Loughborough University

This item was submitted to Loughborough's Research Repository by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Explicit vs tacit collusion: The effects of firm numbers and asymmetries

PLEASE CITE THE PUBLISHED VERSION

https://doi.org/10.1016/j.ijindorg.2017.10.006

PUBLISHER

© Elsevier

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This paper was accepted for publication in the journal International Journal of Industrial Organization and the definitive published version is available at https://doi.org/10.1016/j.ijindorg.2017.10.006.

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Garrod, Luke, and Matthew Olczak. 2017. “Explicit Vs Tacit Collusion: The Effects of Firm Numbers and Asymmetries”. Loughborough University. https://hdl.handle.net/2134/27366.

Explicit vs Tacit Collusion: The Effects of Firm Numbers and Asymmetries*

Luke Garrod† and Matthew Olczak‡

August 31, 2017

Abstract

In an infinitely repeated game where firms with (possibly asymmetric) capacity constraints can make secret price cuts, we analyse the incentives for explicit collusion when firms can alternatively collude tacitly. Tacit collusion can involve price wars on the equilibrium path. Explicit collusion involves firms secretly sharing their private information to avoid such price wars, but this is illegal and runs the risk of sanctions. We find that, in contrast to the conventional wisdom but consistent with some empirical evidence, illegal cartels are least likely to arise in markets with a few symmetric firms, because tacit collusion is relatively more appealing in such markets. We discuss the implications for anti-cartel enforcement policy.

JEL classification: D43, D82, K21, L44

Key words: cartels, tacit collusion, imperfect monitoring, capacity constraints

*We are grateful for comments from two anonymous referees and co-editor Giacomo Calzolari. We have also benefitted from the comments of seminar participants at the Centre for Competition Policy, the Bergen Center for Competition Law and Economics (BECCLE) conference 2016, the Competition and Regulation European Summer School (CRESSE) conference 2016, and the European Association of Research in Industrial Economics (EARIE) conference 2016. The usual disclaimer applies.

†Corresponding author. School of Business and Economics, Loughborough University, LE11 3TU, UK, email: l.garrod@lboro.ac.uk

‡Aston Business School, Aston University, Birmingham, B4 7ET, UK, email: m.olczak@aston.ac.uk

1 Introduction

Under competition law, there is an important distinction between explicit and tacit collusion. Explicit collusion is where a group of firms directly communicate with each other, usually with the intention of coordinating and/or monitoring their actions to raise profits above competitive levels. This is prohibited in the EU by Article 101 and the US by the Sherman Act. Hence, detected cartels usually face hefty sanctions to punish and deter such conduct, and undetected cartels are incentivised to reveal themselves through leniency programmes that grant immunity to the first cartel member to turn themselves in. In contrast, tacit collusion is where firms coordinate and monitor their actions without such direct communication.[1] This is not usually considered illegal, so firms guilty of tacit collusion should not face any penalties, despite their conduct causing similar economic effects to explicit collusion.[2]

Despite its importance in terms of the law, the role of communication is often absent in theories of collusion (a point lamented by Harrington, 2008). Consequently, such theories have been presumed to be applicable to both forms of collusion, with the same models often used to derive predictions in applied work that relate specifically to tacit collusion (see Ivaldi *et al.*, 2003) or to explicit collusion (see Grout and Sonderegger, 2005). While this approach has undoubtedly played a key role in developing an understanding of the circumstances in which tacit or explicit collusion is likely to arise, it has the potential to lead to incorrect predictions, especially those related to explicit collusion. The reason is that "parties might be more likely to engage in overtly collusive practices specifically in those circumstances that are predicted by the theory as being adverse to collusion" because "the need for cartel members to communicate intensifies precisely when collusion is harder to sustain" (Grout and Sonderegger, 2005, p.36). Yet, such incentives are not captured by theories of collusion in which communication is not modelled.

