THE ECONOMIC JOURNAL

The Economic Journal, **123** (*November*), F558–F581. Doi: 10.1111/ecoj.12075 © 2013 The Author(s). The Economic Journal Published by John Wiley & Sons on behalf of the Royal Economic Society. Published by John Wiley & Sons, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

ANTITRUST PENALTIES AND THE IMPLICATIONS OF EMPIRICAL EVIDENCE ON CARTEL OVERCHARGES*

Yannis Katsoulacos and David Ulph

This article makes two contributions to the literature linking penalties charged by competition authorities to observed cartel price overcharges. (*i*) It extends the theory of optimal penalties by introducing new considerations regarding the timing of penalty decisions. Drawing on a new European data set to calculate these additional factors, the optimal penalty is shown to be approximately 75% of that implied by the conventional formula. (*ii*) It shows that because penalties are typically imposed on revenue, a tougher regime may increase cartel overcharges. This calls into question some recent empirical findings on this issue and the potential benefits of raising penalties.

The issue of optimal penalties for antitrust violations has attracted a lot of attention among economists[1] at least since Landes (1983) applied Becker's (1968)[2] analysis to provide a first theoretical treatment. In particular, in a series of recent papers, Connor and Lande (2005, 2006, 2008, 2012) have argued that the prevailing US and European penalties for antitrust violations – particularly cartels – are too low to generate the optimal level of deterrence. They draw on a range of empirical evidence to support this conclusion, a crucial figure being that for the cartel overcharge – the percentage amount by which the collusive price exceeds that which would obtain in the absence of the cartel. They draw on an extensive survey by Bulotova and Connor (2006), which reports a mean value for the cartel overcharge of 29%.[3]

A significant recent challenge to these conclusions has been made by Allain *et al.* (2011), who question two key elements in the Connor and Lande calculations. The first is the estimate of the average cartel overcharge. Drawing on Boyer and Kotchoni (2011), they argue that Bulotova and Connor's estimate suffers from a significant upward bias, and report that, when these biases are corrected, the mean value for the cartel overcharge is 17.5%. The second is the use of an annual probability of detection. They argue that the appropriate probability is that of detection over the lifetime of the

* Correponding author: David Ulph, School of Economics and Finance, University of St Andrews, St Andrews, Fife KY16 9AL, UK. Email: du1@st-andrews.ac.uk.

Research for this paper was supported by a grant from the Economic and Social Research Council RES-062-23-2211 'Optimal Decision and Enforcement Structures for Competition Policy'. We are grateful to Kelly Benetatou, Vasiliki Bageri and Jacob Seifert for research assistance. An earlier version of the paper was presented at the CRESSE Workshop on 'Optimal Decision and Enforcement Structures for Competition Policy' on 8 July 2012 in Chania. We thank Emanuele Giovanetti, other Workshop participants, Rachel Griffith and a referee for helpful comments. All errors remain our own.

[1] For a recent review of the literature on optimal fines, see Polinsky and Shavell (2000), while for discussions concentrating on optimal fines in the case of cartels, see Buccirossi and Spagnolo (2007). Other recent contributions which focus on the dynamic aspects of the issue include Motchenkova and Kort (2006) and Motchenkova (2008).

[2] See also Stigler (1970) for an early important contribution.

[3] Connor and Lande (2012) report on data from another survey, which gives a mean value of 31%.

cartel, which, being much larger, gives a lower optimal penalty. They conclude that there is no reason to think existing penalties are too low.

Rather than going from observed levels of overcharges to implied optimal penalties, another natural and related question to ask is the reverse one: if cartels faced a tougher penalty regime – by which we mean not just the level of penalty but the anticipated probability of successful antitrust enforcement – would this result in a lower cartel overcharge? There has also been some limited empirical work on this issue. Thus, drawing on their database of 800 cartels across a large number of countries and over a very long period of time Bulotova and Connor (2006, pp. 1133, 1134) report that 'cartels … tend to achieve lower overcharges in jurisdictions with strongly enforced antitrust laws' and that the size of the overcharge is lower in more recent periods, especially 1990–2004 'when antitrust sanctions were harshest'.

It should be noted that Bulotova and Connor do not attempt to find alternative explanations of why cartel price overcharges have been falling in more recent decades. For example, increased globalisation may imply that domestic cartels are formed under tighter competitive constraints, that is, they are more difficult to sustain, with excessive overcharges, in markets increasingly open to international competition. In addition, Boyer and Kotchoni (2011, p. 49) detect no evidence that cartel overcharges have been affected by the toughness of the penalty regime and report 'a fairly homogeneous behavior of cartels across different types, geographical locations and periods'.

The aim of this article is to contribute to both of these issues. In Section 1, we revisit the determination of the optimal penalty for antitrust violations and introduce a number of factors relating to the way that competition cases are handled that have not so far been taken systematically into account. The existing literature, based on the economics of crime, effectively assumes that the detection and prosecution of cases takes place immediately after the action has come to its natural end. However, antitrust violations can last for many years and competition authorities (CAs) sometimes intervene and terminate actions before they have come to a natural end.[4] On the other hand, a CA may only reach a decision on a case and impose a penalty long after the antitrust action has terminated. Each of these two factors raises considerations that point in different directions for the optimal level of penalty.

If desistance takes place and so an action is stopped before it has reached its natural end, then the firm will suffer a loss of profits relative to what otherwise might have happened and so the penalty does not need to be so high to generate the same level of deterrence. However, on the other hand the revenue base on which the penalty will be imposed is smaller than it would otherwise have been had the action lasted its natural life and so the penalty rate has to be higher to achieve the same level of deterrence.

If a decision can be reached and a penalty imposed long after the action has come to a natural end then this implies that the probability of effective action ever being taken is higher than if the action is taken only when the action has reached its natural life – pointing to a lower penalty. However, the fact that the penalty is imposed much later

[4] Put differently, CAs can stop antitrust actions through both desistance and deterrence.

means that, discounted back to the present, it represents a lower potential cost to the firm contemplating taking the action, and so the penalty rate needs to be raised to have the same deterrence effect.

In Section 2, we quantify the resulting optimal penalty, drawing in part on existing empirical evidence but also on a new database of antitrust actions prosecuted under Article 102 TFEU by various European competition authorities.[5] These data enable us to calculate the natural life of actions, the fraction terminated prior to the natural life – and how much sooner this happens – and the fraction that are penalised after the natural life – and how much later this happens. We show that the combined effect of these two missing factors is to lower optimal penalties to around 75% of what would have been predicted using the conventional approach. Overall, our conclusion supports that of Allain *et al.* (2011) – that existing penalties are within the range supported by calculations of optimal penalties.

In Sections 3 and 4, we provide a theoretical framework for thinking about the link between the 'toughness of the penalty regime' and the observed level of cartel overcharge. We show that this link is far from straightforward and may indeed go in the opposite direction to that expected. This provides additional reasons for being cautious about prescriptions for significantly raising the level of penalties.

Our analysis highlights a number of important considerations that need to be taken into account when considering the effect of fines on cartel overcharges.

(*i*) First, the effect of fines on overcharges depends on the incidence of the former, i.e. whether they are imposed on profits or revenues;
(*ii*) Second, while Bulotova and Connor (2006) associate more 'strongly enforced' or 'more effective' antitrust law implementation with higher penalties, more generally, a 'harsher antitrust enforcement regime' can also be associated with a higher probability of being convicted;[6]
(*iii*) Third, we distinguish between the case where the penalty regime is independent of a firm's actions, and that where either the size of the penalty (as a proportion of revenue/profits) or the probability of conviction depends on the price overcharge;
(*iv*) Fourth, we allow for deterrence effects by examining the impact of the penalty regime on not just the overcharge set by non-deterred cartels but also on the number and types of cartels that form.

Section 3 of our article examines the effects of penalties on non-deterred cartels – the first three issues – while Section 4 considers the implications of deterrence.

