Review of Industrial Organization (2022) 61:341–371
https://doi.org/10.1007/s11151-022-09877-8

Cartel Penalties Under Endogenous Detection

Berkay Akyapi[1] · **Douglas C. Turner**[1]

Accepted: 21 July 2022 / Published online: 6 August 2022

Abstract

We compare the performance of cartel penalties that are proportional to a cartel's revenue and cartel penalties that are proportional to the difference between the cartel price and the competitive price: the overcharge. Prior literature has shown that when the probability of cartel detection does not depend on the cartel price, penalties that are based on a cartel's overcharge generate greater total surplus and consumer surplus than do penalties that are based on a cartel's revenue. In contrast, we find that when the probability of detection depends on the cartel price, penalties that are based on revenue can generate greater total surplus and consumer surplus.

Keywords Antitrust · Cartel penalties · Cartels · Collusion

JEL Codes L4 · K2 · C7

✉ Douglas C. Turner
douglasturner@ufl.edu

Berkay Akyapi
berkayakyapi@ufl.edu

1 Department of Economics, University of Florida, PO Box 117140, Gainesville, FL 32611, USA

1 Introduction

The recent discovery of a number of large and harmful cartels demonstrates that illegal price fixing remains a significant concern for antitrust authorities.[1] To deter cartels, antitrust authorities penalize firms that are proven guilty of illegal collusion. The structure of the penalties that are imposed by an antitrust authority affects both cartel formation and pricing.[2]

Revenue-based and aggregate overcharge-based penalties are two often-proposed types of cartel penalties. A revenue-based penalty is a multiple of cartel revenue. An aggregate overcharge-based penalty is a multiple of the difference between the collusive price and the competitive price (the overcharge) multiplied by some base volume of sales. Revenue-based penalties are the current practice in all major jurisdictions.[3]

Recent research (Bageri et al., 2013; Katsoulacos et al., 2015) suggests that cartel penalties that are based on revenue generate a lower level of social surplus than do penalties that are based on the cartel aggregate overcharge.[4] This is the case because revenue-based penalties induce cartels to price above the monopoly level in order to reduce revenue and thereby reduce the potential penalty. Intuitively, because a cartel that is maximizing profits already prices on the elastic part of its demand curve, the cartel *increases* its price so as to reduce revenue. Conversely, aggregate overcharge-based penalties induce cartels to price *below* the monopoly level so as to reduce the measured overcharge. These studies assume a constant and exogenous probability of cartel detection.

In this study, we compare revenue-based and overcharge-based cartel penalties while allowing the probability of detection to depend on the cartel price.[5] We consider two specifications for the probability of detecting a cartel: In the first specification, the probability of detection is increasing in the cartel (per-unit) overcharge.[6] In

[1] For example, Taro Pharmaceuticals U.S.A., Inc. and Sandoz, Inc. were fined \$205.7 million and \$195 million respectively for involvement in a price fixing conspiracy (*U.S. vs. Taro Pharmaceuticals U.S.A, Inc.* (E.D. Pa., No. 2:20-CR-00214-RBS 7/23/20) and *U.S. vs. Sandoz Inc.* (E.D. Pa., No. 2:20-CR-00111-RBS 3/02/20)). Starkist Co. admitted to conspiring with other major canned tuna suppliers to “fix, raise, and maintain the prices of packaged seafood products” and agreed to a fine of \$100 million (*U.S. vs. Starkist Co.* (N.D. Ca., No. 3:18-CR-00513-EMC 11/14/18)).

[2] See Block et al. (1981) and Bolotova et al. (2009) for empirical evidence that expected antitrust penalties affect the cartel price.

[3] International Competition Network Cartels Working Group, “Setting Fines for Cartels in ICN Jurisdictions”, Report to the 16th Annual Conference, Porto (2017).

[4] See also: Bageri and Katsoulacos (2014), Katsoulacos et al. (2018), Katsoulacos et al. (2019a), Katsoulacos et al. (2019b), Houba et al. (2010), Garrod and Olczak (2018) and Dargaud et al. (2016).

[5] A number of cartel cases suggest that the probability of cartel detection is endogenous and depends on the cartel price. In the stainless steel cartel case, an investigation began as a result of buyers reporting a suspiciously large increase in prices to the European Commission (Levenstein et al., 2004). Anomalous pricing also created suspicions of collusion in the Nasdaq case (Christie and Schultz, 1995). In the auto parts cartel, the cartel members specifically set prices in order to avoid detection by buyers *(In re Automotive Parts Antitrust Litigation* (E.D. Mich,. No. 12-md-02311, 03/08/17)).

[6] See Block et al. (1981), Houba et al. (2010, 2012, 2015), Katsoulacos and Ulph (2013) and Bos et al. (2018).

the second, the probability of detection is increasing in the *rate* at which the cartel price increases, as in Harrington (2004, 2005). Intuitively, relatively high prices or large price increases are more likely to attract the attention of an antitrust authority and lead to the cartel's detection. For each specification, we analyze and compare cartel pricing, deterrence, total surplus and consumer surplus under revenue-based and overcharge-based penalties.

We find that cartel penalties that are based on revenue generate a higher level of total and consumer surplus when the probability of detection is sufficiently sensitive to price under both specifications. Cartels do not charge prices above the monopoly price because such a high price is likely to raise suspicions of collusion and may cause the cartel's detection. Thus, the pricing result of Katsoulacos et al. (2015) does not occur. Instead, cartels set low prices to avoid detection. However, a revenue-based penalty is larger than an aggregate overcharge-based penalty for low cartel prices. As a result, collusion is more difficult to sustain under revenue-based penalties. In addition, cartels that face revenue-based penalties have a stronger incentive to reduce price further so as to avoid detection and the payment of larger revenue-based penalties.

The rest of the article is organized as follows: Sect. 2 presents the model. Section 3 compares the two types of penalties when the probability of detection depends on the cartel overcharge. Section 4 compares the two types of penalties when the probability of detection depends on the rate at which the cartel price increases. Section 5 concludes. All proofs can be found in "Appendix 1".

2 Base Model

The model closely follows Katsoulacos et al. (2015, 2020a, 2020b).[7] $N \geq 2$ symmetric firms compete in prices in each of infinitely many periods and have a common discount factor $\delta < 1$. Firms produce an homogeneous product at a constant marginal cost $c > 0$. Market demand when firm i charges price p_i is $D(p_1, \dots p_N) = N\left[1 - \underline{p}\right]$ where $\underline{p}$ is the lowest price that is charged by firms. Market demand is split evenly among all firms that charge the lowest price. All other firms receive zero demand. Let $D(p) = 1 - p$ denote per-firm demand when all firms charge a common price p. When all firms charge price p, the profit of each firm is $\pi(p) = D(p)\left[p - c\right]$. Demand and cost conditions do not change over time.

Industry suppliers have the opportunity to form an illegal cartel at the beginning of the initial period. If a cartel forms, the firms agree to set a common price. An antitrust authority may detect and penalize a cartel. Cartels can be detected in a variety of ways. A cartel could be detected because of: a report from an internal whistleblower; a complaint from a buyer; or the discovery of evidence during a merger review or separate antitrust investigation. Let $\phi(p_{t-1}, p_t)$ denote the probability of

[7] We analyze and build on the model of Katsoulacos et al. (2015) for comparability and analytical tractability. However, as the primary forces that underly our main results are more general, our conclusions are expected to hold under other demand or cost assumptions.

detection when the price in the current period is p_t and the price in the preceding period is p_{t-1}. We consider two different specifications for ϕ in the following sections. If a cartel is detected, we assume that each firm is prosecuted, convicted and penalized with probability 1.[8]

We compare two penalty structures: revenue-based penalties and overcharge-based penalties. Revenue-based penalties are of the form $x_R(p) = \gamma_R D(p)p$ where γ_R is a positive constant and p is the cartel price during the period of detection.[9] Overcharge-based cartel penalties are of the form $x_O(p) = \gamma_O D(p^{BF})\left[p - p^{BF}\right]$ where p^{BF} is the but-for price and γ_O is a positive constant. The but-for price is the price that would prevail in the absence of collusion. Following Katsoulacos et al. (2015), we assume that the but-for price is the Nash equilibrium price c which implies $x_O(p) = \gamma_O Q_N \left[p - c\right]$ where $Q_N = 1 - c$. Following Motta and Polo (2003), Katsoulacos et al. (2015), Chen and Rey (2013) and Katsoulacos et al. (2020b), we assume that a cartel has a renewed opportunity to form after detection.

A firm will defect from the cartel if the expected present discounted value of the payoff from collusion is less than the payoff from defection. After defection, the market reverts to Nash competition for all future periods. A defecting firm undercuts the cartel price and serves all demand. Defection profit for a firm when all rivals charge price p is[10]

$$\pi^D(p) = \begin{cases} ND(p)\left[p - c\right] & \text{if } p \le p^m \\ ND(p^m)\left[p^m - c\right] & \text{if } p > p^m \end{cases},$$

where p^m is the monopoly price. We assume that cartels can be detected only when all firms charge the collusive price: The cartel cannot be detected in a period where any firm defects or after the breakdown of a cartel.[11] Once a firm has defected, we assume the cartel cannot be detected.[12]

If a cartel forms, the firms set a common price to maximize the expected present discounted value of per-firm profit less penalties, subject to the constraint that no cartel member wishes to defect. Cartel prices satisfy the Bellman equation

[8] We consider leniency programs, which allow firms to report the cartel in exchange for reduced penalties, in the online appendix (see https://www.douglascturner.com/endogenous-detection-online-appendix/). Our results are qualitatively unchanged.

[9] We assume that penalties depend only on the price in the period of detection and do not depend on the length of time that a cartel has operated. See Akyapi and Turner (2021) for an exploration of cartel penalties that depend on duration.

[10] We write the defection profit as a function of the cartel price p. If $p \le p^m$, the defecting firm infinitesimally undercuts the cartel price and the defecting firm's price is $p - \epsilon$. If $p > p^m$, the defecting firm charges a price of p^m.

[11] Motta and Polo (2003), Katsoulacos et al. (2015) and Dargaud et al. (2016) make a similar assumption.

[12] This reflects the fact that, once a cartel has dissolved, evidence of collusion is less likely to be uncovered. Additionally, as the market is now competitive, buyers or other observers are less likely to form or report suspicions of collusion. We explore the possibility of detection after cartel breakdown in the online appendix (https://www.douglascturner.com/endogenous-detection-online-appendix/).

$$\begin{aligned} V_i(p_{t-1}) = \max_{p\in[c,1]} \ & \pi(p) - \phi(p_{t-1},p)x_i(p) + \delta\left[1-\phi(p_{t-1},p)\right]V_i(p) + \delta\phi(p_{t-1},p)V_i(c) \\ \text{s.t.} \quad & \pi(p) - \phi(p_{t-1},p)x_i(p) + \delta\left[1-\phi(p_{t-1},p)\right]V_i(p) + \delta\phi(p_{t-1},p)V_i(c) \\ & \geq \pi^D(p) \end{aligned} \tag{1}$$

where $V_i(p_{t-1})$ denotes the expected present discounted value of the payoff from collusion when the cartel price in the prior period was p_{t-1} and the penalty type is $i \in \{R, O\}$ where R denotes revenue-based penalties and O denotes overcharge-based penalties. We assume that the market is competitive prior to cartel formation so the initial price is the Nash equilibrium price c.[13] We refer to the inequality in Eq. (1) as an incentive compatibility constraint.

Consumers have discount factor $\delta < 1$. The expected present discounted value of consumer surplus (total surplus) under penalty type $i \in \{O, R\}$ when a cartel forms is denoted CS_i (TS_i). CS_i and TS_i are defined and derived in "Appendix 2".

A cartel forms if and only if collusion is sustainable and profitable. Collusion is profitable if $V_i(c) > 0$. Collusion is sustainable if there exists a sequence of prices that satisfy the constraint in (1) in every period. As δ increases, the constraint in (1) is more easily satisfied. The critical discount factor δ_i is the discount factor such that a cartel forms under penalty type $i \in \{O, R\}$ for all $\delta > \delta_i$ and does not form under penalty type $i \in \{O, R\}$ if $\delta < \delta_i$.[14] If a cartel does not form, firms engage in Nash competition in all future periods and earn zero profit.