Our aim in this paper is to develop a simple framework that captures the incentives for explicit collusion when firms can alternatively collude tacitly, so we can revisit the question: under which types of market structures are cartels most likely to arise? There is a need to revisit this question due to mounting evidence that some cartels do not seem to arise in markets that standard theory predicts. For instance, a well-established result of collusion theory is that when there are a large number of symmetric firms or when firms are asymmetric, it is more difficult to prevent some members from deviating from the collusive agreement. Consequently,

[1] Throughout the paper, we use tacit collusion as synonymous for all other terms that have been used to refer to collusion without explicit communication, such as tacit coordination and conscious parallelism, for example.

[2] This approach to the law is arguably desirable because it ensures legal certainty over prohibited conduct, and it prevents competitive behaviour from being punished erroneously (see Motta, 2004, p.185-190).

the conventional wisdom is that cartels are most likely to arise in markets with a few firms, often considered to be no more than three or four, that are relatively symmetric. However, there is a long established literature into the features of prosecuted cartels in the US that has uncovered some puzzling results (see Levenstein and Suslow, 2006, for a review). In particular, across a number of different samples from throughout the 20th century, the median number of cartel members varies between six and ten, and there is even evidence that cartels with many members are likely to last longer than those with fewer (see Posner, 1970; Hay and Kelley, 1974; Fraas and Greer, 1977). More recently, Davies and Olczak (2008) analysed the market shares of cartels prosecuted by the European Commission between 1990 and 2006. They found that the medium number of firms was five and that the extent of the asymmetries among the cartel members were so large that it "calls into question whether symmetry of market shares is a pervasive feature of real world cartels" (p.198).

One possible explanation for this puzzling evidence, first raised by Posner (1970), is that it results from a sample selection bias: cartels with few relatively symmetric firms may be harder to detect and are consequently under-represented in any sample of prosecuted cartels. However, evidence that is inconsistent with the conventional wisdom is still observed for samples of legal cartels, where no such selection bias should exist (see Dick, 1996; and Symeonidis, 2003). An alternative explanation raised by Hay and Kelly (1974, p.14) is that: "firms in highly concentrated industries often do not need to collude explicitly, but can rely on tacit collusion". This is the conjecture that we explore in this paper, and it is interesting to note that the (admittedly limited) evidence on tacit collusion is more consistent with the predictions of standard theory. For example, Davies *et al.* (2011) show that the European Commission's interventions in mergers due to the increased likelihood of tacit collusion have almost always been confined to cases where there would have been only two relatively symmetric players post-merger. Similarly, numerous laboratory experiments have shown that collusion without communication is unsuccessful at raising prices above competitive levels if there are more than two participants or if there are asymmetries among them (see for example, Huck *et al.*, 2004; and Fonseca and Normann, 2008).

We explore this issue by analysing an infinitely repeated game with asymmetrically capacity constrained firms that have the potential to make secret price cuts. This may be the case, for example, in upstream business-to-business markets where transaction prices can be unrelated to posted prices. We draw on earlier work (Garrod and Olczak, 2017), where we analysed tacit collusion in a private monitoring setting. In this setting, similar to Stigler (1964), market demand is unobservable and fluctuates over time, and firms never directly observe their rivals' prices or

sales. Consequently, without more information, firms must monitor a collusive agreement through their own sales, and, consistent with other models of imperfect monitoring (see Green and Porter, 1984; and Tirole, 1988), the firms may need to initiate costly price wars on the equilibrium path when at least one firm receives sufficiently low sales. In this paper, we build on our previous analysis by also modelling explicit collusion in the same framework. In this form of collusion, each firm can secretly share its private information with their rivals to improve their ability to monitoring each other's actions. As a result, explicit collusion can raise profits relative to tacit collusion by avoiding price wars, but this is illegal and runs the risk of sanctions. Thus, this tradeoff is likely to capture the incentives of cartels that are set up with the main purpose of monitoring each other's sales.[3]