We show that the effect of harsher antitrust regimes on cartel overcharges is highly ambiguous, though generally our results tend to show no effect or the opposite effect to that reported by Bulotova and Connor (2006). It is mainly when firms anticipate a stricter regime when their overcharge is higher that we may find theoretical support for their finding that tougher regimes will lower overcharges.

[5] Thus, we no longer rely on evidence relating to just one form of antitrust violation – cartels.
[6] This is the probability of being detected multiplied by the probability of being convicted (and hence penalised) conditional on detection.

Taken together with the results of Sections 1 and 2 this suggests that one has to be extremely careful when drawing antitrust fining policy recommendations on the basis of existing economic models and empirical evidence.

1. Theoretical Derivation of Optimal Antitrust Penalties on the Basis of Price Overcharges and Other Factors

1.1. *Background*

The context in which our analysis is set is the following: CAs set out guidance on how they decide what penalty to impose when a firm[7] takes some anti-competitive action that violates Competition Law. The same guidance applies whatever the nature of this action.

The ultimate fine paid by any firm depends on a variety of factors which reflect the extent to which the CA wants to penalise or reward firms for taking other ancillary actions: e.g. the extent to which they facilitate or hinder an investigation; or whether they are serial offenders. However, typically CAs start by calculating a basic penalty that is imposed for taking the anti-competitive action and this is then adjusted to take account of these other factors.[8] It is the calculation of this basic penalty with which we are concerned.

The starting point for calculating this basic fine is normally the revenue[9] that the firm made in the last year in which the action took place. This could either be the last year of the natural life of the action if the authority intervenes only after the action has come to an end, or it could be the last year of an on-going action which the authority has ordered the firm to cease before it has to come to a natural end. There is then an adjustment made to take account of the duration of the action – which typically takes the form of just multiplying the revenue by the number of years over which the action took place. The basic fine that is set is calculated as a proportion of the last year revenue adjusted for duration.[10] The question is what factor of proportionality should the CA use to set its penalty?

As we will show, depending on the welfare standard used by a given CA the penalty will be related to either the harm that the action has caused to others – particularly consumers – or the benefit that the firm has derived from taking the action. In either case, CAs start by turning to evidence about price overcharges – the extent to which prices have been driven by anti-competitive actions – and in particular to evidence about cartel overcharges, as these have been most extensively analysed.

In this Section, we examine systematically how such overcharge information could be used to calculate the penalty. We argue that the formula that is often used for doing this misses some important considerations regarding the timing of decisions by the CA as to whether or not an action is anti-competitive. For, if it is deemed to be so, then the

[7] This should be understood to include the possibility that, in certain situations – e.g. mergers, cartels – the action may be taken by a group of firms.

[8] Bageri *et al.* (2013) contains a review of how CAs in Europe and US set fines in practice.

[9] Bageri *et al.* (2013) make clear how levying fines on revenue, rather than profits, introduces a number of distortions.

[10] More complex penalty structures have been analysed by, e.g. Motchenkova (2008) and Houba *et al.* (2012).

timing of this decision could potentially curtail the stream of profits arising from the action and will certainly determine when the fine is imposed. We derive a more general formula that captures these considerations.

1.1.1. *Price overcharge*

In common with a lot of the literature, we make the following simplifying assumptions: there is a market for a homogeneous product in which the production technology is characterised by constant unit costs; there is an inter-temporally constant price and output in both the situation where the firm has acted anti-competitively and that in which it has not.

In the counterfactual situation in which the potentially anti-competitive action had not been taken, we assume that the firm would have had constant unit costs $c_0 > 0$, the equilibrium price would have been $p_0 \geq c_0$ and the equilibrium output $Q_0 > 0$. The counterfactual price p_0 is sometimes known as the 'but-for' price. Notice that there is no presumption that the counter-factual situation is that of perfect competition. There may be 'natural' forces of competition – barriers to entry, limited number of firms – that would have produced an outcome other than perfect competition.

The firm takes some anti-competitive action which has the effect of not only raising the price-cost margin but may also have an efficiency effect of lowering costs. As we are looking at the role of price-overcharge information in setting penalties, we assume that the former effect dominates the latter and, overall, price increases.[11] So formally, once the action has been taken unit costs are $c_1, 0 < c_1 \leq c_0$ the equilibrium price is $p_1 > p_0$ and the equilibrium output is $Q_1, 0 < Q_1 < Q_0$.

The price overcharge, θ, is the extent to which price is raised above its 'but-for' level, expressed as a fraction of the 'but-for' price and so is defined as:

$$\theta = \frac{\Delta p}{p_0} = \frac{p_1 - p_0}{p_0} > 0. \tag{1}$$

Without loss of generality, we normalise prices by assuming that

$$p_0 = 1, \tag{2}$$

so the price overcharge reflects both the absolute and the percentage increase in price.

Associated with these two equilibria are revenues $R_i = p_i Q_i, i = 0, 1$, and profits $\pi_i = R_i - c_i Q_i = (p_i - c_i) Q_i, i = 0, 1$. We assume that, in the absence of intervention by a CA, the firm would want to take the action because then it generates a positive change in profits, i.e. $\Delta\pi = \pi_1 - \pi_0 > 0$.

We assume that this action imposes harm on others which is not corrected through a successful claim for private damage, and so constitutes a genuine externality. In this article we focus solely on the harm to consumers through the loss of consumer surplus. To calculate this and relate it in a straightforward way to the price overcharge, assume a linear demand function which, by a suitable choice of units, can be written as:

[11] In related work, Katsoulacos *et al.* (2011), we consider the more general case where actions may be on balance pro-competitive and part of the role of CAs is to try to discriminate between pro and anti-competitive actions.

$$p = (1 + \varepsilon) - Q, \quad \varepsilon > 0. \tag{3}$$

Given the normalisation in (2), $\varepsilon = Q_0 = -(\mathrm{d}p/\mathrm{d}Q)(Q_0/p_0)$ and so measures the inverse elasticity of demand in the 'competitive' equilibrium, and is thus a measure of the underlying competitiveness of the industry in which the action is taking place.[12]

It also follows from (2) and (3) that:

$$Q_1 = \varepsilon - \theta, \tag{4}$$

and so, to ensure positive output and profits when the potentially anti-competitive action is taken, it must be the case that $\theta < \varepsilon$.

Given this demand function it follows that $CS_i = \frac{1}{2}Q_i^2$, $i = 0, 1$ and so

$$\Delta CS = CS_0 - CS_1 = \frac{1}{2}\left(Q_0^2 - Q_1^2\right) > 0, \tag{5}$$

measures the loss of consumer surplus – the harm – caused by the action.

1.2. *Deriving the Optimal Penalty: The Conventional Analysis*

To go from this information to a calculation of the required penalty, involves consideration of a number of different factors. We start by setting out the conventional analysis in this subsection and then, in the following subsection, consider the extensions that need to be made to reflect the fact that anti-competitive competitive actions last a long time and that CAs can intervene either before or after they have come to a natural end.

1.2.1. *Assumptions*

To start with, we assume that:

(*i*) The anti-competitive action lasts for just a single period;
(*ii*) At the end of the period the firm taking the action faces a probability χ, $0 < \chi \leq 1$ of having its action investigated by a CA. We refer to χ as the coverage rate;
(*iii*) If a firm is investigated, the CA will be able to determine for sure that the action is anti-competitive (has imposed harm on consumers) and impose a penalty.

We first consider the implications of two different welfare standards.