3 Overcharge (Price-Level) Detection Specification

In this section, the probability of detection depends on the per-unit cartel overcharge: $[p_t - c]$. Since c is exogenous, the probability of detection effectively depends on the price: p_t. A large overcharge—a relatively high price—is more likely to attract the attention of an antitrust authority—either directly, or through the complaints of buyers or other industry observers – and cause the detection of the cartel.

Let

$$\phi(\cdot, p_t) = \min\left\{\alpha_0 + \alpha_1\left[p_t - c\right]^2, 1\right\} \tag{2}$$

where $\alpha_1 > 0$ measures the sensitivity of the probability of detection to the cartel overcharge.[15] This specification is not a function of p_{t-1}—the cartel price in the

[13] When a cartel reforms after detection, we again consider the initial price to be c. This reflects the expectation of competitive market conditions after detection. A large deviation from marginal cost pricing, or a high value of $\phi(c, p)$, indicates a cartel may have reformed. Alternatively, we could assume that the market operates competitively for one period after detection and then a cartel reforms. Results are robust to this alternative assumption.

[14] If a cartel does not form for any discount factor, we set $\delta_i = 1$ as a convention.

[15] The overcharge is squared in Eq. (2) for consistency with prior literature (Harrington, 2005) and analytical tractability: concavity of the maximization problem.

prior period—and depends only on the level of the cartel price in the current period. For example, in industries where a high overcharge (a relatively high price) is considered anomalous and indicative of collusion, α_1 is high. On the other hand, in industries where a high overcharge is unlikely to attract the attention of an antitrust authority and lead to detection, α_1 is low. $\alpha_0 > 0$ represents the probability of detection that does not depend on the cartel price.[16]

Houba et al. (2010, 2012, 2015) also analyze a specification where the probability of detection is increasing in the cartel price, but they do not compare revenue and overcharge-based penalties.[17] We assume $\gamma_R > 1$ and $\gamma_O > 1$ to ensure penalties are large enough that cartels do not choose a price that would cause immediate detection.[18] We also assume $\alpha_0\gamma_R < 1 - c$ and $\alpha_0\gamma_O < 1$ to ensure expected penalties are not so large as to preclude all collusion.[19]

3.1 Cartel Formation Under Price-Level Specification

In this section, we analyze the critical discount factor under revenue and aggregate overcharge-based cartel penalties when the probability of detection is given by Eq. (2). A cartel forms if collusion is sustainable—no cartel member wishes to defect in any period – and profitable.

Theorem 1 $\delta_O = \frac{N-1}{N} + \frac{\alpha_0\gamma_O}{N}$.

$\delta_O < 1$ because $\alpha_0\gamma_O < 1$ by assumption. A cartel forms for any $\delta > \delta_O$. The critical discount factor under overcharge-based penalties does not depend on α_1.

As α_1 increases, cartels must reduce the cartel price to avoid detection (i.e., $p \to c$). However, a reduction in the cartel price does not make collusion unprofitable. This is the case because while a reduction in the cartel price diminishes profit, it also causes a commensurate drop in the aggregate overcharge-based penalty. Cartel profit approaches 0 as $p \to c$ but—as is illustrated in Fig. 1—aggregate overcharge based penalties also approach 0 as $p \to c$. This is the case because both a cartel's profit and the aggregate overcharge-based penalty are tied to the same outcome—the markup $p - c$. In essence, the benefit (cartel profit) and cost (the expected penalty) of collusion decrease in tandem as the cartel price declines. Thus, although the payoff from collusion declines as $\alpha_1 \to \infty$, it remains positive if the cartel charges a sufficiently low price.

[16] Certain mechanisms of cartel detection—such as internal whistleblowers, random auditing, or the uncovering of evidence in a separate antitrust investigation—are likely unrelated to the cartel price. α_0 captures the probability of detection due to these mechanisms.

[17] See also: Block et al. (1981), Katsoulacos and Ulph (2013) and Bos et al. (2018).

[18] If $\gamma_R < 1$ or $\gamma_O < 1$, cartels could profitably collude by ignoring the possibility of raising suspicions of collusion and setting prices such that the cartel is detected with certainty in every period: a price greater than $c + \sqrt{\frac{1-\alpha_0}{\alpha_1}}$. Immediate detection after one period of collusion is inconsistent with empirical evidence (Levenstein and Suslow, 2006). In addition, the examples in footnote 5 suggest that cartels do not ignore the possibility of raising suspicions of collusion.

[19] The continued detection of illegal cartels suggests this is a reasonable assumption.

When we turn to sustainability, a similar intuition holds: As $\alpha_1 \to \infty$, the cartel price falls which reduces profits. However, the payoff from defection is also reduced proportionately. Thus, the incentives to defect are limited as $\alpha_1 \to \infty$, and collusion remains sustainable—provided firms are sufficiently patient: $\delta > \delta_O$. And from Theorem 1, the critical discount factor above which a cartel forms—δ_O—does not depend on α_1.

Theorem 2 $\delta_R \to 1$ *monotonically as* $\alpha_1 \to \infty$.[20]

Under revenue-based penalties, the critical discount factor increases as the sensitivity of the probability of detection to the cartel price increases. When the probability of detection is highly sensitive to price, cartels must set low prices to avoid detection (as was the case under overcharge-based penalties). However, in contrast to the setting with overcharge-based penalties, collusion is unprofitable under revenue-based penalties when the cartel price is low.

To understand why, observe that the per-period profit from collusion—$\pi(p)$—approaches 0 as $p \to c$. However, the penalty to be paid if detected—$\gamma_R pD(p)$—approaches $\gamma_R cD(c) > 0$ as $p \to c$ (see Fig. 1) because, even when the cartel price is low, revenue is strictly positive. Thus, the expected penalty is greater than the per-period profit from collusion when the cartel price is sufficiently low (which occurs when α_1 is sufficiently large)—which implies that collusion is unprofitable and a cartel does not form.[21]

Intuitively, revenue-based penalties more effectively penalize cartels that set low prices because, as is shown in Fig. 1, revenue-based penalties are greater than overcharge-based penalties for low cartel prices. However, overcharge-based penalties are typically more effective at penalizing cartels that set high prices because, as is shown in Fig. 1, overcharge-based penalties exceed revenue-based penalties for high cartel prices. Crucially, when the probability of detection is highly sensitive to the cartel price, cartels must set low prices to avoid detection. Therefore, revenue-based penalties are superior and deter the formation of a greater number of cartels.

3.2 Cartel Pricing Under Price-Level Specification

In this section, we assume that a cartel forms under both penalty types, and we compare cartel prices under revenue and overcharge-based penalties. Note that the cartel's problem under the ϕ specification in Eq. (2) is identical in every period because the probability of detection does not depend on the cartel price in the prior period p_{t-1}. Therefore, the cartel sets the same price in each period.

[20] Recall that no cartel forms for any δ when $\delta_R = 1$.

[21] Theorem 2 can also be understood by analyzing the ratio of a cartel's penalty to its per-period profit: $\frac{x_R(p)}{\pi(p)} = \frac{\gamma_R D(p)p}{D(p)(p-c)} = \frac{\gamma_R p}{p-c}$. As $\alpha_1 \to \infty$ and $p \to c$, this ratio approaches ∞ which implies that the penalty from collusion becomes infinitely larger then per-period profit.

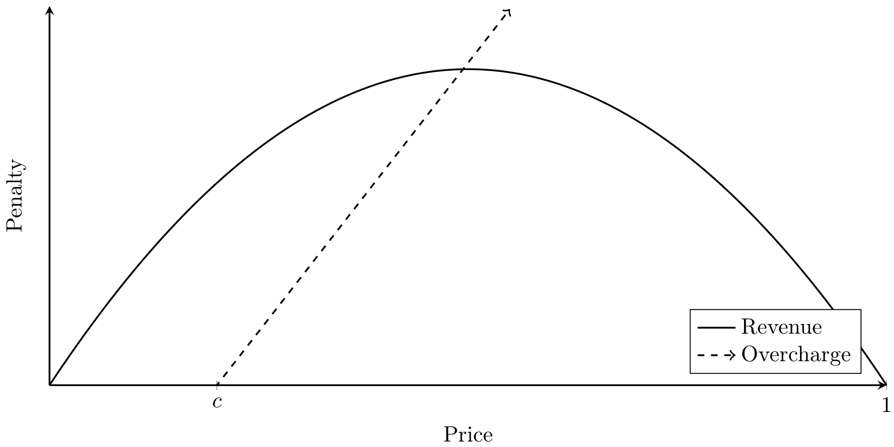

Fig. 1 Revenue-based penalty ($\gamma_R p(1-p)$) and aggregate overcharge-based penalty ($\gamma_O Q_N(p-c)$)

Let p_O denote the price under overcharge-based penalties and let p_R denote the price under revenue-based penalties.[22] Katsoulacos et al. (2015) show that $p_O < p^m < p_R$ if $\alpha_1 = 0$. Intuitively, cartels increase price above the monopoly price in order to reduce revenue under revenue-based penalties and set a price below the monopoly price in order to reduce the overcharge under overcharge-based penalties. However, we show in this section that cartel prices may be lower under revenue-based penalties when $\alpha_1 > 0$.

Theorem 3 *There exists an $\bar{\alpha}_1^L$ such that $p_R < p_O$ if $\alpha_1 > \bar{\alpha}_1^L$ and a cartel forms under both penalty types.*

Theorem 3 implies that prices are lower under revenue-based penalties than under overcharge-based penalties when a cartel forms under both penalty types and the probability of detection is sufficiently sensitive to the cartel overcharge.

As Katsoulacos et al. (2015) show, cartels set a high price under revenue-based penalties—$p_R > p^m$—when $\alpha_1 = 0$. However, when α_1 is sufficiently large, such a high price is likely to raise suspicions of collusion and increase the likelihood of the cartel's detection. Instead, cartels set low prices when α_1 is high so as to reduce the likelihood of detection. When cartel prices are relatively low—close to marginal cost—revenue-based penalties are strictly larger than overcharge-based penalties (see Fig. 1). Faced with the possibility of a relatively large penalty if detected, cartels under revenue-based penalties further reduce price so as to reduce the likelihood of detection and penalization.

Intuitively, the sensitivity of the probability of detection to the price level encourages cartels to set prices in a region where prices are relatively close to marginal

[22] Recall that the price is c if a cartel does not form. Also, note that the cartel price must exceed marginal cost if a cartel forms because collusion must be profitable.

cost—but this is also a region where revenue-based penalties are large. Recognizing this, cartels that face revenue-based penalties are more exposed to detection and, as a result, reduce their price to a level that is below the price that is set under overcharge-based penalties so as to disguise collusion and reduce the likelihood of detection.

3.3 Surplus Under Price-Level Specification

In this section, we compare total surplus and consumer surplus under overcharge-based penalties and revenue-based penalties. There are two cases to consider: $\delta \leq \delta_O$ and $\delta > \delta_O$. First, we consider the case of $\delta \leq \delta_O$. Theorem 2 implies that $\delta_R > \delta_O$ if α_1 is sufficiently high. Therefore, no cartel forms under either penalty type and both penalties generate the same level of total and consumer surplus when $\delta \leq \delta_O$ and the probability of detection is sufficiently sensitive to the cartel price.

Next, we consider the case of $\delta > \delta_O$. When $\delta > \delta_O$, a cartel forms under overcharge-based penalties. However, a cartel does not form under revenue-based penalties when α_1 is sufficiently high, by Theorem 2. Therefore, both total and consumer surplus are greater under revenue-based penalties than under overcharge-based penalties when $\delta > \delta_O$ and the probability of detection is sufficiently sensitive to the cartel price.