We analyse how this tradeoff changes as asymmetries in capacities between firms become greater. Similar to the previous literature, we show that both forms of collusion are easiest to sustain when capacities are distributed symmetrically. However, in our model, under explicit collusion, firms will always optimally share the monopoly profit when such collusion is sustainable. In contrast, under tacit collusion, the optimal equilibrium profits depend upon the capacity of the smallest firm. The reason is that under tacit collusion deviations by the smallest firm are most difficult for rivals to detect, because each rival's resultant sales are most similar to its collusive sales. Thus, monitoring is more difficult when the smallest firm has less capacity, so punishment phases occur more often on the equilibrium path. Consequently, tacit collusion is most profitable in a symmetric duopoly, when the smallest firm is as large as possible, and as the smallest firm gets smaller, the profits from tacit collusion fall. As a result, the marginal benefit of explicit collusion is smaller in markets where tacit collusion is more profitable. Therefore, in contrast to the conventional wisdom, the incentive to form a cartel can be small in markets with a few symmetric firms, and it increases when there are greater asymmetries between the smallest and largest firm or when there are a greater number of symmetric firms. This is similar to the experimental results of Fonseca and Normann (2012), who show that communication in laboratory markets with four or six symmetric participants is more effective in facilitating collusion than those with just two.

Our analysis has three main policy implications. First, it questions the effectiveness of structural screens that use industry characteristics associated with collusion to search proactively for as yet undetected cartels (see for example Petit, 2012). Our results suggest that if such screens

[3]For example, see the discussion by Harrington (2006) of the following cartels: carbonless paper, choline chloride, copper plumbing tubes, graphite electrodes, plasterboard, vitamins, and zinc phosphate.

are constructed according to the conventional wisdom, then they have the potential to highlight markets conducive to tacit collusion, where explicit collusion is unnecessary, and to overlook markets where cartels are most likely to occur. Second, if the counterfactual to a cartel is assumed to be competition, such that the possibility of tacit collusion absent a cartel is ignored, then both the fines necessary to deter cartels as well as the benefits from deterrence will be overestimated. Third, we show that basing cartel fines on revenue, as is currently the case in the EU and US (see Bageri *et al.*, 2013), is only effective at deterring cartels that cause the smallest marginal detriment to consumer surplus. In contrast, an alternative approach where fines reflect the marginal harm caused by the cartel could be more effective at deterring cartels that cause the greatest marginal detriment to consumer surplus.

The rest of the paper is organised as follows. We review the related literature in Section 2. In section 3, we set out the assumptions of the model and outline the static Nash equilibrium. In section 4, we analyse the two forms of collusion and consider how the capacity distribution affects them separately. In section 5, we analyse firms' incentives to form a cartel when firms can alternatively collude tacitly, and examine how such incentives change as the capacity distribution varies. In section 6, we discuss the implications for anti-cartel enforcement, and conclude in section 7. All proofs are relegated to the appendix.

2 Related Literature

Our paper is related to a small number of papers that model the difference between explicit and tacit collusion (see Athey and Bagwell, 2001, 2008; Martin, 2006; Mouraviev, 2013; Awaya and Krishna, 2014; and Spector, 2015). While the underlying competition game and the information exchanged among the firms varies from model to model, the general approach of this literature is to analyse how exchanging private information facilitates collusion by improving the ability of firms to monitor each other's actions. This is also our approach, but the novelty of our model is that it is the only one in which firms can be *ex ante* asymmetric. Consequently, our model is ideal to consider how the effects of communication on collusion vary with firm numbers and asymmetries, which has not been adequately addressed by the previous literature.

The closest paper to ours is Martin (2006), who analyses the difference between explicit and tacit collusion when symmetric firms exchange verifiable sales information in a Cournot model similar to that of Green and Porter (1984). His main result is that higher fines for collusion can reduce the profits from explicit collusion relative to tacit, provided the probability that tacit

collusion is erroneously prosecuted is sufficiently low. In a similar setting to that just described, but without the chance of prosecution for tacit collusion, Mouraviev (2013) investigates how frequently firms will want to exchange verifiable information with each other when such conduct causes a positive probability of detection. His simulations show that firms may meet more or less often as the number of symmetric firms rises.

The focus of the other papers (Athey and Bagwell, 2001, 2008; Awaya and Krishna, 2014; and Spector, 2015) has been on the role of communication in facilitating explicit collusion over tacit collusion when the information exchanged is non-verifiable. This is closely related to research into the role of communication that solely focuses on explicit collusion (see for example Harrington and Skrzypacz, 2011, and references therein). While we assume in the main body of the paper that sales are verifiable for simplicity, we show in the working paper version of this article (Garrod and Olczak, 2016) that this assumption is not crucial in our framework. The reason is that firms cannot gain under explicit collusion by undercutting the collusive price and submitting false sales reports when such information is non-verifiable. Intuitively, a non-deviant firm can infer from its own sales what other firms will report if they have set the collusive price. In contrast, a deviant will sell its full capacity, so its sales provide it with no information on the level of the market demand. Consequently, the deviant does not know the level of sales that it should report to masquerade as a non-deviant firm.