1.2.2. *Total welfare standard – restitutive penalties*

Here a firm should take (not take) the action according to whether the private benefit from doing so is greater than (less than) the social harm. This is the welfare standard proposed by Connor and Lande (2005, 2006, 2008, 2012), following Landes (1983):

[12] Nothing depends on the assumption of a linear demand and indeed much of the literature effectively uses a linear demand curve as a first-order approximation when calculating the impact of the overcharge on profits and consumer surplus. If the demand curve is linear, then the units in which price and quantity are measured can be chosen to reduce this to a single parameter representation and our simple functional form ensures that this parameter captures an essential feature of the industry – how intrinsically competitive it is, as measured by elasticity of demand at the but-for price. By measuring elasticity at the but-for price we avoid the cellophane fallacy.

$$\Delta\pi - \Delta CS \gtreqless 0. \tag{6}$$

If a firm that obtains the full benefit, $\Delta\pi > 0$, of taking the action faces the possibility of paying a fine $F > 0$ with probability χ, $0 < \chi \leq 1$, then it will take the action as long as the net benefit is positive, i.e. it will take/not take the action according to whether:

$$\Delta\pi - \chi F \gtreqless 0. \tag{7}$$

To line up public with private incentives it is therefore necessary that

$$\chi F = \Delta CS. \tag{8}$$

The penalty here is playing the role of a Pigovian tax on an externality and is not designed to stop all actions, just to ensure that private decisions to take an action are lined up with the net social benefits. Expressed as a fraction of revenue earned by taking the action, the optimal penalty rate under a Total Welfare standard is as follows:

$$\varphi^{TW} = \frac{F}{R_1} = \frac{\Delta CS/R_1}{\chi}. \tag{9}$$

To relate this to the price overcharge note that from (3), (4) and (5) we have the following:

$$\Delta CS = \frac{1}{2}\left[(Q_1 + \Delta p)^2 - Q_1^2\right] = \Delta p Q_1 + \frac{(\Delta p)^2}{2},$$

and so:

$$\frac{\Delta CS}{R_1} = \frac{\Delta p}{p_1} + \frac{1}{2}\frac{\Delta p}{p_1}\frac{\Delta p}{Q_1} = \frac{\theta}{1+\theta}\left(1 + \frac{1}{2}\frac{\theta}{\varepsilon - \theta}\right). \tag{10}$$

Substituting (10) into (9) we get the following:

$$\varphi^{TW} = \frac{\frac{\theta}{1+\theta}\left(1 + \frac{1}{2}\frac{\theta}{\varepsilon - \theta}\right)}{\chi}. \tag{11}$$

Notice that if $2(\varepsilon - \theta) \approx 1$ then we get the approximation:

$$\tilde{\varphi}^{TW} = \frac{\theta}{\chi}, \tag{12}$$

which provides the basis for using the price overcharge as the starting point for calculating optimal penalties.

1.2.3. *Consumer surplus standard – dissuasive penalties*

Most CAs use a consumer surplus welfare standard for deciding whether or not actions are anti-competitive and so should be subject to penalties.[13] For example, the 2008

[13] See Salop (2010) for arguments in favour of such a standard and Carlton (2007) for arguments in favour of a total welfare standard.

European Commission's Guidance Paper on Art. 102 TFEU[14] states (in para. 5) that the Commission 'will focus on those types of conduct that are most harmful to consumers'. The latest version of Merger Guidelines in US[15] also clearly states that 'the Agency considers whether cognisable efficiencies likely would be sufficient to reverse the merger's potential harm to consumers in the relevant market, e.g. by preventing price increases in that market'.

As all actions of the type we have been considering are harmful, the CA would want to set a penalty that ensures that no action is profitable. Consequently, expressed as a fraction of the revenue earned by taking the action, the optimal penalty under a consumer surplus standard is as follows:

$$\varphi^{CS} = \frac{F}{R_1} = \frac{\Delta\pi/R_1}{\chi}. \tag{13}$$

Now it is straightforward to show that

$$\frac{\Delta\pi}{R_1} = \frac{p_1 - p_0}{p_1} - \frac{(p_0 - c_0)(Q_0 - Q_1)}{p_1 Q_1} + \frac{(c_0 - c_1)}{p_1}, \tag{14}$$

where the first term represents the increase in profits through charging a higher price; the second term represents the extent to which profits would have fallen in the counterfactual through having lower output, and the third is the increase in profits through greater efficiency.

If the counterfactual situation were perfectly competitive and there were no efficiency gains through the action then we would have the approximation:

$$\tilde{\varphi}^{CS} = \frac{\theta/(1+\theta)}{\chi}, \tag{15}$$

which, from (11), is smaller than the penalty under a total welfare standard because of the deadweight loss that is reflected in the total welfare standard. It is also smaller than the approximation in (12) that gives the conventional approach of using the price overcharge as the starting point for calculating the optimal penalty.

The expression in (15) will also be correct if it is assumed that, on average, any efficiency enhancing effects of actions are offset by the reduction in profits in the counterfactual. Because CAs typically use a consumer surplus welfare standard, in what follows we will use (15) as the basic formula for the optimal penalty factor, and consider how this has to be further adjusted in the light of other considerations that follow from systematically relaxing the assumptions in the base case.

1.2.4. *Imperfect decision-making by CA and legal uncertainty*

So far we have assumed that the CA can accurately determine whether or not an action is harmful and that firms know this, so the only uncertainty they face is whether or not

[14] Guidance on the Commission's Enforcement Priorities in Applying Article 102 EC Treaty to Abusive Exclusionary Conduct by Dominant Undertakings, Commission of the European Communities, Brussels, 3 December 2008.

[15] US Department of Justice and the Federal Trade Commission, Horizontal Merger Guidelines (Revised 1997) available at: http://www.justice.gov/atr/public/guidelines/horiz_book/hmg1.html.

they will be investigated. In Katsoulacos and Ulph (2012*a*) we have examined the optimal penalties and the optimal legal standards (effects-based or *per se*) in situations where

(*i*) the CA may make Type I and/or Type II errors;
(*ii*) there may be legal uncertainty so firms may not fully understand
  (*a*) whether or not their actions are truly harmful and/or
  (*b*) the basis on which the CA will determine whether or not their actions are harmful.

We show that if the CA can set the optimal penalty then

(*i*) the optimal legal standard is unambiguously an effects-based standard;
(*ii*) while the degree of legal uncertainty might affect the optimal penalty that should be imposed,[16] if different firms face different degrees of legal uncertainty then the CA will still achieve the optimal deterrence by setting a penalty which, expressed as a fraction of revenue, is given by:

$$\varphi^{TW} = \frac{\Delta CS/R_1}{\chi\overline{\rho}}; \quad \varphi^{CS} = \frac{\Delta\pi/R_1}{\chi\overline{\rho}}, \tag{16}$$

where $\overline{\rho}, 0<\overline{\rho}<1$ is the average probability of the CA deciding that an action that has been investigated is anti-competitive and a penalty should be imposed.

Using the approximations in (12) and (15) these equations become as follows:

$$\tilde{\varphi}^{TW} = \frac{\theta}{\chi\overline{\rho}}; \quad \tilde{\varphi}^{CS} = \frac{\theta/(1+\theta)}{\chi\overline{\rho}}. \tag{17}$$

The formula in (16) for the optimal penalty under a total welfare standard is that used by Connor and Lande as a basis for calculating the optimal penalty. Rather than profilerating formulae and numerical values, in what follows we prefer to use the formula applicable under a consumer surplus standard, because that is the *de facto* standard used by most authorities. Using (17) it is straightforward to calculate the appropriate value of the optimal penalty under a total welfare standard.

1.3. *Duration*

So far we have assumed that the action lasts for a single period, at the end of which the CA may investigate and impose a penalty. We now take seriously the idea that, in the absence of any action by a CA, anti-competitive actions may last for many years but that a CA may either intervene and stop the action before it would otherwise have ended, or may reach a decision and impose a penalty many years after the action has come to a natural end.

[16] For example, if no firm faced any legal uncertainty then the optimal penalty is given by (15).

1.3.1. *Assumptions*

Suppose a firm takes an action. There are a number of features of this action.