Theorem 4 *When $\delta > \delta_O$, there exists an $\tilde{\alpha}_1^L$ such that $CS_R > CS_O$ and $TS_R > TS_O$ if $\alpha_1 > \tilde{\alpha}_1^L$. When $\delta \leq \delta_O$, there exists an $\tilde{\alpha}_1^L$ such that $CS_R = CS_O$ and $TS_R = TS_O$ if $\alpha_1 > \tilde{\alpha}_1^L$.*

In summary, the optimal penalty type depends on α_1. As Katsoulacos et al. (2015) show, overcharge-based penalties generate a greater level of surplus when the probability of detection does not depend on the cartel overcharge ($\alpha_1 = 0$). Conversely, surplus is greater under revenue-based penalties when the probability of detection is sufficiently sensitive to the cartel overcharge.

4 Price-Change Detection Specification

In this section, the probability of detection depends on the difference between the current period price and the price in the preceding period. We assume that large price increases are more likely to lead to detection than is true for smaller increases. As noted above, a large increase in price might create suspicions of collusion among buyers or other industry observers which may be reported to a competition authority.[23] Alternatively, a large and anomalous price increase could directly attract the attention of competition authorities.

[23] See Harrington (2004, 2005) for further discussion.

Let

$$\phi(p_{t-1}, p_t) = \begin{cases} \min\left\{\alpha_0 + \alpha_1 \left[p_t - p_{t-1}\right]^2, 1\right\} & \text{if } p_t > p_{t-1} \\ \alpha_0 & \text{if } p_t \leq p_{t-1} \end{cases} \quad (3)$$

where $\alpha_1 \geq 0$ represents the sensitivity of the probability of detection to the price increase.[24] α_0 represents the probability of detection when the cartel price is constant or decreasing. Harrington (2004, 2005) first introduced and analyzed a similar specification, but those studies do not consider revenue-based penalties. As in the previous section, we assume $\gamma_R > 1$ and $\gamma_O > 1$ to ensure that a cartel does not choose a price that would cause detection with probability 1. In addition, we assume that $\alpha_0 \gamma_R < 1 - c$ and $\alpha_0 \gamma_O < 1$ so as to ensure that the expected penalties are not so large as to preclude all collusion.

4.1 Cartel Formation Under Price-Change Specification

In this section, we analyze the critical discount factor under revenue and overcharge-based cartel penalties when the probability of detection is given by Eq. (3). A cartel forms if collusion is sustainable—no cartel member wishes to defect in any period – and profitable.

Theorem 5 $\delta_O = \frac{N-1}{N} + \frac{\alpha_0 \gamma_O}{N}$.

$\delta_O < 1$ because $\alpha_0 \gamma_O < 1$ by assumption. A cartel forms for any $\delta > \delta_O$. As was the case under the price-level specification, the critical discount factor under overcharge-based penalties does not depend on α_1.

As α_1 increases, cartels must increase the cartel price more gradually to avoid detection: $p_t \to p_{t-1}$. Raising the cartel price more gradually effectively reduces the cartel price in each period. However, the more gradual price increase does not render collusion unprofitable. This is the case because while increasing the cartel price more gradually diminishes profit, it also causes a commensurate drop in the overcharge-based penalty that the cartel faces in each period: The benefit (cartel profit) and cost (the expected penalty) of collusion decrease in tandem as the cartel price rises more slowly. Thus, while the payoff from collusion declines as $\alpha_1 \to \infty$, the payoff from collusion remains positive if the cartel increases price at a sufficiently slow rate.

When we turn to sustainability, a similar intuition holds: As $\alpha_1 \to \infty$, the cartel price falls in each period, which reduces profit. However, the payoff from defection is also reduced proportionately. Thus, the incentives to defect are limited as $\alpha_1 \to \infty$, and collusion remains sustainable if firms are sufficiently patient: $\delta > \delta_O$. The critical discount factor above which a cartel forms—δ_O—does not depend on α_1.

[24] The change in price is squared in Eq. (3) for consistency with prior literature (Harrington, 2004, 2005) and analytical tractability: concavity of the maximization problem.

Theorem 6 $\delta_R \rightarrow 1$ *monotonically as* $\alpha_1 \rightarrow \infty$.[25]

Under revenue-based penalties, the critical discount factor is increasing in the sensitivity of the probability of detection to changes in the cartel price. When the probability of detection is highly sensitive to changes in the cartel price, cartels must increase price slowly in order to avoid detection (as was the case under overcharge-based penalties). However, unlike under overcharge-based penalties, collusion becomes unprofitable under revenue-based penalties when the cartel is forced to increase price slowly.

To see this, first note that gradually increasing the cartel price effectively reduces the price in each period. Next, recall that the per-period profit from collusion—$\pi(p_t)$—approaches 0 as $p_t \rightarrow c$. However the penalty—$\gamma_R p_t D(p_t)$ – approaches $\gamma_R cD(c) > 0$ as $p_t \rightarrow c$ (see Fig. 1) because, even when the cartel price is low, revenue is strictly positive. Thus, the expected penalty is greater than the per-period profit from collusion when the cartel price is sufficiently low.[26]

As $\alpha_1 \rightarrow \infty$, the cartel is induced to price in a region (low prices) where the expected penalty exceeds cartel profit in an increasingly large number of periods. Rapidly increasing the price to a level where collusion is profitable would cause detection. Thus, as α_1 increases, the cartel incurs a per-period loss in increasingly more and more periods; and, for sufficiently large α_1, collusion is altogether unprofitable, and a cartel does not form.

4.2 Cartel Pricing Under Price-Change Specification

In this section, we assume that a cartel forms under both penalty types and compare cartel prices under revenue and overcharge-based penalties. Under the ϕ specification in (3), the Bellman equation in (1) does not readily generate tractable closed-form solutions or analytical results.

However, cartel prices under each penalty type can be determined through numerical solutions. Specifically, we solve the Bellman equation in (1) with the use of value function iteration.[27] We first analyze this issue in an illustrative baseline setting. Next, we consider modifications of the baseline setting to assess the robustness of our findings.

Baseline Setting $\delta = 0.9, c = 0.1, N = 2, \gamma_R = 5, \gamma_O = 3.05, \alpha_0 = 0.05$

The baseline setting represents a duopoly. γ_O and γ_R are chosen to ensure that the cartel penalty is the same under the two penalty types when each member of the cartel charges the monopoly price.[28] We consider a range of α_1 values. Let p_i^t denote

25 Recall that no cartel forms for any δ when $\delta_R = 1$.

26 Per-period profit in period t is $\pi(p_t)$, and the expected penalty in period t is $\phi(p_{t-1}, p_t)x_R(p_t)$. As $p_t, p_{t-1} \rightarrow c$, $\pi(p_t) \rightarrow \pi(c) = 0$ and $\phi(p_{t-1}, p_t)x_R(p_t) \rightarrow \alpha_0 x_R(c) > 0$.

27 See the online appendix (https://www.douglascturner.com/endogenous-detection-online-appendix/) for details.

28 The monopoly price is 0.55. The overcharge-based cartel penalty when $p = 0.55$ is $3.05(.55 - 0.1)D(0.1) \approx 1.24$. The revenue-based cartel penalty when $p = 0.55$ is $5(0.55)D(0.55) \approx 1.24$.

the cartel price under penalty type $i \in \{O, R\}$ in period t. Figure 2 shows that the cartel price is constant and higher under revenue-based penalties than for aggregate overcharge-based penalties when $\alpha_1 = 0$, as is shown in Katsoulacos et al. (2015). However, Fig. 2 also shows that cartel prices are lower under revenue-based penalties in early periods and lower under aggregate overcharge-based penalties in later periods when $\alpha_1 = 5, \alpha_1 = 10$ or $\alpha_1 = 15$.

To understand these results, first note that cartels increase their price gradually as bigger price movements are likely to attract attention and lead to the cartel's detection. As Fig. 1 shows, revenue-based penalties exceed aggregate overcharge-based penalties when the cartel price is low. Thus, a cartel that is detected in one of the first few periods of collusion pays a large penalty under revenue-based penalties and a smaller penalty under overcharge-based penalties. Recognizing this, cartels that face revenue-based penalties increase the cartel price more slowly so as to disguise collusion and avoid the payment of relatively high penalties. Conversely, cartels that face aggregate overcharge-based penalties are less wary of rapidly increasing their price and raising suspicions of collusion because they face a smaller penalty. These considerations cause the cartel price to be lower under revenue-based penalties during the early periods of collusion.

In the later periods of collusion, $p_t^R > p_t^O$. This occurs because the cartel price converges to the price that is derived in Katsoulacos et al. (2015)—the cartel price when $\alpha_1 = 0$—as $t \to \infty$.[29] Increasing the cartel price any further would reduce the payoff from collusion. Thus, $p_t^R > p_t^O$ in later periods when the cartel price has converged to its limiting value. We demonstrate the robustness of our results to alternative parameter configurations in the online appendix.[30]

In summary, the following conclusion arises in every numerical solution we conducted:

Conclusion 1 If α_1 is sufficiently large and a cartel forms under both penalty types, then $p_t^R < p_t^O$ in early periods and $p_t^R > p_t^O$ in later periods.

4.3 Surplus Under Price-Change Specification

In this section, we compare total and consumer surplus under overcharge-based penalties and revenue-based penalties. There are two cases to consider: $\delta \leq \delta_O$ and $\delta > \delta_O$. First, consider the case of $\delta \leq \delta_O$. Theorem 6 implies that $\delta_R > \delta_O$ when α_1 is sufficiently high. Therefore, no cartel forms under either penalty type and both penalties generate the same level of total and consumer surplus when $\delta \leq \delta_O$ and the probability of detection is sufficiently sensitive to the change in cartel price.

[29] Harrington (2004) makes a similar observation with regard to the limiting (or steady-state) cartel price (see Theorem 2 in Harrington (2004)). Recall that, as Katsoulacos et al. (2015) showed, the cartel price under revenue-based penalties is greater than the cartel price under overcharge-based penalties when the probability of detection is constant.

[30] See https://www.douglascturner.com/endogenous-detection-online-appendix/.

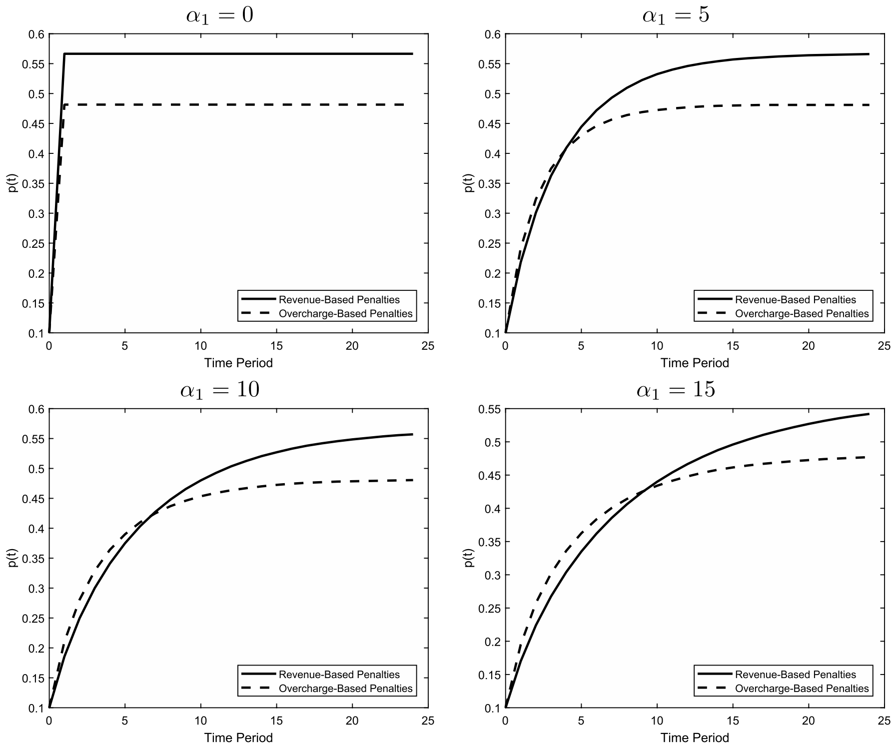

Fig. 2 Cartel prices in the baseline setting

If $\delta > \delta_O$, a cartel forms under overcharge-based penalties. However, a cartel does not form under revenue-based penalties when α_1 is sufficiently high, by Theorem 6. Therefore, both total and consumer surplus are greater under revenue-based penalties than under overcharge-based penalties when $\delta > \delta_O$ and the probability of detection is sufficiently sensitive to the change in cartel price.[31]

Theorem 7 *When $\delta > \delta_O$, there exists an $\tilde{\alpha}_1^C$ such that $CS_R > CS_O$ and $TS_R > TS_O$ if $\alpha_1 > \tilde{\alpha}_1^C$. When $\delta \le \delta_O$, there exists an $\tilde{\alpha}_1^C$ such that $CS_R = CS_O$ and $TS_R = TS_O$ if $\alpha_1 > \tilde{\alpha}_1^C$.*

As under the price-level specification, the optimal penalty type depends on α_1. As Katsoulacos et al. (2015) show, overcharge-based penalties induce a greater level of surplus when the probability of detection does not depend on the cartel price ($\alpha_1 = 0$). Conversely, surplus is greater under revenue-based penalties when the probability of detection is sufficiently sensitive to price increases.