Finally, our paper is also related to the theoretical literature that analyses collusion when firms are asymmetrically capacity constrained (see Compte *et al.,* 2002; Vasconcelos, 2005; and Bos and Harrington, 2010 and 2015). In particular, our framework is an extension of Compte *et al.* (2002) to a setting of imperfect monitoring. Their model (which could be interpreted as explicit or tacit collusion) shows that, in contrast to the conventional wisdom, collusion can become easier to sustain as the number of symmetric firms increase, holding the total capacity constant.[4] Our model has the same feature under explicit collusion, but we also show that an increase in the number of symmetric firms can make tacit collusion more difficult to sustain and less profitable. These differences between explicit and tacit collusion seem more consistent with the evidence discussed above. Furthermore, our result relating to the impact of firm numbers on cartel activity is caused by an effect that is new to the literature. The reason is that we consider the incentives for explicit collusion when firms can alternatively collude tacitly. Thus, in our model, explicit collusion is relatively more appealing as the number of symmetric firms rises, because the profits from tacit collusion fall when monitoring is imperfect due to harsher

[4]Kühn (2012) highlights this result in a more general Bertrand-Edgeworth framework with perfect observability.

price wars on the equilibrium path.

3 The Model

To model the difference between explicit and tacit collusion under different market structures, we use the capacity-constrained private monitoring setting in Garrod and Olczak (2017), where we developed a model of tacit collusion. In this paper, we develop a model of explicit collusion in the same framework and compare the two.

3.1 Basic assumptions

Consider a market in which a fixed number of $n \geq 2$ capacity-constrained firms compete on price to supply a homogeneous product over an infinite number of periods. Firm $i \in \{1, \ldots, n\}$ can produce a unit of the product at a constant marginal cost but the maximum it can produce in any period is k_i. We denote the total industry capacity as $K \equiv \sum_i k_i$, the sum of firm i's rivals' capacities as $K_{-i} \equiv \sum_{j \neq i} k_j$, and we let $k_n \geq k_{n-1} \geq \ldots \geq k_1 > 0$, without loss of generality. In any period t, firms set prices simultaneously, where $\mathbf{p}_t = \{p_{it}, \mathbf{p}_{-it}\}$ is the vector of prices, p_{it} is the price of firm i and $\mathbf{p}_{-it}$ is the vector of prices of all of firm i's rivals. Firms have a common discount factor, $\delta \in (0, 1)$, and we normalise their marginal costs to zero.

Market demand consists of a mass of m_t (infinitesimally small) buyers, each of whom are willing to buy one unit provided the price does not exceed their reservation price, which we normalise to 1. We assume that firms do not observe m_τ, for all $\tau \in \{0, \ldots, t\}$, but they know that m_t is independently drawn from a continuous distribution $G(m)$, with mean $\widehat{m}$ and density $g(m) > 0$ on the interval $[\underline{m}, \overline{m}]$. Furthermore, firm i never observes firm j's prices, $p_{j\tau}$, or sales, $s_{j\tau}$, $j \neq i$, for all $\tau \in \{0, \ldots, t\}$. In contrast, buyers are informed of prices, so they will want to buy from the cheapest firm. Thus, this setting is consistent with a market where all buyers are willing to check the prices of every firm in each period to find discounts from posted prices, but actual transaction prices are never public information.[5]

3.2 Demand allocation and sales

We assume demand is allocated according to the proportional allocation rule. Specifically, unsupplied buyers want to buy from the firm(s) with the lowest price among those with spare capacity.

[5] Our assumptions rule out the possibility of firms pretending to be buyers in order to discover whether their rivals are offering discounts.