(*i*) It has a natural lifetime of $L > 0$ periods;[17]

(*ii*) There is a probability $\chi_b, 0 \leq \chi_b < 1$ that, at or before it reaches its natural life, the CA will have detected that the action has been taken and, following an investigation, reached a decision as to whether the action is deemed to be harmful (anti-competitive) or benign (not anti-competitive);

(*iii*) There is a probability $\chi_a, 0 \leq \chi_a \leq 1 - \chi_b$ that the CA will reach a decision on the action after it reaches its natural life. If we assume that CAs can keep identifying and investigating actions forever, then eventually all actions will be investigated and so $\chi_a = 1 - \chi_b$. However, if there is some time limit beyond which the CA will not pursue an investigation into an action, then $\chi_a < 1 - \chi_b$ and $1 - \chi_a - \chi_b$ represent the probability that an action will go unchallenged by the CA, in which case the firm will derive the per-period increased flow of profits $\Delta\pi = \pi_1 - \pi_0 > 0$ throughout the natural life of the action;

(*iv*) If the CA's decision is reached at or before the action has reached its natural life then, on average, this occurs at date L_b, $0 < L_b \leq L$, whereas if it is reached after the action has reached its natural life then, on average, this occurs at date $L_a > L$;

(*v*) Irrespective of when the CA reaches its decision, there is a probability $\overline{\rho}, 0 < \overline{\rho} < 1$ that it will be found to be anti-competitive.

To understand the implications of these assumptions, consider separately the two possible timing outcomes.

(*i*) The decision is reached at or before the action has reached its natural life. With probability $(1 - \overline{\rho})$ the CA will conclude that it is benign and so neither stop nor change the action, which will therefore generate the per-period flow of profits $\Delta\pi = \pi_1 - \pi_0 > 0$ throughout its natural lifetime. However, with probability, $\overline{\rho}$, the CA will deem the action is anti-competitive and order the firm to stop the action at date L_b[18] and to pay a penalty that is a proportion $\varphi > 0$ of the revenue it has earned up until that date. So the firm will make a flow of (net) profits that will be $\Delta\pi - \varphi R_1$ up until L_b and zero thereafter;

(*ii*) The CA reaches its decision after the natural life of the action. With probability $(1 - \overline{\rho})$ the CA will retrospectively approve it, and the firm will simply have earned the per-period flow of increased profits, $\Delta\pi = \pi_1 - \pi_0 > 0$ for L periods. However, with probability, $\overline{\rho}$, $0 < \overline{\rho} < 1$, the CA will disapprove of the action. The firm will still have made the per-period flow of increased profits, $\Delta\pi = \pi_1 - \pi_0 > 0$ throughout the natural lifetime of the action but, at the time it makes its decision, the CA will impose a penalty that is proportional to the value of the revenue that the firm earned over the natural lifetime of the

[17] In practice this will vary across actions but, for simplicity, we assume all actions have the same natural life, L.

[18] We subsume in this the possibility that the CA may impose remedies that remove the anti-competitive features of the action. Effectively we assume that, in terms of its impact on profits, this is tantamount to stopping the action.

action, namely $R_1[e^{rL} - 1/(r)]$. Discounted back to the time when the firm takes the action the present value of this penalty is given as

$$e^{-rL_a}\varphi R_1[e^{rL} - 1/(r)] = e^{-r(L_a - L)}\varphi R_1[1 - e^{-rL}/(r)].$$

It follows from this that the expected present value of the change in profits from taking the action is as follows

$$\Delta\tilde{\Pi}^e = \chi_b\left[\overline{\rho}\left(\frac{1 - e^{-rL_b}}{r}\right)(\Delta\pi - \varphi R_1) + (1 - \overline{\rho})\left(\frac{1 - e^{-rL}}{r}\right)\Delta\pi\right]$$
$$+ \chi_a\left(\frac{1 - e^{-rL}}{r}\right)\left(\Delta\pi - \overline{\rho}\varphi e^{-r(L_a - L)}R_1\right) + (1 - \chi_a - \chi_b)\left(\frac{1 - e^{-rL}}{r}\right)\Delta\pi.$$

This is equivalent to getting, over the natural life of the action, the per-period flow of profits

$$\Delta\tilde{\pi}^e = [1 - \overline{\rho}\chi_b(1 - \lambda_b)]\Delta\pi - \overline{\rho}\chi_b\varphi R_1\left(\frac{\chi_a}{\chi_b}\lambda_a + \lambda_b\right), \tag{18}$$

where

$$\lambda_a = e^{-r(L_a - L)}; \quad \lambda_b = \frac{1 - e^{-rL_b}}{1 - e^{-rL}}; \quad 0 < \lambda_j \leq 1,\ j = a, b.^{19} \tag{19}$$

It is easy to see that the conventional approach is just a special case of this. In that approach all decisions are made just as soon as the natural life of the action has expired – neither before, nor after. If no decisions are made after the natural life of the action then $\chi_a = 0$, while if none are reached before the natural life then $L_b = L \Rightarrow \lambda_b = 1$. In this case, the expression in (18) just reduces to the standard formula for expected profits:

$$\Delta\tilde{\pi}^e = \Delta\pi - \chi_b\overline{\rho}\varphi R_1, \tag{20}$$

where the relevant probability of being investigated is χ_b – the probability of being detected by the time the natural life has been reached.

1.3.2. *The optimal penalty*

If we use a consumer surplus welfare standard the optimal penalty is one that deters all firms, so we want the lowest penalty that makes expected profits as given by (18) zero. This implies that

$$\varphi^{CS} = \left(\frac{\Delta\pi/R_1}{\overline{\rho}\chi_b}\right)\left[\frac{1 - \overline{\rho}\chi_b(1 - \lambda_b)}{\frac{\chi_a}{\chi_b}\lambda_a + \lambda_b}\right]. \tag{21}$$

From (16) the first term on the RHS of (21) is just the conventional formula for the optimal penalty under a consumer-surplus standard,[20] where the notation χ_b just emphasises that, under the conventional approach, the relevant detection probability is that of being detected by the time the natural life of the action has been reached.

[19] To first order we have approximation: $\lambda_b \approx L_b/L$.

[20] It is the value of the penalty that makes the conventional expression for profits as given by (20) exactly zero.

The second term on the RHS of (21) therefore captures all the adjustments that need to be made to this formula to take account of the possibility that decisions can be made both before and after the natural life of the action has been reached.

The fact that decisions can be reached before the natural life of the action has been reached means that $L_b < L$ and so $\lambda_b < 1$, whereas in the conventional approach $L_b = L \Rightarrow \lambda_b = 1$. We can see that this has two effects: it reduces both the numerator and the denominator of the second term. The first effect lowers the optimal penalty – reflecting the fact that the loss of profits that firms suffer from early intervention itself acts as a penalty, and so means that the penalty applied by the CA does not need to be so high to have same deterrent effect. However, the second effect raises the optimal penalty as the revenue base to which the penalty is applied is also smaller because of the early intervention.

The fact that the decision can be reached after the natural life of the action has two effects. First $L_a > L \Rightarrow \lambda_a < 1$ whereas, under the conventional approach, $\lambda_a = 1$. As λ_a enters the denominator this factor raises the optimal penalty. This reflects the fact that, with discounting, the additional delay in imposing the penalty reduces its present value, so the penalty itself has to be higher to have the same deterrent effect. However, the second effect of having decisions reached after the natural life is that there is now a higher probability that an action will be dectected at some point – which is reflected in the fact that $\chi_a > 0$, whereas, under the conventional approach, $\chi_a = 0$. This higher probability of detection means that the penalty can be reduced and still have the same level of deterrence.

So, as explained in the introduction, when we allow the possibility that decisions can be made both before and after an action reaches its natural life, the optimal penalty has to be adjusted to reflect these two timing considerations, each of which introduces two facets, one of which increases the penalty and one of which reduces it.

To relate the optimal penalty to the price overcharge, we can use the approximations in (15) and (17) to get the formula

$$\tilde{\varphi}^{CS} = \left[\frac{\theta/(1+\theta)}{\overline{\rho}\chi_b}\right]\left[\frac{1-\overline{\rho}\chi_b(1-\lambda_b)}{\frac{\chi_a}{\chi_b}\lambda_a + \lambda_b}\right]. \tag{22}$$

2. Using Empirical Evidence on Price Overcharge and Other Factors to Calculate the Optimal Penalty

To obtain the optimum penalty as given by (22), it is necessary to calculate the following:

- the penalty that would emerge from the conventional formula – the expression in the first bracket on RHS of (22);
- the adjustment that needs to be made to account for the timing considerations introduced in subsection 1.3 – the expression in the second bracket on RHS of (22).

