Rosati (2009). Science, Myth, and Fines: Do Cartels Typically Raise Prices by 25%? *Concurrences - Competition Law Journal* N° 4-2009,

Finkelstein, M. and H. Levenback (1983). Regression Estimates of Damages in Price-Fixing Cases. *Law and Contemporary Problems* 46: 145-169.

Ginsburg, D.H. and J.D. Wright (2010). Antitrust Sanctions. *Competition Policy International* 6, No. 2, pp. 3-39.

Hall, R.E. (1988). The relation between price and marginal cost in U.S. industry. J. Polit. Econ. 96, 921–947.

Harrington, J.E. (2005). Optimal Cartel Pricing In The Presence Of An Antitrust Authority. *International Economic Review*, vol. 46(1), pages 145-169, 02.

Harrington, J.E. (2010). Comment on 'Antitrust Sanctions' by Douglas Ginsburg and Joshua Wright. *Competition Policy International*, Vol 6, No. 2, pp. 41-51.

Harrington, J.E. (2014). Penalties and the deterrence of unlawful collusion. *Economic Letters* 124 (2014) 33–36.

Heckman, James (1979). Sample Selection Bias as a Specification Error. *Econometrica.*47**,** 153-161.

Houba, H., Motchenkova, E. and Q. Wen (2013). Legal Principles in Antitrust Enforcement. TI Discussion Paper 13-178/II, Amsterdam: Tinbergen Institute. http://papers.tinbergen.nl/13178.pdf

Hunter, J. and F. Schmidt (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings,* Second Edition, SAGE Publications.

Hüschelrath, K., Müller K. and T. Veith (2012). Concrete Shoes for Competition: The Effect of the German Cement Cartel on Market Price. *Center for European Economic Research*, Discussion Paper No. 12-035.

Ivaldi, M., Jullien, B., Rey, P., Seabright, P. and J Tirole (2003). The Economics of Tacit Collusion. *Monograph IDEI*. Final Report for DG Competition, European Commission.

Katsoulacos, Y. and D. Ulph (2013). Antitrust Penalties and the Implications of Empirical Evidence on Cartel Overcharges. *The Economic Journal* 123, Issue 572.

Kearney, Elisa (2009). The Competition Bureau '20 Percent Solution': Does the Punishment Fit the Crime? *The Marker*, CBA National Competition Law Section, March.

Kovacic, W. E. and C. Shapiro (2000). Antitrust Policy: A Century of Economic and Legal Thinking. *Journal of Economic Perspectives* 14, Number 1, Pages 43–60

Kullback S. and R. A. Leibler. 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics*, Volume 22, Number 1 (1951), 1-164.

Landes, William H. (1983). Optimal Sanctions for Antitrust Violations. *University of Chicago Law Review* 50: 652-678.

Levenstein, Margaret C. and Valerie Y. Suslow (2003). Contemporary International Cartels and Developing Countries: Economic Effects and Implications for Competition Policy. *Antitrust Law Journal* 71: 801-852.

Morrison, Catherine J. (1990). Market Power, Economic Profitability, and Productivity Growth Measurement: An Integrated Structural Approach, *NBER Working Paper* No. 3355.

Nieberding, J.F. (2006). Estimating Overcharges in Antitrust Cases Using a Reduced-Form Approach: Methods and Issues. *Journal of Applied Economics*, Vol IX, No. 2, 361-380.

White, J. Lawrence (2001). Lysine and Price Fixing: How Long? How Severe? *Review of Industrial Organization* 18: 23–31.

1130, rue Sherbrooke Ouest, bureau 1400, Montréal (Québec) H3A 2M8
Tél. : 514-985-4000 • Téléc. : 514-985-4039
www.cirano.qc.ca • info@cirano.qc.ca

Centre interuniversitaire de recherche en analyse des organisations
Center for Interuniversity Research and Analysis on Organizations