If the joint capacity of such firms is insufficient to supply all of the unsupplied buyers, then such capacity is exhausted, and the remaining unsupplied buyers now want to purchase from the firm(s) with the next lowest price among those with spare capacity, and so on. Whereas, if the joint capacity of such firms suffices to supply all of the unsupplied buyers, then each firm supplies an amount of buyers equal to its proportion of the joint capacity.

We also place the following potentially restrictive assumption on the capacity distribution that we justify below:

Assumption 1. $K_{-1} \leq \underline{m}$.

This says that the joint capacity of the smallest firm's rivals should not exceed the minimum market demand. This ensures that firm i's sales in period t are strictly positive, for all i and all $m_t > \underline{m}$, even if it is the highest-priced firm. An implication of Assumption 1 is that if $\underline{m} < K$, then there is a restriction on the size of the smallest firm in that it cannot be too small.

Denoting $\Omega(p_{it})$ as the set of firms that price strictly below p_{it} and $p_t^{\max} \equiv \max\{\mathbf{p}_t\}$, Assumption 1 and the proportional allocation rule together imply that firm i's sales in period t, $s_{it}(p_{it}, \mathbf{p}_{-it}; m_t)$, for any $p_{it} \leq 1$, are:

$$s_{it}\left(p_{it}, \mathbf{p}_{-it}; m_t\right) = \begin{cases} k_i & \text{if } p_{it} < p_t^{\max} \\ \min\left\{ \frac{k_i}{K - \sum_{j \in \Omega(p_{it})} k_j} \left(m_t - \sum_{j \in \Omega(p_{it})} k_j \right), k_i \right\} \geq 0 & \text{if } p_{it} = p_t^{\max}. \end{cases} \tag{1}$$

This says that a firm will supply its proportion of the residual demand if it is the highest-priced firm in the market and if capacity is not exhausted, otherwise it will supply its full capacity. This implies that firm i's expected per-period profit is $\pi_{it}\left(p_{it}, \mathbf{p}_{-it}\right) = p_{it} \int_{\underline{m}}^{\overline{m}} s_{it}\left(p_{it}, \mathbf{p}_{-it}; m\right) g(m) dm$, where we drop time subscripts if there is no ambiguity. We write $\pi_i\left(p\right) = k_i p S\left(p\right)$ if $p_j = p$ for all j, where $S\left(p\right)$ is the expected sales per unit of capacity, such that:

$$S\left(p\right) = \begin{cases} 1 & \text{if} \quad K \leq \underline{m} \\ \int_{\underline{m}}^{K} \frac{m}{K} g(m) dm + \int_{K}^{\overline{m}} g(m) dm & \text{if} \quad \underline{m} < K < \overline{m} \\ \frac{\widehat{m}}{K} & \text{if} \quad \overline{m} \leq K. \end{cases} \tag{2}$$

So, such profits are maximised for $p^m \equiv 1$.

Before moving on, we should take a moment to defend the above assumptions. First, the proportional allocation rule is not crucial to our main story, because we can generate similar results to those presented below when demand is instead allocated equally among firms with a common price. We prefer the proportional allocation rule, because it is more tractable than the

alternative, it is consistent with the literature, and there is also anecdotal evidence of cartels allocated the demand in this manner (see Vasconcelos, 2005, and Bos and Harrington, 2010). Second, Assumption 1 is imposed as it simplifies the characterisation of the mixed strategy Nash equilibrium analysis when there are more than two firms with asymmetric capacities.[6] It also simplifies the characterisation of the optimal collusive scheme under tacit collusion. However, Assumption 1 is a sufficient condition for our main results, because they can be generated without it when there are two asymmetric firms or $n \geq 2$ symmetric firms.[7] Moreover, Assumption 1 is not very restrictive for the capacity distributions where tacit and explicit collusion are likely to be substitutes. The reason is that the evidence on tacit collusion indicates that it is most likely to occur in markets with two or three relatively symmetric firms, so the smallest firm is likely to be relatively large. Finally, the main purpose of the assumption that buyers are infinitesimally small is to develop the simplest conceivable model to compare tacit and explicit collusion when firms are asymmetric. However, it is somewhat restrictive as it ensures that buyers do not have the power to allocate their demand strategically, such that it appears to colluding firms that there has been a deviation, even when there has not. Thus, an implication of this assumption is that our analysis primarily applies to settings where the supplier side of the market is concentrated and the buyer side is unconcentrated.[8]