As there is an extensive literature on the penalty implied by the conventional formula, our primary focus will be on the adjustment factor.

2.1. *Adjustment Factor*

To calculate this we have used a new data set that we have compiled. This relates to abuse of dominance cases investigated by various European national competition authorities between 1990 and 2010. We have a sample of 32 such cases where we know the dates on which the action started and stopped; the date on which the decision was made by the CA and what that decision was. A Table setting out the basic data on which all the calculations set out below are based is presented in Appendix A.

Using this information we identified two subgroups of cases. The first was a group of cases which had run their natural life. We put an action in this category if either

(*i*) the date at which the CA made its decision was after the date at which the action had ceased or

(*ii*) the CA did not find the action to be anti-competitive as reflected in the fact that the CA imposed neither any penalty nor any conditions on the action.

For actions satisfying this second condition, the date of the decision could be later than, the same as or earlier than the date at which the action came to an end and, indeed, in some instances there was no stop date for the action which was presumed to be still active at the time the data were gathered – 2010. Of the 32 cases, 23 fell into this first category and had an average duration of 5.8 years. Given that some of them could have continued beyond 2010, in what follows we will take it that the natural life of actions is six years. That is $L = 6$.[21] For those cases where the decision was reached after the action had come to an end, the decision came 3.6 years after the action had ceased. So $L_a = 9.6 \Rightarrow L_a - L = 3.6$.

The second subgroup of cases was actions which had ended through the actions of the CA before they reached their natural life. An action was put in this category if the date at which the action came to an end was the same as that when the CA made its decision and the CA decided that the action was anti-competitive and imposed either a penalty or certain conditions that effectively brought the action in its original form to an end. There were nine actions in this category and their average duration was 5.1 years. So $L_b = 5.1$.

The implied values of λ_a, λ_b depend on the assumed interest rate. From (19) a central value of $r = 0.05$ would imply that: $\lambda_a = 0.835$; $\lambda_b = 0.868$; a value of $r = 0.025$ would imply $\lambda_a = 0.914$; $\lambda_b = 0.859$; while a value $r = 0.1$ would imply $\lambda_a = 0.698$; $\lambda_b = 0.885$. So the two parameters move in different directions with the interest rate, with λ_a being the more sensitive of the two.

As 20 of the 32 actions were found to be anti-competitive it follows that the average probability of an action being judged anti-competitive conditional on being investigated is $\overline{p} = \frac{5}{8} = 0.625$.[22]

[21] This is consistent with the evidence for cartels. In the sample of cartels used by Boyer and Kotchoni (2011) the average length of life of a cartel is nine years, though Allain *et al.* (2011) use a figure of six years in their calculation. Connor (2011) reports that other studies put cartel life between two and eight years, while, using his sample, he calculated a mean duration of seven years and a median duration of just under five.

[22] This is arguably compatible with the evidence for cartels. Boyer and Kotchoni (2011) show that the average probability of being disallowed or convicted is 0.66 for the entire sample used by Bulotova and Connor (2006) but that it is 0.72 for the more restricted set of cartels they use for their analysis. The somewhat lower figure of 0.625 for abuse of dominance cases could be argued to reflect the less cut-and-dried nature of the offence.

The final parameters we need to calculate are χ_a and χ_b. To do so we exploit two final pieces of information from our data set

- In our sample, in 21 of the 32 cases a decision was reached in less than six years of the start of the action, so $\chi_b/(\chi_a + \chi_b) = \frac{21}{32} = 0.656 \approx \frac{2}{3}$.
- The maximum time lapse between an action's starting and a decision being reached was 15 years.[23] So in what follows we will assume that if a decision has not been made on an action within 15 years of its initiation then the action will have escaped the attention of the CA.

As reported in Allain *et al.* (2011), a number of studies suggest that the annual probability of detection is around 0.15.[24] Assuming this was also the annual probability of a decision being reached on a case, then the probability that a decision would be made within 15 years would be $\chi_a + \chi_b = 0.913$,[25] while the probability that it would be reached within six years would be $\chi_b = 0.623$. This implies a value of $\sigma = \chi_b/\chi = 0.682$, which is somewhat higher that the the value of $\sigma = 0.66$ that we observe on our data set.

However, it is arguable that, given the delay in carrying out an investigation and reaching a decision, the annual probability of reaching a decision should be lower than the annual probability of detection alone. Assuming that the annual probability of a decision being reached takes the slightly lower value of 0.14, then we get $\chi = \chi_a + \chi_b = 0.896$; $\chi_b = 0.596$ and so we get the observed value of $\sigma = 0.66$. Consequently in what follows we will assume an annual probability of decision-making of 0.14 and that $\chi_a + \chi_b = 0.9; \chi_b = 0.6 \Rightarrow \chi_a = 0.3$.

Putting all this together, it turns out that with our central value of $r = 0.05$ and with $L_b = 5.1; L = 6; L = 9.6; \overline{\rho} = 0.625; \chi_a = 0.3; \chi_b = 0.6$ then, from the formula given in (22) the value of the adjustment factor is 0.74.

To assess how sensitive this adjustment factor is to the assumed values of some of the parameters, the calculation was performed for three values of the interest rate, namely $r = 0.025,\ 0.05,\ 0.1$ and three values of the probability of conviction $\overline{\rho} = 0.625,\ 0.65,\ 0.7$ where the latter reflect values reported in other studies. The results are shown in Table 1.

The conclusion is that the adjustment factor does not seem very sensitive to either of these two parameters.

[23] Again this is not inconsistent with evidence from the literature on cartels. For example, commenting on the Saint-Gobain Car Glass Cartel, Stephan (2009) reports 'the French glass producer Saint-Gobain was fined a staggering €896 million for its involvement in the infringement. Typically, this fine was imposed more than a decade after the anti-competitive behaviour was first instigated and some nine years after the infringement ceased'.

[24] An early paper by Bryant and Eckard (1991) put the figure between 0.13 and 0.17. Similar figuress have been reported by Combe *et al.* (2008), Connor (2011) and Ormosi (2011). Of course, although a widely used assumption, there is in general no reason to think that the annual probability of detection is constant. However, given that we do not observe how many actions go undetected, we need such an identifying restriction to calculate χ_a, χ_b and hence the fraction of undetected actions, which our calculations suggest is 10%.

[25] This follows from the formula that if q is the (constant) annual probability of a decision being reached, then the probability of a decision being reached within T years is $1 - (1 - q)^T$.

Table 1

Sensitivity of Adjustment Factor

	$r = 0.025$	$r = 0.05$	$r = 0.1$
$\overline{\rho} = 0.625$	0.72	0.74	0.78
$\overline{\rho} = 0.65$	0.72	0.74	0.77
$\overline{\rho} = 0.7$	0.71	0.73	0.77

2.2. *Conventional Formula*

If we use the Bulotova and Connor (2006) value of $\theta = 0.3$ for the price overcharge, then, with $\chi_b = 0.6$ and $\overline{\rho} = 0.625$, the optimal penalty as given by the conventional formula (17) would be $\widetilde{\varphi}^{CS} = 0.62$. If instead we used the Boyer and Kotchoni (2011) value for the overcharge of $\theta = 0.175$, this would produce a figure of $\widetilde{\varphi}^{CS} = 0.40$. Using a higher figure for the probability of conviction that has been found in cartel cases, namely, $\overline{\rho} = 0.7$, would lower these figures to $\widetilde{\varphi}^{CS} = 0.55$ and 0.35[26] respectively.

Combining the conventional formula and the adjustment factor we see that the optimal penalty would lie between

- 25% and 31% if the overcharge were 17.5%;
- 40% and 48% if the overcharge were 30%.