[31] The proof of Theorem 7 does not depend on the numerical results of Sect. 4.2; see "Appendix 1".

5 Conclusion

We have examined revenue- and aggregate overcharge-based cartel penalties when the probability of detection depends on the cartel price. We have compared the two penalty structures on the basis of both consumer surplus and total surplus. First, we consider a model in which the probability of detection depends on the cartel overcharge. Next, we analyze a model in which the probability of detection depends on the rate at which the cartel price increases. Under both models, we obtain a similar result: When the probability of detection is sufficiently sensitive to the cartel price, both consumer surplus and total surplus are higher under revenue-based penalties.

In summary, which penalty structure secures a higher level of surplus depends crucially on the detection process. In industries where a large price increase or a high cartel overcharge is considered anomalous, indicative of collusion, and likely to lead to detection, revenue-based cartel penalties generate a higher level of surplus. For example, in a stable industry which does not experience significant price variation when competitive, a large price increase is likely to attract the attention of an antitrust authority and lead to detection. In this case, revenue-based penalties are better.[32]

In industries where a large price increase or high overcharge is not likely to lead to detection, overcharge-based penalties generate a higher level of surplus. For instance, in a turbulent industry which experiences regular price variation even when competitive, a large price increase is unlikely to raise suspicions of collusion. In this case, aggregate overcharge-based penalties are better.

Appendix 1: Proofs

Price-Level Specification

Let $W_R(p)$ denote the present discounted value of profit less expected penalties under revenue-based penalties when the cartel price is p. Let $W_O(p)$ denote the present discounted value of profit less expected penalties under overcharge-based penalties when the cartel price is p. $p_R = \text{argmax}_{p\in[c,1]} W_R(p)$ subject to $W_R(p) \geq \pi^D(p)$ is the cartel price under revenue-based penalties, if a cartel forms. $p_O = \text{argmax}_{p\in[c,1]} W_O(p)$ subject to $W_O(p) \geq \pi^D(p)$ is the cartel price under overcharge-based penalties, if a cartel forms.

Lemma 1 provides bounds on the cartel price, which necessarily hold if a cartel forms.

Lemma 1 *$p_O \in \left(c, c+\sqrt{\frac{1-\alpha_0}{\alpha_1}}\right)$ and $p_R \in \left(c, c+\sqrt{\frac{1-\alpha_0}{\alpha_1}}\right)$ if a cartel forms.*

[32] Revenue-based penalties have an additional advantage. Revenue-based penalties are less costly to calculate than alternative penalty types, because they require no information about a firm's marginal cost or but-for price. Information on firm revenue is often readily available to investigators.

Proof If $p_R = c$, the payoff from collusion is

$$W_R(c) = \frac{(1-c)(c-c) - \alpha_0\gamma_R c(1-c)}{1-\delta} = -\frac{\alpha_0\gamma_R c(1-c)}{1-\delta} < 0,$$

which implies that collusion is not profitable.

If $p_R \geq c + \sqrt{\frac{1-\alpha_0}{\alpha_1}}$, then the probability of detection is 1. Therefore, the payoff from collusion, when the cartel price is p_R, is

$$\begin{aligned} W_R(p_R) =& \frac{(1-p_R)(p_R-c) - \gamma_R p_R(1-p_R)}{1-\delta} < \frac{(1-p_R)(p_R-c) - p_R(1-p_R)}{1-\delta} \\ =& \frac{-(1-p_R)c}{1-\delta} \leq 0, \end{aligned}$$

which implies that collusion is not profitable (the first inequality follows from $\gamma_R > 1$).

If $p_O = c$, the payoff from collusion is

$$W_O(c) = \frac{(1-c)(c-c) - \alpha_0\gamma_O Q_N(c-c)}{1-\delta} = 0,$$

which implies that collusion is not profitable.

If $p_O \geq c + \sqrt{\frac{1-\alpha_0}{\alpha_1}}$, then the probability of detection is 1. Therefore, the payoff from collusion, when the cartel price is p_O, is

$$\begin{aligned} W_O(p_O) =& \frac{(1-p_O)(p_O-c) - \gamma_O Q_N(p_O-c)}{1-\delta} < \frac{(1-p_O)(p_O-c) - (1-c)(p_O-c)}{1-\delta} \\ =& \frac{-(p_O-c)^2}{1-\delta} < 0, \end{aligned}$$

which implies that collusion is not profitable (the first inequality follows from $\gamma_O > 1$). □

Lemma 2 *$W_R(p)$ and $W_O(p)$ are strictly concave on $\left[c, c+\sqrt{\frac{1-\alpha_0}{\alpha_1}}\right]$ if $\alpha_1 > \max\left\{\left(\frac{5}{H(c)}\right)^2(1-\alpha_0), \frac{4(1-\alpha_0)}{(1-c)^2}\right\}$ where $H(c) = \min\left\{c(1-c), \left(\frac{1+c}{2}\right)\left(1-\frac{1+c}{2}\right)\right\}$.*

Proof First, we consider revenue-based penalties. Note that $W_R(p) = \frac{1}{1-\delta}\left((1-p)(p-c) - \left(\alpha_0 + \alpha_1(p-c)^2\right)\gamma_R p(1-p)\right)$ when $p \in \left[c, c+\sqrt{\frac{1-\alpha_0}{\alpha_1}}\right]$. The second derivative of $W_R(p)$ when $p \in \left[c, c+\sqrt{\frac{1-\alpha_0}{\alpha_1}}\right]$ is

$$\begin{aligned}
\frac{\partial^2 W_R(p)}{\partial p^2} &= \frac{\partial^2}{\partial p^2}\frac{1}{1-\delta}\big((1-p)(p-c) - \big(\alpha_0 + \alpha_1(p-c)^2\big)\gamma_R p(1-p)\big) \\
&= \frac{\partial}{\partial p}\frac{1}{1-\delta}\big(1 - 2p + c - \alpha_1\gamma_R\big((p-c)^2(1-2p) + 2p(1-p)(p-c)\big) \\
&\quad -\alpha_0\gamma_R(1-2p)\big) \\
&= \frac{\partial}{\partial p}\frac{1}{1-\delta}\big(1 - 2p + c - \alpha_1\gamma_R \\
&\quad \big((p-c)^2 - 2p(p-c)^2 + 2p(p-c) - 2p^2(p-c)\big) - \alpha_0\gamma_R(1-2p)\big) \\
&= \frac{1}{1-\delta}\big(-2 - \alpha_1\gamma_R\big(2(p-c) - 2(p-c)^2 \\
&\quad -4p(p-c) + 2(p-c) + 2p - 4p(p-c) - 2p^2\big) + 2\alpha_0\gamma_R\big) \\
&= \frac{1}{1-\delta}\big(-2\big(1-\alpha_0\gamma_R\big) - \alpha_1\gamma_R\big(4(p-c) - 2(p-c)^2 - 8p(p-c) + 2p - 2p^2\big)\big)
\end{aligned}$$

which is less than 0 if $4(p-c) - 2(p-c)^2 - 8p(p-c) + 2p - 2p^2 \geq 0$. Note that $p(1-p) \geq H(c)$ because $c \leq p \leq c + \sqrt{\frac{1-\alpha_0}{\alpha_1}} < \frac{1+c}{2}$ where the last inequality follows from $\alpha_1 > \frac{4(1-\alpha_0)}{(1-c)^2}$.[33] Therefore,

$$\begin{aligned}
&4(p-c) - 2(p-c)^2 - 8p(p-c) + 2p - 2p^2 \\
\geq& -2(p-c)^2 - 8p(p-c) + 2p(1-p) \\
\geq& -2(p-c)^2 - 8p(p-c) + 2H(c) && \text{by } p(1-p) \geq H(c) \\
\geq& -2(p-c) - 8(p-c) + 2H(c) \\
\geq& -10(p-c) + 2H(c) \\
\geq& -10\sqrt{\frac{1-\alpha_0}{\alpha_1}} + 2H(c) \\
\geq& -10\sqrt{\frac{1-\alpha_0}{\left(\frac{5}{H(c)}\right)^2(1-\alpha_0)}} + 2H(c) && \text{by } \alpha_1 > \left(\frac{5}{H(c)}\right)^2(1-\alpha_0) \\
\geq& -10\left(\frac{H(c)}{5}\right) + 2H(c) \\
\geq& -2H(c) + 2H(c) = 0.
\end{aligned}$$

Therefore, $W_R(p)$ is strictly concave on $\left[c, c + \sqrt{\frac{1-\alpha_0}{\alpha_1}}\right]$.

Next, we show $W_O(p)$ is strictly concave. Note that $W_O(p) = \frac{1}{1-\delta}\big((1-p)(p-c) - \big(\alpha_0 + \alpha_1(p-c)^2\big)\gamma_O Q_N(p-c)\big)$ when $p \in \left[c, c + \sqrt{\frac{1-\alpha_0}{\alpha_1}}\right]$. The second derivative is

[33] $c \leq p < \frac{1+c}{2}$ implies $p(1-p) \geq H(c)$ because $p(1-p) \geq c(1-c)$ if $p < \frac{1}{2}$ and $p(1-p) \geq \frac{1+c}{2}(1 - \frac{1+c}{2})$ if $p \geq \frac{1}{2}$.

$$\begin{aligned}
\frac{\partial^2 W_O(p)}{\partial p^2} &= \frac{\partial^2}{\partial p^2}\frac{1}{1-\delta}\big((1-p)(p-c)-\alpha_1(p-c)^2\gamma_O Q_N(p-c)-\alpha_0\gamma_O Q_N(p-c)\big)\\
&= \frac{\partial}{\partial p}\frac{1}{1-\delta}\big(1-2p+c-3\alpha_1(p-c)^2\gamma_O Q_N-\alpha_0\gamma_O Q_N\big)\\
&= \frac{1}{1-\delta}\big(-2-6\alpha_1(p-c)Q_N\gamma_O\big) < 0.
\end{aligned}$$

Therefore, $W_O(p)$ is strictly concave when $p \in \left[c, c+\sqrt{\frac{1-\alpha_0}{\alpha_1}}\right]$. □