3.3 Static Nash equilibrium

Lemma 1 states the static Nash equilibrium profits, which can result from pure or mixed strategies. An important part of the equilibrium analysis is the minimax payoff, which for all i is:

$$\overline{\pi}_i \equiv \begin{cases} \widehat{m} - K_{-i} & \text{if} \quad \overline{m} \leq K, \\ \int_{\underline{m}}^{K} (m - K_{-i})\, g(m) dm + k_i \int_{K}^{\overline{m}} g(m) dm & \text{if} \quad \underline{m} < K < \overline{m} \\ k_i & \text{if} \quad K \leq \underline{m} < \overline{m}. \end{cases}$$

The intuition is that if the realisation of market demand is below total capacity, $m_t < K$, then a firm that sets the monopoly price expects to supply the residual demand, otherwise it expects

[6] For more details on the complexities when Assumption 1 does not hold, see Hirata (2009) who characterises the static mixed strategy Nash equilibrium with three asymmetrically capacity constrained firms but without demand uncertainty.

[7] See the online appendix of our companion paper where we show that it is still possible to characterise the optimal collusive scheme when this assumption is relaxed.

[8] Similar to the results of Bergemann and Välimäki (2002), we would expect that the presence of such strategic buyers would make collusion more difficult to sustain. For example, if there were only one buyer, then it could ensure that it exhausts the capacity of at least one firm in any period where all firms set a collusive price. This would make it look to that firm's rivals that the firm had undercut them, even though it had not. Consequently, collusion would be impossible to sustain in such a setting.

to supply its full capacity.

Lemma 1. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m}$:*
i) if $\underline{m} \geq K$, the unique pure strategy Nash equilibrium profits are $\pi_i^N = k_i \ \forall \ i$;
ii) if $\underline{m} < K$, the mixed strategy Nash equilibrium profits are $\pi_i^N(k_n) = k_i \frac{\overline{\pi}_n}{k_n} \ \forall \ i$.

We fully characterise the static Nash equilibrium in our companion paper. Consequently, for brevity, we only provide the intuition here. Competition is not effective if the minimum market demand is above total capacity, $\underline{m} \geq K$, so firms set $p_i = 1$ in equilibrium and receive $\pi_i^N = k_i$ for all i. In contrast, if $\underline{m} < K$, firms have incentives to undercut each other. However, by charging $p_i = 1$, firm i can obtain $\overline{\pi}_i$, where $\overline{\pi}_i > 0$ for all $\overline{m} > \underline{m}$ and any i from Assumption 1. Then, the largest firm will never set a price below $\underline{p} \equiv \overline{\pi}_n / k_n$ in an attempt to be the lowest-priced firm. Thus, smaller firms $i < n$ can sell their full capacity with certainty by charging a price slightly below $\underline{p}$ to obtain a profit of $k_i \underline{p} \geq \overline{\pi}_i$. The mixed strategy Nash equilibrium profits are given by $\pi_i^N(k_n) = k_i \underline{p}$ and the lower bound of the support is $\underline{p}$.

4 Two Forms of Collusion

We now analyse the repeated game. In any period, each firm's prices and sales are initially private information, but they may want to share this information with their rivals to facilitate collusion. Below, we first set out the assumptions regarding the exchange of this information. In section 4.1, we solve for the optimal equilibrium when firms exchange their private information with each other, and we refer to this as collusion under explicit monitoring. In section 4.2, we then outline the optimal equilibrium analysis in the absence of this information exchange, and we refer to this as collusion under tacit monitoring. In both cases, the optimal equilibrium has the highest equilibrium profits and the lowest critical discount factor. Finally, in section 4.3, we consider how the capacity distribution affects these equilibria separately. Henceforth, we impose $\underline{m} < K$, as collusion is unnecessary otherwise from Lemma 1.