So we conclude that:

(*i*) the conventional approach ignores important features relating to the timing of decisions by competitions authorities and consequently produces a figure for the optimal penalty that is too high, the correct figure being around 75% of that generated by the conventional formula;

(*ii*) the optimal penalty is sensitive to the price overcharge, but would be around 30% if we use Boyer and Kotchoni (2011) figure which corrects for the biases in the Bulotova and Connor (2006) figures.

3. The Impact of Tougher Penalty Regimes on the Price Overcharge: Non-deterred Cartels

In this Section and in the next we derive comparative static predictions about the effect of the toughness of the penalty regime on the cartel overcharge. Throughout, we simplify by assuming that

(*i*) production takes place under constant average and marginal costs, c,

(*ii*) the counterfactual is perfectly competitive and

(*iii*) there are no efficiency effects of cartels.

In this Section we also neglect the deterrence effects of penalties – we return to this in Section 5. If a cartel is detected and convicted, it will be liable to a penalty, which, in

[26] Given that, for example, the Office of Fair Trading and the European Commission currently impose baseline penalties of 10% of revenue, these figures illustrate how Connor and Lande (2012) come to the conclusion that, to generate effective deterrence, penalties should be at least quintupled.

general, can have a fixed part and a variable part, where the latter can be a percentage of the profits or revenues of the colluding firms. In this Section the fixed part plays no role so we assume it is zero. A crucial issue is whether the penalty is imposed on profits or revenue. We distinguish a number of cases.

3.1. *Toughness of Antitrust Regime Unrelated to Overcharge*

The question we are interested in is how, anticipating the probability of being investigated, found to be anti-competitive, and penalised, the cartel will set its profit-maximising price and output.

3.1.1. *Case A: Fines on revenues*

If the fine is levied on revenue then, from (18), expected cartel profits are as follows:

$$\Pi(Q) = [1 - \beta(1 + \psi)]\, R(Q) - (1 - \beta)cQ, \tag{23}$$

where $\beta = \overline{\rho}\chi_b(1 - \lambda_b)$ is a measure of the probability of effective enforcement, and

$$\psi = \varphi\left(\frac{\frac{\chi_a}{\chi_b}\lambda_a + \lambda_b}{1 - \lambda_b}\right)$$

measures the penalty rate – both being scaled to reflect the factors relating to timing of investigations introduced in subsection 1.3.

We can re-write (23) as

$$\Pi(Q) = (1 - \beta)[(1 - \alpha)R(Q) - cQ], \tag{24}$$

where $\alpha = \beta\psi/(1 - \beta)$, is increasing in both β and ψ and serves as a combined measure of the toughness of the penalty regime. Assuming declining marginal revenue – the usual second-order condition for profit maximisation – then, as noted first in Bageri, Katsoulacos and Spagnolo (2013), it is clear that:

(*i*) the cartel output (price) is lower (higher) if the penalty is imposed on revenue than if it is imposed on profits;
(*ii*) if the penalty is imposed on revenue, a tougher penalty regime – higher value of α – will result in a higher cartel price/overcharge.

The intuition is straightforward. With a penalty applied to revenue, a profit maximising firm will seek to reduce the penalty, which it does by reducing revenue. But since, with positive marginal costs, marginal revenue is positive, to reduce revenue it reduces output and raises price. The tougher the penalty the greater is this effect.

3.1.2. *Case B: Fines on profits*

If, instead, the fine is levied on profits, then from (23) the expected cartel profits are given by:

$$\Pi(Q) = [1 - \beta(1 + \psi)][R(Q) - cQ] = (1 - \beta)(1 - \alpha)[R(Q) - cQ]. \tag{25}$$

Clearly the optimal cartel output is the (un-distorted) monopoly output that the cartel would choose in the absence of penalties, and consequently the cartel price/overcharge is unaffected by the toughness of the penalty regime.

So we have the following:

PROPOSITION 1. *If the fine is proportional to profit the toughness of the penalty regime does not affect the cartel overcharge. If it is proportional to revenue, a tougher penalty regime – higher value of* α *– will result in a higher cartel price/overcharge.*

Both parts of this result are inconsistent with the conclusions of Bulotova and Connor (2006). In part this could be because, as indicated in the introduction, the empirical evidence supporting this conclusion is rather limited. There are some comparisons across time and across jurisdictions but we lack a wealth of studies that carefully control for all the potential explanatory factors that might be driving overcharges – e.g. globalisation. It is possible that when these are included the empirical results would be consistent with our prediction. However, it may also be that this is just too simple a model, so in what follows we extend the model in a number of directions.

3.2. *Toughness of Antitrust Regime Endogenous*

One possible explanation for the above finding is that firms taking anti-competitive actions do not expect that the probability of being caught or the penalty imposed if caught might depend on the extent of the price overcharge. To allow for this possibility, assume, first, that the probability of effective enforcement is a strictly increasing function of the cartel price/overcharge,[27] p, and that the function $\beta(p)$ satisfies the equation:

$$\beta(1) = 0, \quad \text{and} \quad \forall p \geq 1, \quad \beta(p) < 1; \quad \beta'(p) > 0; \quad \beta''(p) < 0.$$

3.2.1. *Case A: Fines on revenue*

When fines are levied on revenues then, from (23), expected profit becomes as follows:

$$\Pi(Q) = \{1 - \beta[p(Q)](1 + \psi)\}\ R(Q) - c\{1 - \beta[p(Q)]\}\ Q, \tag{26}$$

so profit maximisation implies that:

$$(1 - \beta)[(1 - \alpha)R'(Q) - c] - (\beta' p'\{[R(Q) - cQ] + \psi R(Q)\}) = 0. \tag{27}$$

The second term on the LHS of (27) is negative and so implies that the cartel sets a higher output and lower price than it would have done if the probability of successful enforcement were independent of the cartel price. However, because of the distortion implied by the first term on the LHS of (27) this could still be consistent with marginal revenue being above marginal cost and so the effect of the penalty's being levied on revenue could still be to push price above the monopoly price. Furthermore,

[27] Houba *et al.* (2010, 2012) also explore the possibility that penalties might be related to the cartel's price/overcharge.

$$\Pi_{Q\psi} = -\beta[R'(Q) - c] - \beta' p' R(Q). \tag{28}$$

Given that $p' < 0$ and $\beta' < 0$, this expression is unambiguously positive if $R'(Q) - c < 0$ and so the cartel sets a price below the monopoly price but otherwise the sign depends on the relative magnitude of the two terms.

PROPOSITION 2.

(*i*) *If the penalty is imposed on revenue and if the probability of effective enforcement is greater the higher is the price, then price will be lower than it would have been had the probability of effective enforcement been unaffected by the price overcharge.*

(*ii*) *If price is below the monopoly price then an increase in the penalty unambiguously lowers the price overcharge. Otherwise, it could still be the case that tougher penalties are associated with higher price overcharges.*

3.2.2. *Case B: Fines on profits*

From (25), expected profit becomes as follows:

$$\Pi(Q) = \{1 - \beta[p(Q)](1 + \psi)\}[R(Q) - cQ], \tag{29}$$

and is therefore maximised when:

$$(\{1 - \beta[p(Q)](1 + \psi)\}[R'(Q) - c]) - \{\beta' p'(1 + \psi)[R(Q) - cQ]\} = 0, \tag{30}$$

and so, as the second term in braces on the LHS of (30) is negative, it must be the case that the cartel produces where $R'(Q) - c < 0$ – i.e. marginal revenue is less than marginal cost, and so operates with a higher output, lower price, than in the case where the probability of investigation is unaffected by the price overcharge. Furthermore, it is easy to see from (30), that, because $R'(Q) - c < 0$, it follows that:

$$\Pi_{Q\psi} = -\beta[R'(Q) - c] - \beta' p'[R(Q) - cQ] > 0, \tag{31}$$

so a higher penalty will cause the cartel to increase its output and lower its price/overcharge. So we have the following:

PROPOSITION 3.