***Proof of Theorem* 1** We show that a cartel forms when $\delta > \delta_O$ and does not form when $\delta \le \delta_O$. A cartel forms if collusion is sustainable and profitable. For collusion to be sustainable, collusion must be incentive-compatible: The constraint in Eq. (1) must hold. Note that the cartel price p satisfies both $c < p < c + \sqrt{\frac{1-\alpha_0}{\alpha_1}}$, by Lemma 1, and $p \le p^m$.[34] Collusion is sustainable at a price $p = c + \epsilon$ where $p \le p^m$ and $p \in \left(c, c+\sqrt{\frac{1-\alpha_0}{\alpha_1}}\right)$ if

$$\begin{aligned}
N\pi(p) &\le \frac{1}{1-\delta}\pi(p) - \frac{\alpha_0 x_O(p)}{1-\delta} - \frac{\alpha_1(p-c)^2 x_O(p)}{1-\delta}\\
1-\delta &\le \frac{1}{N} - \frac{\alpha_0 x_O(p)}{N\pi(p)} - \frac{\alpha_1(p-c)^2 x_O(p)}{N\pi(p)}\\
\delta &\ge \frac{N-1}{N} + \frac{\alpha_0 x_O(p)}{N\pi(p)} + \frac{\alpha_1(p-c)^2 x_O(p)}{N\pi(p)}\\
\delta &\ge \frac{N-1}{N} + \frac{\alpha_0\gamma_O Q_N\epsilon}{N\epsilon D(c+\epsilon)} + \frac{\alpha_1\epsilon^2\gamma_O Q_N\epsilon}{N\epsilon D(c+\epsilon)}\\
\delta &\ge \frac{N-1}{N} + \frac{\alpha_0\gamma_O Q_N}{ND(c+\epsilon)} + \frac{\alpha_1\epsilon^2\gamma_O Q_N}{ND(c+\epsilon)}\\
\delta &\ge \frac{N-1}{N} + \frac{\alpha_0\gamma_O(1-c-\epsilon+\epsilon)}{ND(c+\epsilon)} + \frac{\alpha_1\epsilon^2\gamma_O Q_N}{ND(c+\epsilon)}\\
\delta &\ge \frac{N-1}{N} + \frac{\alpha_0\gamma_O D(c+\epsilon)}{ND(c+\epsilon)} + \epsilon\frac{\alpha_0\gamma_O}{ND(c+\epsilon)} + \frac{\alpha_1\epsilon^2\gamma_O Q_N}{ND(c+\epsilon)}\\
\delta &\ge \frac{N-1}{N} + \frac{\alpha_0\gamma_O}{N} + \epsilon\frac{\alpha_0\gamma_O}{ND(c+\epsilon)} + \epsilon^2\frac{\alpha_1\gamma_O Q_N}{ND(c+\epsilon)}.
\end{aligned} \tag{4}$$

If $\delta > \delta_O = \frac{N-1}{N} + \frac{\alpha_0\gamma_O}{N}$, then there exists a sufficiently small $\epsilon > 0$ such that inequality (4) holds and collusion is sustainable. Collusion is also profitable at price $p = c + \epsilon$ because

[34] The optimal price under aggregate overcharge-based penalties does not exceed the monopoly price. To see this, suppose that the optimal cartel price is $p > p^m$. The cartel could increase profit and reduce the expected penalty by reducing the cartel price to p^m. In addition, the payoff from defection is unchanged. Thus, $p > p^m$ is not optimal.

$$0 < N\pi(p) \leq \frac{1}{1-\delta}\pi(p) - \frac{\alpha_0 x_O(p)}{1-\delta} - \frac{\alpha_1 (p-c)^2 x_O(p)}{1-\delta},$$

where the first inequality follows from $c < p \leq p^m$ and the second inequality follows from the sustainability of collusion at price $p = c + \epsilon$. Therefore, collusion is profitable and sustainable if $\delta > \delta_O$. If $\delta \leq \frac{N-1}{N} + \frac{\alpha_0 \gamma_O}{N}$, collusion is unsustainable for all $\epsilon \in \left(0, \sqrt{\frac{1-\alpha_0}{\alpha_1}}\right)$ (see inequality (4)) and unprofitable for $\epsilon = 0$ and $\epsilon \geq \sqrt{\frac{1-\alpha_0}{\alpha_1}}$ (see Lemma 1). Thus, a cartel does not form. □

***Proof of Theorem* 2** First, we show $\delta_R \to 1$ as $\alpha_1 \to \infty$. A cartel forms if collusion is sustainable and profitable. By Lemma 1, $p \in \left(c, c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right)$. We show collusion is unprofitable for all $\delta < 1$ if $\alpha_1 > \frac{(1-\alpha_0)\left(1-\alpha_0\gamma_R\right)^2}{\left(\alpha_0\gamma_R c\right)^2}$. Therefore, a cartel does not form and $\delta_R = 1$ if $\alpha_1 > \frac{(1-\alpha_0)\left(1-\alpha_0\gamma_R\right)^2}{\left(\alpha_0\gamma_R c\right)^2}$. For collusion to be profitable, $\pi(p) - \alpha_0 x_R(p) > 0$ must hold for some price $p \in \left(c, c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right)$.[35] Note that

$$\begin{aligned}
\pi(p) - \alpha_0 x_R(p) &= D(p)(p-c) - \alpha_0\gamma_R D(p)p \\
&= D(p)\left[p\left[1-\alpha_0\gamma_R\right] - c\right] \\
&\leq D(p)\left[\left[c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right]\left[1-\alpha_0\gamma_R\right] - c\right] \\
&= D(p)\left[c - \alpha_0\gamma_R c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\left[1-\alpha_0\gamma_R\right] - c\right] \\
&= D(p)\left[-\alpha_0\gamma_R c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\left[1-\alpha_0\gamma_R\right]\right] \\
&\leq D(p)\left[-\alpha_0\gamma_R c + \frac{\alpha_0\gamma_R c}{1-\alpha_0\gamma_R}\left[1-\alpha_0\gamma_R\right]\right] \\
&= 0
\end{aligned}$$

for all $p \in \left(c, c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right)$ where the last inequality follows from $\alpha_1 > \frac{(1-\alpha_0)\left(1-\alpha_0\gamma_R\right)^2}{\left(\alpha_0\gamma_R c\right)^2}$. Thus, collusion is unprofitable and a cartel does not form – $\delta_R = 1$—if $\alpha_1 > \frac{(1-\alpha_0)\left(1-\alpha_0\gamma_R\right)^2}{\left(\alpha_0\gamma_R c\right)^2}$, which implies $\delta_R \to 1$ as $\alpha_1 \to \infty$.

Next, we show $\delta_R \to 1$ monotonically as $\alpha_1 \to \infty$. Let $\delta_R(\alpha_1)$ denote the critical discount factor when the sensitivity of the probability of detection to the overcharge is α_1. Suppose that $\delta_R(\alpha_1') > \delta_R(\alpha_1'')$ for some $\alpha_1' < \alpha_1''$. Let δ be such that

[35] If $\pi(p) - \alpha_0 x_R(p) \leq 0$ for all $p \in \left(c, c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right)$, then $W_R(p) = \frac{1}{1-\delta}\left[\pi(p) - \phi x_R(p)\right] \leq 0$ for all $p \in \left(c, c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right)$ (by $\phi \geq \alpha_0$) and collusion is unprofitable.

$\delta_R(\alpha_1'') < \delta < \delta_R(\alpha_1')$. A cartel forms when the discount factor is δ if $\alpha_1 = \alpha_1''$ and does not form if $\alpha_1 = \alpha_1'$. Let p'' denote the cartel price when $\alpha_1 = \alpha_1''$ and the discount factor is δ. Let $W_R(p;\alpha_1,\delta)$ denote the payoff from collusion when the cartel price is p, the sensitivity of the probability of detection to the overcharge is α_1 and the discount factor is δ. Note that $\pi^D(p'') \leq W_R(p'';\alpha_1'',\delta)$ because collusion is sustainable when the cartel price is p''. Also, note that $0 < W_R(p'';\alpha_1'',\delta)$ because collusion is profitable when the cartel price is p''. However, $\pi^D(p'') \leq W_R(p'';\alpha_1'',\delta) \leq W_R(p'';\alpha_1',\delta)$ and $0 < W_R(p'';\alpha_1'',\delta) \leq W_R(p'';\alpha_1',\delta)$ because $\alpha_1' < \alpha_1''$.[36] Thus, collusion is sustainable and profitable at a price of p'' when $\alpha_1 = \alpha_1'$. This implies that a cartel forms when $\delta < \delta_R(\alpha_1')$, which is a contradiction. Therefore, we conclude that $\delta_R \to 1$ monotonically as $\alpha_1 \to \infty$. □

Proof of Theorem 3 Let

$$\bar{\alpha}_1^L = \max\left\{\frac{9\left(1-\alpha_0\right)}{(1-c)^2}, \frac{1}{3\gamma_O Q_N^2\left(1-\alpha_0\gamma_O\right)}\left(\left(\frac{3\gamma_O Q_N}{2\gamma_R H(c)}\left(1-c+\gamma_R\right)+1\right)^2-1\right), \left(\frac{5}{H(c)}\right)^2(1-\alpha_0), \left(\frac{\gamma_O Q_N}{\gamma_R H(c)}\right)^2\left(1-\alpha_0\right)\right\} \tag{5}$$

where $H(c) = \min\left\{c(1-c), \left(\frac{1+c}{2}\right)\left(1-\frac{1+c}{2}\right)\right\}$. Suppose that $\alpha_1 > \bar{\alpha}_1^L$.

The proof involves three steps: In step 1, we establish that the cartel price under overcharge-based penalties exceeds the cartel price under revenue-based penalties when $\alpha_1 > \bar{\alpha}_1^L$ and incentive compatibility constraints do not bind. This is proven by establishing that the derivative of the payoff from collusion under revenue-based penalties is negative at the unconstrained cartel price under overcharge-based penalties. In step 2, we establish that, when $\alpha_1 > \bar{\alpha}_1^L$, revenue-based penalties exceed overcharge-based cartel penalties. In step 3, we show, using the results from step 1 and step 2, that $p_R < p_O$ generally.

Step 1: Let p_i^{NB} denote the cartel price when the incentive compatibility constraint does not bind under penalty type $i \in \{R, O\}$. First, we show $p_R^{NB} < p_O^{NB}$. Note that $p_O^{NB}, p_R^{NB} \in \left(c, c+\frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right)$ by Lemma 1. By the concavity of $W_O(p)$ (Lemma 2), the first order condition which determines p_O^{NB} is

$$\frac{\partial W_O(p)}{\partial p}(1-\delta) = 1 - 2p + c - 3\alpha_1\gamma_O Q_N (p-c)^2 - \alpha_0\gamma_O Q_N = 0.$$

Solving for p yields

[36] $W_R(p;\alpha_1'',\delta) \leq W_R(p;\alpha'_1,\delta)$ if $\alpha_1' < \alpha_1''$ because

$$W_R(p;\alpha_1'',\delta) = \frac{(1-p)(p-c) - \phi(p;\alpha_1'')x_R(p)}{1-\delta} \leq \frac{(1-p)(p-c) - \phi(p;\alpha_1')x_R(p)}{1-\delta} = W_R(p;\alpha'_1,\delta)$$

which follows from $\phi(p;\alpha_1'') \geq \phi(p;\alpha_1')$, where $\phi(p;\alpha_1) = \min\{\alpha_0 + \alpha_1(p-c)^2, 1\}$.

$$p_O^{NB} = \frac{\sqrt{1 + 3\alpha_1 \gamma_O Q_N^2 \big(1 - \alpha_0 \gamma_O\big)} - 1}{3\alpha_1 \gamma_O Q_N} + c. \tag{6}$$

By the strict concavity of $W_R(p)$ (Lemma 2), the first-order condition that determines p_R^{NB} is

$$\begin{aligned}\frac{\partial W_R(p)}{\partial p}(1-\delta) = 1 - 2p + c - \big(\alpha_0 + \alpha_1 (p-c)^2\big)\gamma_R (1-2p) \\ - 2\alpha_1 (p-c)\gamma_R p(1-p) = 0.\end{aligned} \tag{7}$$

Note that

$$\begin{aligned}\frac{\partial W_R(p)}{\partial p}(1-\delta) &\le 1 - 2p + c + \gamma_R - 2\alpha_1 (p-c)\gamma_R p(1-p) \\ &\le 1 - c + \gamma_R - 2\alpha_1 (p-c)\gamma_R p(1-p),\end{aligned}$$

where the first inequality follows from $-\big(\alpha_0 + \alpha_1 (p-c)^2\big)\gamma_R (1-2p) < \gamma_R$ and the second inequality follows from $p > c$. By the strict concavity of $W_R(p)$, $p_R^{NB} < p_O^{NB}$ if $\frac{\partial W_R(p_O^{NB})}{\partial p} < 0$.