We use the term cartel to refer to a group of firms that exchange their private information with each other. We say that a cartel is active in period t, if at the start of the period there is a chance that the firms will exchange their private information. Otherwise, the cartel is inactive. An active cartel is subject to enforcement and can face fines if detected. In addition, each firm can inform the authorities of an active cartel in return for leniency. Consistent with leniency programmes in the EU and US, we assume that applying for leniency is publicly observable.

The timing of the game for every period $t \geq 0$ is as follows:

Pricing stage: Firms set prices simultaneously and then realise their sales and profits privately.

If there is not an active cartel, period t then ends and period $t+1$ begins. In contrast, if there is an active cartel, then the game continues to the communication stage.

Communication stage: Each firm simultaneously chooses whether to share its private and verifiable information with its rivals secretly and whether to inform the authorities of the cartel publicly in return for leniency. After this, enforcement is realised:

- If no firm has informed the authorities of the cartel, then with a probability $\theta \in (0,1)$ the cartel is detected and all firms are each fined $k_i F$, and with a probability $1-\theta$ the cartel is not detected and no firm is fined.
- If at least one firm has informed the authorities of the cartel, then the cartel is detected and convicted with probability 1. Leniency is given to only one informant and the competition agency selects the informant with the lowest price (or randomly selects among these informants with equal probability if there is more than one). This selected informant is not fined and all other firms are each fined $k_i F$.

Finally, period t ends and period $t+1$ begins.

There are two assumptions that are worth discussing. First, the information that firms exchange is verifiable. This assumption is made for simplicity and we can show that our results are robust to when these sales reports are non-verifiable (see Garrod and Olczak, 2016). Second, firms with more capacity receive a larger proportion of the total cartel fine. Given each firm's collusive sales are also in proportion to its capacity, this assumption is consistent with most jurisdictions, including the EU and US, where the fine for each cartel member is initially linked to the size of its sales (see International Competition Network, 2008).

4.1 Explicit monitoring

We first analyse collusion under explicit monitoring, where there is an active cartel in period 0. We consider the following strategy profile, which we refer to as explicit monitoring strategies. In any period $t \geq 0$ in which the cartel is active, each firm sets the collusive price p^m in the pricing stage. Then in the communication stage, each firm secretly shares its private information with

its rivals and does not apply for leniency. If each firm j shares $p_{jt} = p^m$ and $s_{jt} = \frac{k_j}{K}\Sigma_j s_{jt}$ for all j and if no firms apply for leniency, then the cartel remains active in period $t+1$. Otherwise, the cartel becomes inactive forever. Once the cartel is inactive, each firm prices according to the static Nash equilibrium in each period and never exchanges its private information.

There are three comments to make regarding this strategy profile. First, Nash reversion is the harshest possible punishment under our assumptions.[9] Second, the fact that firms set the monopoly price is uncontroversial in our setting, because it is easy to check that if firms set a lower price, then the equilibrium profits are lower and the critical discount factor is higher. Thus, the optimal equilibrium profits under explicit monitoring strategies will always result from firms setting the monopoly price every period. Third, a cartel does not become inactive if it is detected and convicted by the competition agency. This is a natural assumption, because there is widespread evidence of recidivism among detected cartels (see Connor, 2010) and it is consistent with the approach of a number of papers on leniency (see Motta and Polo, 2003; Spagnolo, 2005; and Chen and Rey, 2013). This implies that, under certain conditions, cartel members will want to exploit the leniency programme by applying for leniency in every period in an attempt to reduce their expected fines. As we demonstrate in our working paper, a sufficient condition to ensure that cartel members cannot profit by exploiting the leniency programme is $\theta < \frac{1}{2}$. Given the low detection rates of cartels, this seems likely to hold in most jurisdictions, hence the reason why we restrict attention to explicit monitoring strategies here.[10]

We now solve for the optimal equilibrium under explicit monitoring strategies. Given firms can observe whether or not they share information or apply for leniency, the game is one of observable actions. Let $k_i V^e$ denote firm i's expected (normalised) profits in a period in which the cartel is active, where $k_i V^e = (1-\delta)\left(\pi_i\left(p^m\right) - \theta k_i F\right) + \delta k_i V^e$ such that solving yields:

$$k_i V^e = \pi_i\left(p^m\right) - \theta k_i F. \tag{3}$$

This says that when the cartel is active firm i receives the expected per-period profits from setting p^m minus its expected fine. We must find the conditions under which no firm will deviate

[9] The reason is that, as showed by Lambson (1994), the harshest punishments under the proportional allocation rule are such that the largest firm receives the stream of profits from its minimax strategy. In our setting, the per-period minimax payoff of the largest firm is equivalent to its static Nash equilibrium profits, so it is not possible to implement a harsher punishment than Nash reversion given the proportional allocation rule.