(*i*) *If the penalty is imposed on profits and if the probability of effective enforcement is greater the higher is the price, then price will be lower than the monopoly price that would have been set had the probability of effective enforcement been unaffected by the price overcharge;*

(*ii*) *an increase in the penalty unambiguously lowers the price overcharge.*

While this result is consistent with the Bulotova and Connor (2006) finding that tougher antitrust enforcement lowers the price overcharge, in practice CAs levy penalties on revenue and not on profits and, as we have seen above in that case it could still be true that tougher penalties are associated with higher price overcharges.

In the above analysis it is the probability of effective enforcement that depends on the price overcharge. If the penalty (as a proportion of revenue or profits) depends on price overcharge then essentially the same results go through – see Katsoualcos and Ulph (2012*b*).

4. The Impact of Tougher Penalty Regimes on the Price Overcharge: Deterrence Effects

To introduce deterrence effects in the analysis stated above we examine a less general model than we used in previous Section. Specifically, we:

(*i*) restrict attention to the realistic case where the penalty is imposed on revenue;
(*ii*) introduce a fixed component to the penalty, so the fine is now

$$\psi R(Q) + \kappa, \quad \psi > 0, \quad \kappa > 0;$$[28]

(*iii*) use the demand function $p = (1 + \varepsilon) - Q, \quad \varepsilon > 0$ that we introduced – (3) – in Section 2. As we noted there, the parameter ε provides a one-dimensional measure of the degree of competitiveness of different *industries* and the larger is ε the more uncompetitive is the industry.

Thus, expected cartel profits are given by:

$$\Pi(Q) = (1 - \beta)\{(1 - \alpha)[(1 + \varepsilon)Q - Q^2] - Q\} - \beta\kappa, \tag{32}$$

where, as before, $\alpha = \beta\psi/(1 - \beta)$ is increasing in both β and ψ and provides a measure of the toughness of the penalty regime. It is easy to see[29] that this implies that the profit-maximising output will be positive iff:

$$\varepsilon > \underline{\varepsilon} = \frac{\alpha}{1 - \alpha} > 0, \tag{33}$$

in which case the profit-maximising output and price are as follows:

$$Q^* = \frac{1}{2}\left(\varepsilon - \frac{\alpha}{1 - \alpha}\right); \quad p^* = \frac{1}{2}\left[(1 + \varepsilon) + \frac{1}{1 - \alpha}\right], \tag{34}$$

and so, as we know from Result 1, the price set by non-deterred cartels, is a strictly increasing function of the penalty toughness parameter, α, as well as the inverse elasticity, ε.

The maximum profits made by the cartel are given by

$$\Pi^*(\varepsilon; \beta, \psi, \kappa) = \frac{1}{4}(1 - \beta)(1 - \alpha)\left(\varepsilon - \frac{\alpha}{1 - \alpha}\right)^2 - \beta\kappa, \tag{35}$$

and so a cartel will form only if these are positive, which is true iff

$$\varepsilon > \overline{\varepsilon} = \frac{\alpha}{1 - \alpha} + 2\sqrt{\frac{\beta\kappa}{(1 - \beta)(1 - \alpha)}} > \underline{\varepsilon}, \tag{36}$$

in which case the cartel's price and output are given by (34). Notice that $\overline{\varepsilon}$ is a strictly increasing function of the penalty regime parameters (ψ, κ, β).

We can now investigate the implications of changes in the penalty regime parameters for the average price overcharge taking into account the effect of the changes in the penalty regime on deterrence – i.e. on the number of cartels actually formed.

[28] Alternatively, this could be thought of as some other form of fixed cost.
[29] For a more detailed exposition of the section, see Katsoualcos and Ulph (2012*b*).

4.1. *Antitrust Regime Independent of Cartel Price/Overcharge*

Assume first that none of the penalty regime parameters depends on the price overcharge – the antitrust regime is exogenous.

4.1.1. *Effect of an increase in* κ

This has no direct effect on the price overcharge of any cartel that is formed, however, it increases $\bar{\varepsilon}$ and so fewer cartels form and those that do form come from less competitive environments, so, on average, the price overcharge will be increased.

4.1.2. *Effect of an increases in* β *and/or* ψ

As noted in Result 1 this has the direct effect of increasing the price/overcharge of those cartels that form but also has the indirect effect of increasing $\bar{\varepsilon}$, so fewer cartels form and those that do form come from less competitive environments, thus reinforcing the direct effect. So unambiguously the average price overcharge is increased.

So we reach the following overall conclusion:

PROPOSITION 4. *With endogenous deterrence, an increase in the toughness of the antitrust regime (increase in parameters* β, ψ, κ*) will lead to an increase in the average price overcharge, as fewer cartels form, and those that do form come from environments that are, on average, less competitive. This positive indirect deterrence effect reinforces whatever direct effects the tougher penalty regime might have on the cartel overcharge – Result* 1.

4.2. *Antitrust Regime Depends on the Overcharge*

We saw in the previous Section that if penalties are imposed on revenue then if either the probability of effective enforcement, β, or the factor relating the penalty to revenue, ψ, depend on the cartel price/overcharge then although this will lead the cartel to set a lower price than it might otherwise have done, and although this could lead to the possibility that a tougher regime results in a lower overcharge, nevertheless the forces at work in Result 1 were still in play and it is possible that the tougher regime results in a higher overcharge.

In this section, we explore the implications of allowing the fixed component of the penalty,[30] κ, to depend on the price overcharge. So assume now that κ is made up of a fixed part and a part that increases with the price overcharge – decreases with cartel output. Specifically let us suppose that the penalty regime can be characterised by

$$\kappa = \kappa_0 - \kappa_1 Q. \tag{37}$$

An increase in κ_1 toughens the penalty regime by making the penalty that the cartel pays more sensitive to the price overcharge: lowering Q, thus increasing the overcharge, increases the penalty more, the greater is κ_1.

It is straightforward to show that the critical value of the inverse elasticity above which a cartel will form is given by

[30] That is, the component of the penalty that does not depend on revenue.

$$\bar{\varepsilon} = \frac{\alpha}{1-\alpha} - \frac{\xi}{1-\alpha} + 2\sqrt{\frac{\beta\kappa_0}{(1-\beta)(1-\alpha)}}, \tag{38}$$

and that the profit-maximising output and price of those cartels that do form are given by:

$$Q^* = \frac{1}{2}\left(\varepsilon - \frac{\alpha}{1-\alpha} + \frac{\xi}{1-\alpha}\right); \quad p^* = \frac{1}{2}\left[(1+\varepsilon) + \frac{1}{1-\alpha} - \frac{\xi}{1-\alpha}\right], \tag{39}$$

where

$$\xi = \frac{\beta\kappa_1}{(1-\beta)}. \tag{40}$$

It follows that as κ_1 increases, this has two effects: it has a direct effect of reducing the price of those cartels that do form (from (39); but it also has the indirect effect of inducing more cartels to form and, moreover, they come on average from more competitive environments (from (38)), which causes the average overcharge of observed cartels to fall. The two effects reinforce one another so, on average, the cartel overcharge falls – which is the Bulotova and Connor result. So we have the following:

PROPOSITION 5. *An increase in the toughness of the penalty regime, that arises through raising the rate at which the fixed component of the penalty falls with cartel output, will unambiguously cause the average observed cartel overcharge to fall. It has both a direct incentive effect that induces those cartels that form to set a lower price/overcharge and an indirect effect of inducing more cartels to form from environments that are, on average, more competitive.*

We can summarise this Section by saying that for the range of cases we have examined, the deterrence effects of tougher antitrust regimes reinforce their direct impact on the prices set be non-deterred cartels. While we can now find one dimension of the penalty regime – the extent to which any fixed penalty is related to the price/overcharge set by the cartel – where a toughening of the regime unambiguously results in a lower cartel overcharge, for many other dimensions it remains the case that a tougher regime increases the overcharge.