$$\begin{aligned}\frac{\partial W_R(p_O^{NB})}{\partial p}(1-\delta) &\le 1 - c + \gamma_R - 2\alpha_1 (p_O^{NB} - c)\gamma_R p_O^{NB}(1 - p_O^{NB}) \\ &= 1 - c + \gamma_R - 2\alpha_1 \frac{\sqrt{1 + 3\alpha_1 \gamma_O Q_N^2 \big(1 - \alpha_0 \gamma_O\big)} - 1}{3\alpha_1 \gamma_O Q_N} \gamma_R p_O^{NB}(1 - p_O^{NB}) \\ &= 1 - c + \gamma_R - 2 \frac{\sqrt{1 + 3\alpha_1 \gamma_O Q_N^2 \big(1 - \alpha_0 \gamma_O\big)} - 1}{3\gamma_O Q_N} \gamma_R p_O^{NB}(1 - p_O^{NB}) \\ &\le 1 - c + \gamma_R - 2 \frac{\sqrt{1 + 3\alpha_1 \gamma_O Q_N^2 \big(1 - \alpha_0 \gamma_O\big)} - 1}{3\gamma_O Q_N} \gamma_R H(c),\end{aligned}$$

where the last inequality follows from $p_O^{NB}(1 - p_O^{NB}) \ge H(c)$, which follows from $c < p_O^{NB} < \frac{1+c}{2}$.[37] Next, note that

[37] $\alpha_1 > \bar{\alpha}_1^L \ge \frac{9(1-\alpha_0)}{(1-c)^2} > \frac{4(1-\alpha_0)}{(1-c)^2}$ implies that $p < c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} \le c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\frac{4(1-\alpha_0)}{(1-c)^2}}} \le c + (1-c)\frac{\sqrt{1-\alpha_0}}{2\sqrt{1-\alpha_0}} = \frac{1+c}{2}$.

$$
\begin{aligned}
&1 - c + \gamma_R - 2\frac{\sqrt{1 + 3\alpha_1\gamma_O Q_N^2\left(1 - \alpha_0\gamma_O\right)} - 1}{3\gamma_O Q_N}\gamma_R H(c) < 0 \\
&\iff 1 - c + \gamma_R < 2\frac{\sqrt{1 + 3\alpha_1\gamma_O Q_N^2\left(1 - \alpha_0\gamma_O\right)} - 1}{3\gamma_O Q_N}\gamma_R H(c) \\
&\iff \frac{3\gamma_O Q_N}{2\gamma_R H(c)}\left(1 - c + \gamma_R\right) < \sqrt{1 + 3\alpha_1\gamma_O Q_N^2\left(1 - \alpha_0\gamma_O\right)} - 1 \\
&\iff \frac{3\gamma_O Q_N}{2\gamma_R H(c)}\left(1 - c + \gamma_R\right) + 1 < \sqrt{1 + 3\alpha_1\gamma_O Q_N^2\left(1 - \alpha_0\gamma_O\right)} \\
&\iff \left(\frac{3\gamma_O Q_N}{2\gamma_R H(c)}\left(1 - c + \gamma_R\right) + 1\right)^2 < 1 + 3\alpha_1\gamma_O Q_N^2\left(1 - \alpha_0\gamma_O\right) \\
&\iff \frac{1}{3\gamma_O Q_N^2\left(1 - \alpha_0\gamma_O\right)}\left(\left(\frac{3\gamma_O Q_N}{2\gamma_R H(c)}\left(1 - c + \gamma_R\right) + 1\right)^2 - 1\right) < \alpha_1
\end{aligned}
$$

which holds by (5). Therefore, $\frac{\partial W_R(p_O^{NB})}{\partial p} < 0$ and $p_R^{NB} < p_O^{NB}$.

Step 2: Next, we show that revenue-based penalties are greater than overcharge-based penalties for all $p \in \left[c, c + \sqrt{\frac{1-\alpha_0}{\alpha_1}}\right)$. Note that

$$
\begin{aligned}
x_R(p) - x_O(p) &= \gamma_R p(1 - p) - \gamma_O Q_N(p - c) \\
&\geq \gamma_R H(c) - \gamma_O Q_N(p - c) \\
&\geq \gamma_R H(c) - \gamma_O Q_N\sqrt{\frac{1 - \alpha_0}{\alpha_1}}.
\end{aligned}
$$

$\gamma_R H(c) - \gamma_O Q_N\sqrt{\frac{1-\alpha_0}{\alpha_1}} > 0$ holds if

$$
\begin{aligned}
\gamma_R H(c) &> \gamma_O Q_N\sqrt{\frac{1 - \alpha_0}{\alpha_1}} \\
\sqrt{\alpha_1} &> \frac{\gamma_O Q_N}{\gamma_R H(c)}\sqrt{1 - \alpha_0} \\
\alpha_1 &> \left(\frac{\gamma_O Q_N}{\gamma_R H(c)}\right)^2\left(1 - \alpha_0\right),
\end{aligned}
$$

which holds by (5). Thus, $x_R(p) > x_O(p)$ and $W_O(p) > W_R(p)$ for all $p \in \left[c, c + \sqrt{\frac{1-\alpha_0}{\alpha_1}}\right)$.

Step 3: Suppose $p_O < p_R$. We show that $p_R + \epsilon$, for sufficiently small $\epsilon > 0$, is incentive-compatible under aggregate overcharge-based penalties and achieves a greater payoff than p_O. Therefore, p_O is not optimal. Note that

$$
W_O(p_R + \epsilon) > W_R(p_R) \geq \pi^D(p_R),
$$

where the first inequality follows from $W_O(p) > W_R(p)$ and the continuity of $W_O(p)$. The second inequality follows from the incentive-compatibility of p_R.[38] Thus, $p_R + \epsilon$ is incentive compatible under overcharge-based penalties.

Next, note that $p_O < p_R \leq p_R^{NB} < p_O^{NB}$ which implies that $p_O < p_R + \epsilon < p_O^{NB}$ for sufficiently small $\epsilon > 0$.[39] Therefore, $W_O(p_O) < W_O(p_R + \epsilon)$ by the strict concavity of $W_O(p)$. We have shown $p_R + \epsilon$ is incentive-compatible under overcharge-based penalties and results in a greater collusive payoff. Therefore, p_O is not the optimal cartel price, which is a contradiction. □

***Proof of Theorem* 4** First, consider the case of $\delta > \delta_O$. The proof follows from the observation that, by Theorems 1 and 2, there exists an $\tilde{\alpha}_1^L$ such that $\delta_R > \delta > \delta_O$ if $\alpha_1 > \tilde{\alpha}_1^L$. Next, consider the case of $\delta \leq \delta_O$. When $\delta \leq \delta_O$, a cartel does not form under overcharge-based penalties (see the proof of Theorem 1). By Theorem 2, there exists an $\tilde{\alpha}_1^L$ such that $\delta_R > \delta$ if $\alpha_1 > \tilde{\alpha}_1^L$. Thus, a cartel does not form under either penalty type when $\alpha_1 > \tilde{\alpha}_1^L$ which implies $CS_R = CS_O$ and $TS_R = TS_O$. □

Changes Specification

***Proof of Theorem* 5** The proof involves two steps: First, we show a cartel forms when $\delta > \delta_O$. Second, we show that a cartel does not form when $\delta \leq \delta_O$.

Suppose that $\delta > \delta_O$. Let $W_O^j(p)$ denote the payoff from collusion under specification $j \in \{L, C\}$ when the cartel price is p in all periods.[40] A cartel forms under specification L when $\delta > \delta_O$ (see Theorem 1). Let p_O^L denote the cartel price under specification L. $p_O^L > c$, and therefore $\pi^D(p_O^L) > 0$, by Lemma 1. In addition, $W_O^L(p_O^L) \geq \pi^D(p_O^L)$ because collusion must be sustainable for a cartel to form. Lastly, note that $W_O^C(p) \geq W_O^L(p)$ for all p. Thus,

$$W_O^C(p_O^L) \geq W_O^L(p_O^L) \geq \pi^D(p_O^L) > 0,$$

which implies that collusion is sustainable and profitable under specification C when $\delta > \delta_O$ at a constant price of p_O^L. Thus, a cartel forms under specification C when $\delta > \delta_O$.

Next, suppose $\delta \leq \delta_O$ and a price path of $\{p_t\}_{t=1}^{\infty}$ is sustainable and profitable under specification C. Let $W_O^j\left(p_{t-1}, \{p_t\}_{t=1}^{\infty}; \alpha_1\right)$ denote the present discounted value of collusion when: the prior period's price is p_{t-1}; the cartel price path is $\{p_t\}_{t=1}^{\infty}$; the sensitivity of the probability of detection to price is α_1 and the ϕ specification is $j \in \{L, C\}$.

[38] Note that collusion is also profitable at price $p_R + \epsilon$ under overcharge-based penalties because $W_O(p_R + \epsilon) > W_R(p_R) \geq \pi^D(p_R) > 0$ where the last inequality follows from the fact that $p_R > c$ by Lemma 1.

[39] Suppose $p_i > p_i^{NB}$ where $i \in \{O, R\}$. p_i is not optimal because $W_i(p_i^{NB}) > W_i(p_i) \geq \pi^D(p_i) \geq \pi^D(p_i^{NB})$ where the first inequality follows from the definition of p_i^{NB} and the strict concavity of $W_i(p)$, the second inequality follows from the incentive compatibility of p_i and the last inequality follows from the observation that $\pi^D(p_i)$ is nondecreasing. Thus, p_i^{NB} is incentive compatible and yields a greater collusive payoff.

[40] Recall that L denotes the price-level specification and C denotes the price-change specification.

First, note that $W_O^C\left(p_{t-1},\{p_t\}_{t=1}^{\infty};\alpha_1\right) \leq W_O^C\left(p_{t-1},\{p_t\}_{t=1}^{\infty};0\right) = W_O^L\left(p_{t-1},\{p_t\}_{t=1}^{\infty};0\right)$ for all t.[41] As $\{p_t\}_{t=1}^{\infty}$ is sustainable under specification C,

$$\pi^D(p_t) \leq W_O^C\left(p_{t-1},\{p_t\}_{t=1}^{\infty};\alpha_1\right) \leq W_O^C\left(p_{t-1},\{p_t\}_{t=1}^{\infty};0\right) = W_O^L\left(p_{t-1},\{p_t\}_{t=1}^{\infty};0\right)$$

for all t. This implies that collusion is sustainable when $\alpha_1 = 0$ and $\delta \leq \delta_O$ under specification L. Because the price path $\{p_t\}_{t=1}^{\infty}$ is profitable under specification C,

$$0 < W_O^C\left(c,\{p_t\}_{t=1}^{\infty};\alpha_1\right) \leq W_O^C\left(c,\{p_t\}_{t=1}^{\infty};0\right) = W_O^L\left(c,\{p_t\}_{t=1}^{\infty};0\right)$$

which implies collusion is profitable when $\alpha_1 = 0$ and $\delta \leq \delta_O$ under specification L. We have shown the price path $\{p_t\}_{t=1}^{\infty}$ is profitable and sustainable when $\alpha_1 = 0$ under specification L. Therefore, a cartel forms when $\alpha_1 = 0$ and $\delta \leq \delta_O$ under specification L, which is a contradiction. □

Lemma 3 $p_t < p_{t-1} + \sqrt{\frac{1-\alpha_0}{\alpha_1}}$ *for all* $t \geq 1$ *under both overcharge and revenue-based penalties.*