5. Conclusions

The analysis given above provides a number of extensions to the theory of antitrust fines which we have used, with existing and new data sets, to contribute to a better understanding of the current fining policies of CAs. In particular, the analysis extends the theory linking cartel overcharges to optimal fines by introducing a number of additional considerations that are important in fine setting, specifically, legal uncertainty; the fact that CAs sometimes intervene and terminate antitrust actions before they have come to a natural end; and the fact that penalties are often imposed after the antitrust action has terminated. We quantify the resulting optimal penalty, drawing in part on existing empirical evidence but also on a new database of antitrust actions prosecuted under Article 102. We show that the combined effect of the second and third of the above missing factors depends

crucially on what figure is used for the overcharge. Overall our conclusion supports that of Allain *et al.* (2011) – that existing penalties are within the range supported by calculations of optimal penalties.

We also examine the reverse issue of how the toughness of the antitrust regime affects the level of cartel overcharges. We examine this issue under a very wide range of different circumstances and we show that the effects are highly ambiguous, thus again questioning some of the recent empirical findings on this issue. We conclude that there is no good theoretical ground for believing that a tougher regime will necessarily lead to lower overcharges.

Appendix A: Data Appendix

In this Appendix, (Tables A1, A2) we set out for each of the 32 cases in our sample the data relating to:

(*i*) the timing of the actions;
(*ii*) the timing of the relevant CA's decisions on those actions;
(*ii*) the nature of the relevant CA's decision.

We group the data into two subsets: those for whom, using the criteria set out in the text, it could be said that the action lasted its natural life; those for whom it could be said the action was stopped by the CA before it reached its natural life.

Table A1

Cases that Lasted Their Natural Life

Dates of action	Decision date	Decision	Duration of action	Time to penalty
2002–4	2006	F	3	2
2001–5	2005	N	5	n/a
1993–9	2005	F	7	6
1995–9	2005	F	5	6
2001–5	2009	F	5	4
2001–6	2009	F	6	3
2002–6	2006	N	5	n/a
1997–2001	2005	N	5	n/a
2002–5	2005	N	4	n/a
1999–2000	2007	F	2	7
2004–6	2009	N	3	n/a
1999–2003	2004	N	5	n/a
2001–3	2006	N	3	n/a
1997–2000	2004	F	4	4
1997–2000	2006	F	4	6
2002–10	2004	N	9	n/a
2002–10	2004	N	9	n/a
2002–10	2006	N	9	n/a
1991–2003	2006	F	13	3
2002–10	2006	N	9	n/a
2000–10	2010	N	11	n/a
2005	2006	F	1	1
2005–9	2010	F	5	1
Cases: 23		Fines: 11	Average = 5.8	Average = 3.6

Notes. Notation regarding decisions: F, fine imposed; C, conditions imposed; N, no fine, no conditions.

Table A2

Cases Terminated Before End of Natural Life

Dates	Decision date	Decision	Duration
2005–7	2007	C	3
2004–5	2005	C	2
2000–5	2005	C	6
2001–4	2004	F	4
1999–2009	2009	F	11
1999–2004	2004	F	6
2007–10	2010	C	4
2002–5	2005	C	4
1999–2004	2004	F	6
No of Cases = 9			Average = 5.1

Athens University of Economics and Business
University of St Andrews and Scottish Institute for Research in Economics (SIRE)

References

Allain, M.-L., Boyer, M., Kotchoni, R. and Ponssard, J.-P. (2011). 'The determination of optimal fines in cartel cases: The myth of underdeterrence', Discussion Paper 2011s-34, CIRANO.

Bageri, V., Katsoulacos, Y. and Spagnolo, G. (2013). 'The distortive effects of antitrust fines based on revenue', Economic Journal, vol. 123(572), pp. F545–57.

Becker, G. (1968). 'Crime and punishment: an economic approach', *Journal of Political Economy*, vol. 76(2), pp. 169–217.

Boyer, M. and Kotchoni, R. (2011). 'The econometrics of cartel overcharges', Discussion Paper 2011s-35. CIRANO.

Bryant, P.G. and Eckard, E.W. (1991). 'Price fixing: the probability of getting caught', *Review of Economics and Statistics*, vol. 73(3), pp. 531–6.

Buccirossi, P. and Spagnolo, G. (2007). 'Optimal fines in the era of whistleblowers: should price fixers still go to Prison?', in (V. Goshal and J. Stennek, eds.), *The Political Economy of Antitrust*, pp. 81–122, North Holland: Elsevier.

Bulotova, Y. and Connor, J. (2006). 'Cartel overcharges: survey and meta-analysis', *International Journal of Industrial Organisation*, vol. 24(6), pp. 1109–37.

Carlton, D.W. (2007). 'Does antitrust need to be modernized?', Discussion Paper 07-3, Economic Analysis Group.

Combe, E., Monnier, C. and Legal, R. (2008). 'Cartels: the probability of getting caught in Europe', Paper no 12, BEER.

Connor, J.M. (2011). 'Cartel detection and duration worldwide', *CPI Antitrust Chronicle*, September, vol. 1, pp. 1–10.

Connor, J.M. and Lande, R.H. (2005). 'How high do cartels raise prices? Implications for optimal cartel fines', *Tulane Law Review*, vol. 80(4), pp. 513–70.

Connor, J.M. and Lande, R.H. (2006). 'The size of cartel overcharges: implications for US and EU fining policies', *Antitrust Bulletin*, vol. 51(4), pp. 983–1022.

Connor, J.M. and Lande, R.H. (2008). 'Cartel overcharges and optimal cartel fines', in (S. Waller, ed.), *Issues in Competition Law and Policy*, vol. 3, Chapter 88, pp. 2203–18.

Connor, J.M. and Lande, R.H. (2012). 'Cartels as rational business strategy: new data demonstrates that crime pays', mimeo, available at http://ssm.com/abstract-1917657 (last accessed 10 August 2013).

Houba, H., Motchenkova, E. and Wen, Q. (2010). 'Antitrust enforcement with price-dependent fines and detection probabilities', *Economics Bulletin*, vol. 30(3), pp. 2017–27.

Houba, H., Motchenkova, E. and Wen, Q. (2012). 'Antitrust enforcement and marginal deterrence', mimeo, Tinbergen Institute, Amsterdam.

Katsoulacos, Y. and Ulph, D. (2012*a*). 'Legal uncertainty and the choice of enforcement procedures', mimeo, www.cresse.info (last accessed: 10 April 2013).

Katsoualcos, Y. and Ulph, D. (2012*b*). 'Effects of antitrust fines on cartel overcharges', mimeo, University of St Andrews.

Katsoulacos, Y., Ulph, D. and Metsiou, E. (2011). 'Competition authority substantive standards and social Welfare', mimeo, www.cresse.info (last accessed: 20 August 2013).

Landes, W.M. (1983). 'Optimal sanctions for antitrust violations', *University of Chicago Law Review*, vol 50(2), pp. 652–78.

Motchenkova, E. (2008). 'Determination of optimal penalties for antitrust violations in a dynamic setting', *European Journal of Operational Research*, vol. 189(1), pp. 296–291.

Motchenkova, E. and Kort, P. (2006). 'Analysis of current penalty schemes for violation of antitrust laws', *Journal of Optimisation Theory and Apllications*, vol. 128(3), pp. 431–51.

Ormosi, P. (2011). 'How big is a tip of the iceberg? A parsimonious way to estimate cartel detection rate', CCP Working Paper 11-6, University of East Anglia.

Polinsky, M. and Shavell, S. (2000). 'The economic theory of public enforcement of law', *Journal of Economic Literature*, March, vol. 38(1), pp. 45–76.

Salop, S.C. (2010). 'Question: what is the real and proper antitrust welfare standard? Answer: the true consumer welfare tandard', *Loyola Consumer Law Review*, vol. 22(3), pp. 336–53.

Stephan, A. (2009). 'Hear no evil, see no evil: why antitrust compliance programmes may be ineffective at preventing cartels', CCP Working Paper 09-09, University of East Anglia.

Stigler, G.J. (1970). 'The optimum enforcement of laws', *Journal of Political Economy*, vol. 78(3), pp. 526–36.