Proof Let $\{p_t\}_{t=1}^{\infty}$ be an optimal price path where $p_{t'} \geq p_{t'-1} + \sqrt{\frac{1-\alpha_0}{\alpha_1}}$ for some t' (and let t' denote the first such t). Note that $\phi(p_{t'-1}, p_{t'}) = 1$. The optimality of $p_{t'}, p_{t'+1}, p_{t'+2} \dots$ implies that it yields a payoff that is greater than or equal to the payoff from the alternative price sequence of $p_1, p_2, p_3 \dots$ from t' onwards.[42] Thus,

$$\begin{aligned} V(p_{t'-1}) &= \pi(p_{t'}) - x(p_{t'}) + \delta V(c) \\ &= \pi(p_{t'}) - x(p_{t'}) + \delta\left[\pi(p_1) - \phi(c,p_1)x(p_1)\right. \\ &\quad \left. + \delta\left[1-\phi(c,p_1)\right]V(p_1) + \delta\phi(c,p_1)V(c)\right] \\ &\geq \pi(p_1) - \phi(p_{t'-1},p_1)x(p_1) + \delta\left[1-\phi(p_{t'-1},p_1)\right]V(p_1) \\ &\quad + \delta\phi(p_{t'-1},p_1)V(c). \end{aligned} \tag{8}$$

Note that $\pi(p_{t'}) - x(p_{t'}) \leq 0$ because $\gamma_O > 1$ and $\gamma_R > 1$.[43] Thus, inequality (8) implies

$$\delta V(c) \geq \pi(p_1) - \phi(p_{t'-1},p_1)x(p_1) + \delta\left[1-\phi(p_{t'-1},p_1)\right]V(p_1) + \delta\phi(p_{t'-1},p_1)V(c),$$

41 Specification L and specification C are equivalent when $\alpha_1 = 0$.

42 Note that $p_1, p_2, p_3 \dots$ is sustainable when the price in the prior period was p_{t-1} because

$$\begin{aligned} &\pi(p_1) - \phi(p_{t'-1},p_1)x(p_1) + \delta\left[1-\phi(p_{t'-1},p_1)\right]V(p_1) + \delta\phi(p_{t'-1},p_1)V(c) \\ \geq &\pi(p_1) - \phi(c,p_1)x(p_1) + \delta\left[1-\phi(c,p_1)\right]V(p_1) + \delta\phi(c,p_1)V(c) \\ \geq &\pi^D(p_1) \end{aligned}$$

where the first inequality follows from $\phi(p_{t'-1},p_1) \leq \phi(c,p_1)$. The second inequality follows from the sustainability of p_1 when the price in the prior period is c.

43 $\pi(p) - \gamma_R p(1-p) = (1-p)(p(1-\gamma_R) - c) = (1-p)(-p(\gamma_R - 1) - c) = -(1-p)(p(\gamma_R - 1) + c) \leq 0$ and $\pi(p) - \gamma_O Q_N(p-c) = (p-c)(1-p-\gamma_O Q_N) \leq (p-c)(1-c-\gamma_O Q_N) = -(p-c)Q_N(\gamma_O - 1) \leq 0$.

or

$$\begin{aligned}&\pi(p_1)-\phi(p_{t'-1},p_1)x(p_1)+\delta\left[1-\phi(p_{t'-1},p_1)\right]V(p_1)+\delta\phi(p_{t'-1},p_1)V(c)\\ \leq\ &\delta\left[\pi(p_1)-\phi(c,p_1)x(p_1)+\delta\left[1-\phi(c,p_1)\right]V(p_1)+\delta\phi(c,p_1)V(c)\right].\end{aligned}\tag{9}$$

Note that $\pi(p_1)-\phi(p_{t'-1},p_1)x(p_1)\geq\pi(p_1)-\phi(c,p_1)x(p_1)$ because $\phi(p_{t'-1},p_1)\leq\phi(c,p_1)$. Also, note that $\delta<1$. Thus, inequality (9) implies

$$\delta\left[1-\phi(p_{t'-1},p_1)\right]V(p_1)+\delta\phi(p_{t'-1},p_1)V(c)<\delta\left[1-\phi(c,p_1)\right]V(p_1)+\delta\phi(c,p_1)V(c).\tag{10}$$

Given that $V(p_1)\geq V(c)$,[44] inequality (10) holds only if $\phi(p_{t'-1},p_1)>\phi(c,p_1)$ which is false. Therefore, $p_{t'}$ is not optimal and the proof is complete. □

Proof of Theorem 6 Note that, because the initial price is c, the cartel price in period t satisfies

$$p_t<p_{t-1}+\sqrt{\frac{1-\alpha_0}{\alpha_1}}<p_{t-2}+2\sqrt{\frac{1-\alpha_0}{\alpha_1}}<\ldots<c+t\sqrt{\frac{1-\alpha_0}{\alpha_1}},$$

where the inequalities follow from Lemma 3. Let T^* be such that $T^*>2$ and $\frac{\delta^{T^*}}{1-\delta}<\frac{\alpha_0\gamma_R c}{2}D\left(\frac{1+c}{2}\right)$. Let

$$\alpha_1>\hat{\alpha}_1^C=\max\left\{\left[T^*\frac{\sqrt{1-\alpha_0}}{\alpha_0\gamma_R c}\right]^2,\frac{4\left(1-\alpha_0\right)}{(1-c)^2}\right\}.\tag{11}$$

We show that collusion is unprofitable when $\alpha_1>\hat{\alpha}_1^C$, and therefore a cartel does not form for all $\delta<1$.

Suppose a cartel forms for some $\alpha_1>\hat{\alpha}_1^C$. Let $\left\{p_t\right\}_{t=1}^{\infty}$ denote the cartel price path. First, observe that

[44] $V(p_1)\geq V(c)$ because the probability of detection is a decreasing (or constant) function of the price in the previous period.

$$
\begin{aligned}
p_1 &< c + \sqrt{\frac{1-\alpha_0}{\alpha_1}} \\
&< c + \sqrt{\frac{1-\alpha_0}{\frac{4(1-\alpha_0)}{(1-c)^2}}} \\
&= c + \sqrt{\frac{(1-c)^2}{4}} \\
&= c + \frac{1-c}{2} = \frac{1+c}{2},
\end{aligned}
$$

where the first inequality follows from Lemma 3 and the second inequality follows from inequality (11).

Next, observe that

$$
\begin{aligned}
V(c) &= \pi(p_1) - \phi(c,p_1)x_R(p_1) + \phi(c,p_1)\delta V(c) + \left[1 - \phi(c,p_1)\right]\delta V(p_1) \\
&\leq \pi(p_1) - \phi(c,p_1)x_R(p_1) + \delta V(p_1) \\
&= \pi(p_1) - \phi(c,p_1)x_R(p_1) \\
&\quad + \delta\left[\pi(p_2) - \phi(p_1,p_2)x_R(p_2) + \phi(p_1,p_2)\delta V(c) + \left[1 - \phi(p_1,p_2)\right]\delta V(p_2)\right] \\
&\leq \pi(p_1) - \phi(c,p_1)x_R(p_1) \\
&\quad + \delta\left[\pi(p_2) - \phi(p_1,p_2)x_R(p_2) + \delta V(p_2)\right] \\
&\leq \sum_{t=1}^{\infty} \delta^{t-1}\left[\pi(p_t) - \phi(p_{t-1},p_t)x_R(p_t)\right] \\
&\leq \sum_{t=1}^{\infty} \delta^{t-1}\left[\pi(p_t) - \alpha_0 x_R(p_t)\right],
\end{aligned}
$$

where the first, second and third inequalities follow from the observation that $V(p_t) \geq V(c)$ for all t.[45] The last inequality follows from $\phi \geq \alpha_0$. Therefore,

[45] $V(p_t) \geq V(c)$ because the probability of detection is a decreasing (or constant) function of the price in the previous period.

$$
\begin{aligned}
V(c) &\leq \sum_{t=1}^{\infty} \delta^{t-1}\left[D(p_t)(p_t - c) - \alpha_0\gamma_R D(p_t)p_t\right] \\
&= \sum_{t=1}^{\infty} \delta^{t-1}\left[D(p_t)\left(p_t\left[1-\alpha_0\gamma_R\right] - c\right)\right] \\
&= \underbrace{D(p_1)\left(p_1\left[1-\alpha_0\gamma_R\right] - c\right)}_{A} \\
&\quad + \underbrace{\sum_{t=2}^{T^*} \delta^{t-1}\left[D(p_t)\left(p_t\left[1-\alpha_0\gamma_R\right] - c\right)\right]}_{B} \\
&\quad + \underbrace{\sum_{t=T^*+1}^{\infty} \delta^{t-1}\left[D(p_t)\left(p_t\left[1-\alpha_0\gamma_R\right] - c\right)\right]}_{C}.
\end{aligned}
$$

We show $A + C < 0$ and $B \leq 0$which implies $V(c) < 0$ and, therefore, a cartel does not form. First, note that

$$
\begin{aligned}
A &= D(p_1)\left(p_1\left[1-\alpha_0\gamma_R\right] - c\right) \\
&\leq D(p_1)\left[\left(c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right)\left[1-\alpha_0\gamma_R\right] - c\right] \\
&= D(p_1)\left[c + \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} - c\alpha_0\gamma_R - \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\alpha_0\gamma_R - c\right] \\
&= D(p_1)\left[\frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} - c\alpha_0\gamma_R - \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\alpha_0\gamma_R\right] \\
&\leq D(p_1)\left[\frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} - c\alpha_0\gamma_R\right] \\
&< -\frac{c\alpha_0\gamma_R}{2} D(p_1) \\
&< -\frac{c\alpha_0\gamma_R}{2} D\left(\frac{1+c}{2}\right)
\end{aligned}
$$

where the last inequality follows from $p_1 < \frac{1+c}{2}$ and the second to last inequality follows from Eq. (11):

$$
\begin{aligned}
&\left[\frac{T^* \sqrt{1-\alpha_0}}{c\alpha_0\gamma_R}\right]^2 < \alpha_1 \\
&\implies \left[\frac{2\sqrt{1-\alpha_0}}{c\alpha_0\gamma_R}\right]^2 < \alpha_1 \\
&\implies \frac{2\sqrt{1-\alpha_0}}{c\alpha_0\gamma_R} < \sqrt{\alpha_1} \\
&\implies \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} < \frac{c\alpha_0\gamma_R}{2} \\
&\implies \frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} - c\alpha_0\gamma_R < -\frac{c\alpha_0\gamma_R}{2}.
\end{aligned}
$$

Thus, we have shown $A < -\frac{c\alpha_0\gamma_R}{2} D\left(\frac{1+c}{2}\right)$. $B \leq 0$ because

$$
\begin{aligned}
B &= \sum_{t=2}^{T^*} \delta^{t-1}\left[D(p_t)\left(p_t\left[1-\alpha_0\gamma_R\right]-c\right)\right] \\
&\leq \sum_{t=2}^{T^*} \delta^{t-1} D(p_t)\left[\left(c + t\frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right)\left[1-\alpha_0\gamma_R\right] - c\right] \\
&= \sum_{t=2}^{T^*} \delta^{t-1} D(p_t)\left[t\frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} - \alpha_0\gamma_R c - \alpha_0\gamma_R t\frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}\right] \\
&\leq \sum_{t=2}^{T^*} \delta^{t-1} D(p_t)\left[t\frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} - \alpha_0\gamma_R c\right] \\
&\leq \sum_{t=2}^{T^*} \delta^{t-1} D(p_t)\left[T^*\frac{\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}} - \alpha_0\gamma_R c\right] \\
&\leq 0,
\end{aligned}
$$

where the last inequality follows from Eq. (11):

$$
\begin{aligned}
\alpha_1 &> \left[T^*\frac{\sqrt{1-\alpha_0}}{\alpha_0\gamma_R c}\right]^2 \\
\sqrt{\alpha_1} &> T^*\frac{\sqrt{1-\alpha_0}}{\alpha_0\gamma_R c} \\
\alpha_0\gamma_R c &> \frac{T^*\sqrt{1-\alpha_0}}{\sqrt{\alpha_1}}.
\end{aligned}
$$

Lastly, note that

$$
\begin{aligned}
C &= \sum_{t=T^*+1}^{\infty} \delta^{t-1}\left[D(p_t)\left(p_t\left[1-\alpha_0\gamma_R\right]-c\right)\right] \\
&\le \sum_{t=T^*+1}^{\infty} \delta^{t-1} \\
&= \frac{\delta^{T^*}}{1-\delta} < \frac{\alpha_0\gamma_R c}{2} D\left(\frac{1+c}{2}\right),
\end{aligned}
$$

where the first inequality follows from the observation that per-period profit is bounded above by 1. The second inequality follows from the definition of T^*. Therefore, $A + C < -\frac{c\alpha_0\gamma_R}{2} D\left(\frac{1+c}{2}\right) + \frac{c\alpha_0\gamma_R}{2} D\left(\frac{1+c}{2}\right) = 0$. Thus, collusion is not profitable, and no cartel forms when $\alpha_1 > \hat{\alpha}_1^C$. This implies $\delta_R \to 1$ as $\alpha_1 \to \infty$.

Next, we show $\delta_R \to 1$ monotonically as $\alpha_1 \to \infty$. Let $\delta_R(\alpha_1)$ denote the critical discount factor when the sensitivity of the probability of detection to the change in price is α_1. Suppose $\delta_R(\alpha_1') > \delta_R(\alpha_1'')$ for $\alpha_1' < \alpha_1''$. Let δ be such that $\delta_R(\alpha_1'') < \delta < \delta_R(\alpha_1')$. A cartel forms when the discount factor is δ if $\alpha_1 = \alpha_1''$ and does not form if $\alpha_1 = \alpha_1'$. Let $\{p_t''\}_{t=1}^{\infty}$ denote the cartel price path when $\alpha_1 = \alpha_1''$ and the discount factor is δ. Let $W_R(p_{t-1}, \{p_t\}_{t=1}^{\infty}; \alpha_1, \delta)$ denote the payoff from collusion when the cartel price path is $\{p_t\}_{t=1}^{\infty}$, the prior period price is p_{t-1}, the sensitivity of the probability of detection to the change in price is α_1 and the discount factor is δ.

Note that

$$
\pi^D(p_t'') \le W_R(p_{t-1}'', \{p_t''\}_{t=1}^{\infty}; \alpha_1'', \delta) \le W_R(p_{t-1}'', \{p_t''\}_{t=1}^{\infty}; \alpha_1', \delta)
$$

for all t where the first inequality follows from the sustainability of collusion when $\alpha_1 = \alpha_1''$. The second inequality follows from $\alpha_1' < \alpha_1''$. Also, note that

$$
0 < W_R(c, \{p_t''\}_{t=1}^{\infty}; \alpha_1'', \delta) \le W_R(c, \{p_t''\}_{t=1}^{\infty}; \alpha'_1, \delta),
$$

where the first inequality follows from the profitability of collusion when $\alpha_1 = \alpha_1''$. The second inequality follows from $\alpha_1' < \alpha_1''$. Thus, collusion is sustainable and profitable with a price path of $\{p_t''\}_{t=1}^{\infty}$ when $\alpha_1 = \alpha_1'$ and the discount factor is δ. Therefore, a cartel forms when $\alpha_1 = \alpha_1'$ which implies $\delta > \delta_R(\alpha_1')$, a contradiction. We therefore conclude that $\delta_R \to 1$ monotonically as $\alpha_1 \to \infty$. □

Proof of Theorem 7 First, consider the case of $\delta > \delta_O$. The proof follows from the observation that, by Theorems 5 and 6, there exists an $\tilde{\alpha}_1^C$ such that $\delta_R > \delta > \delta_O$ if $\alpha_1 > \tilde{\alpha}_1^C$. Next, consider the case of $\delta \le \delta_O$. When $\delta \le \delta_O$, a cartel does not form under overcharge-based penalties (see the proof of Theorem 5). By Theorem 6, there exists an $\tilde{\alpha}_1^C$ such that $\delta_R > \delta$ if $\alpha_1 > \tilde{\alpha}_1^C$. Thus, a cartel does not form under either penalty type when $\alpha_1 > \tilde{\alpha}_1^C$ which implies $CS_R = CS_O$ and $TS_R = TS_O$. □

Appendix 2: Additional Derivations

In this section, we derive expressions for consumer surplus and total surplus.

Let the cartel's price path be denoted $\{p_t\}_{t=1}^{\infty}$. Let CS_t denote the expected present discounted value of consumer surplus when the price in the prior period was p_{t-1}. CS_1 is the expected present discounted value of consumer surplus when the price in the prior period was c: the initial period or any period immediately after detection. Let $CS(p_t)$ denote per-period consumer surplus when the cartel price is p_t and let $\phi_t = \phi(p_{t-1}, p_t)$.

$$CS_t = CS(p_t) + \delta(1-\phi_t)CS_{t+1} + \delta\phi_t CS_1.$$

If undetected in period t (which happens with probability $(1-\phi_t)$), the cartel continues into the next period (period $t+1$) with an initial price of p_t. If detected (which happens with probability ϕ_t), the cartel reforms with an initial price of c. Note that

$$\begin{aligned} CS_1 &= CS(p_1) + \delta(1-\phi_1)CS_2 + \delta\phi_1 CS_1 \\ &= CS(p_1) + \delta(1-\phi_1)\big(CS(p_2) + \delta(1-\phi_2)CS_3 + \delta\phi_2 CS_1\big) + \delta\phi_1 CS_1 \\ &= CS(p_1) + \delta(1-\phi_1)CS(p_2) + \delta^2(1-\phi_2)(1-\phi_1)CS_3 + \delta^2\phi_2(1-\phi_1)CS_1 + \delta\phi_1 CS_1. \end{aligned}$$

Continuing this pattern yields

$$CS_1 = \sum_{t=1}^{\infty} \delta^{t-1}\left(\prod_{\tau=1}^{t-1}(1-\phi_\tau)\right)CS(p_t) + \sum_{t=1}^{\infty}\delta^t\phi_t\left(\prod_{\tau=1}^{t-1}(1-\phi_\tau)\right)CS_1.$$

Solving for CS_1 yields

$$CS_1 = \frac{\sum_{t=1}^{\infty}\delta^{t-1}\left(\prod_{\tau=1}^{t-1}(1-\phi_\tau)\right)CS(p_t)}{1-\sum_{t=1}^{\infty}\delta^t\phi_t\left(\prod_{\tau=1}^{t-1}(1-\phi_\tau)\right)},$$

where we let $\prod_{\tau=1}^{0}\left[1-\phi_\tau\right] = 1$ as a convention.[46]

Computations analogous to those above show that the expected present discounted value of total surplus—the sum of consumer surplus and aggregate cartel profit—is

$$TS_1 = \frac{\sum_{t=1}^{\infty}\delta^{t-1}\left(\prod_{\tau=1}^{t-1}(1-\phi_\tau)\right)TS(p_t)}{1-\sum_{t=1}^{\infty}\delta^t\phi_t\left(\prod_{\tau=1}^{t-1}(1-\phi_\tau)\right)} \tag{12}$$

where $\phi_t = \phi(p_{t-1}, p_t)$ and $TS(p_t)$ is total surplus when the cartel price is p_t.[47]

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11151-022-09877-8.

[46] If no cartel forms, then $p_t = c$ for all $t \geq 1$, and $CS_1 = \frac{CS(c)}{1-\delta}$.

[47] Specifically, $TS(p) = CS(p) + N\pi(p)$. If no cartel forms, then $p_t = c$ for all $t \geq 1$, and $TS_1 = \frac{TS(c)}{1-\delta}$.

Acknowledgements We thank David Sappington, Jeongwoo Lee, two anonymous refrees and the Editor Lawrence J. White for excellent comments and suggestions.

References

Akyapi, B. & Turner, D. (2021). On the design of cartel penalties. In *Working paper*.

Bageri, V., & Katsoulacos, Y. (2014). A simple quantitative methodology for the setting of optimal fines by antitrust and regulatory authorities. *European Competition Journal, 10*(2), 253–278.

Bageri, V., Katsoulacos, Y., & Spagnolo, G. (2013). The distortive effects of antitrust fines based on revenue. *The Economic Journal, 123*(572), F545–F557.

Block, M. K., Nold, F. C., & Sidak, J. G. (1981). The deterrent effect of antitrust enforcement. *Journal of Political Economy, 89*(3), 429–445.

Bolotova, Y., Connor, J. M., & Miller, D. J. (2009). Factors influencing the magnitude of cartel overcharges: An empirical analysis of the us market. *Journal of Competition Law and Economics, 5*(2), 361–381.

Bos, I., Davies, S., Harrington, J. E., Jr., & Ormosi, P. L. (2018). Does enforcement deter cartels? A tale of two tails. *International Journal of Industrial Organization, 59,* 372–405.

Chen, Z., & Rey, P. (2013). On the design of leniency programs. *The Journal of Law and Economics, 56*(4), 917–957.

Christie, W. G., & Schultz, P. H. (1995). Policy watch: Did Nasdaq market makers implicitly collude? *Journal of Economic Perspectives, 9*(3), 199–208.

Dargaud, E., Mantovani, A., & Reggiani, C. (2016). Cartel punishment and the distortive effects of fines. *Journal of Competition Law & Economics, 12*(2), 375–399.

Garrod, L., & Olczak, M. (2018). Explicit vs tacit collusion: The effects of firm numbers and asymmetries. *International Journal of Industrial Organization, 56,* 1–25.

Harrington, J. E., Jr. (2004). Cartel pricing dynamics in the presence of an antitrust authority. *RAND Journal of Economics, 35,* 651–673.

Harrington, J. E., Jr. (2005). Optimal cartel pricing in the presence of an antitrust authority. *International Economic Review, 46*(1), 145–169.

Houba, H., Motchenkova, E., & Wen, Q. (2010). Antitrust enforcement with price-dependent fines and detection probabilities. *Economics Bulletin, 30*(3), 2017–2027.

Houba, H., Motchenkova, E., & Wen, Q. (2012). Competitive prices as optimal cartel prices. *Economics Letters, 114*(1), 39–42.

Houba, H., Motchenkova, E., & Wen, Q. (2015). The effects of leniency on cartel pricing. *The BE Journal of Theoretical Economics, 15*(2), 351–389.

Katsoulacos, Y., Motchenkova, E., & Ulph, D. (2015). Penalizing cartels: The case for basing penalties on price overcharge. *International Journal of Industrial Organization, 42,* 70–80.

Katsoulacos, Y., Motchenkova, E., & Ulph, D. (2018). Sophisticated revenue-based cartel penalties vs overcharge-based penalties. In *School of Business and Economics Research Memorandum*.

Katsoulacos, Y., Motchenkova, E., & Ulph, D. (2019a). Penalizing cartels: A spectrum of regimes. *Journal of Antitrust Enforcement, 7*(3), 339–351.

Katsoulacos, Y., Motchenkova, E., & Ulph, D. (2019b). Public and private antitrust enforcement for cartels: Should there be a common approach to sanctioning based on the overcharge rate? *Revista de Economia Contemporanea, 23*(2), 1–19.

Katsoulacos, Y., Motchenkova, E., & Ulph, D. (2020a). Combining cartel penalties and private damage actions: The impact on cartel prices. *International Journal of Industrial Organization, 73,* 1–18.

Katsoulacos, Y., Motchenkova, E., & Ulph, D. (2020b). Penalising on the basis of the severity of the offence: A sophisticated revenue-based cartel penalty. *Review of Industrial Organization, 57,* 627–646.

Katsoulacos, Y., & Ulph, D. (2013). Antitrust penalties and the implications of empirical evidence on cartel overcharges. *The Economic Journal, 123*(572), F558–F581.

Levenstein, M. C., & Suslow, V. Y. (2006). What determines cartel success? *Journal of Economic Literature, 44*(1), 43–95.

Levenstein, M., Suslow, V. Y., & Oswald, L. J. (2004). Contemporary international cartels and developing countries: Economic effects and implications for competition policy. *Antitrust Law Journal, 71,* 801–852.

Motta, M., & Polo, M. (2003). Leniency programs and cartel prosecution. *International Journal of Industrial Organization, 21*(3), 347–379.