[10] An important factor in the analysis when firms exploit the leniency programme is which informant is selected for leniency when there is more than one. We assume that the deviant with the lowest price is given leniency. This is consistent with Spagnolo (2005), who assumes that a deviant informant will be given leniency over a colluding informant. This assumption has no effect on the analysis in this paper, because firms do not apply for leniency on the equilibrium path, so deviations from the equilibrium will only ever have at most one informant.

from its prescribed strategy. Firms play the static Nash equilibrium when the cartel is inactive, so we need only consider deviations when it is active. We begin by considering deviations in the communication stage, and then we move back to the pricing stage.

First, suppose that all firms abided by their strategies until the communication stage of period t. In this stage, if firm i applies for leniency, then the cartel becomes inactive, regardless of whether it shares its private information or not, so $\delta\pi_i^N(k_n)$ represents the highest profits from this point on that firm i can get by deviating in this stage.[11] Thus, firm i's "communication" incentive compatibility constraint (ICC) can be written as:

$$k_i V^e \geq (1-\delta)\,\pi_i(p^m) + \delta\pi_i^N(k_n)\,,\ \forall\ i. \tag{4}$$

This says that firm i will not deviate in the communication stage if it cannot gain by applying for leniency, which eliminates its fine but makes the cartel inactive thereafter.

Now suppose that all firms abided by their strategies until the pricing stage of period t. Assuming the communication ICC is satisfied, firm i's "pricing" ICC is:

$$k_i V^e \geq (1-\delta)\,k_i + \delta\pi_i^N(k_n)\,,\ \forall\ i. \tag{5}$$

This says that firm i will not deviate in the pricing stage if it cannot gain by marginally undercutting p^m to supply its full capacity, k_i, and by applying for leniency in the communication stage, which makes the cartel inactive thereafter.[12]

Note that the pricing ICC is more stringent than the communication ICC for all $K > \underline{m}$, so it is the one that binds. Furthermore, if the pricing ICC holds for firm i, then it also holds for all of its rivals $j \neq i$. Thus, let $\delta_e^*(k_n, F)$ be the unique level of δ that ensures (5) binds with equality, and let $\overline{F}(k_n)$ be the unique level of F that sets $k_i V^e = \pi_i^N(k_n)$. We say that collusion under explicit monitoring is not sustainable if the profile of explicit monitoring strategies is not a subgame perfect Nash equilibrium.

[11] If firm i deviates by not sharing its private information without applying for leniency, then its resultant profits would be lower than this due to the positive expected fine.

[12] If firm i deviates in the pricing stage, the cartel will become inactive regardless of firm i's actions in the communication stage. Consequently, firm i will have a dominant strategy to apply for leniency to raise its expected profits by eliminating its fine.

Proposition 1. *For any given* $n \geq 2$, $K_{-1} \leq \underline{m} < K$ *and* $\overline{m} \geq \underline{m}$, *if* $F \in \left[0, \overline{F}(k_n)\right)$ *and* $\delta \geq \delta_e^*(k_n, F) = \frac{1-S(p^m)+\theta F}{1-\underline{p}}$, *then firm* i*'s optimal equilibrium profits under explicit monitoring strategies are* $k_i V_e^* = \pi_i(p^m) - \theta k_i F \in \left(\pi_i^N(k_n), \pi_i(p^m)\right] \ \forall\ i$. *Otherwise, collusion under explicit monitoring is not sustainable.*

4.2 Tacit monitoring

Next, we analyse collusion under tacit monitoring, where there is never an active cartel in any period.[13] In this case, firms must monitor each other through their sales and they do so using firm-specific "trigger levels". These trigger levels are determined by the largest possible sales firms $i > 1$ can make if all such firms set the same price and firm 1 undercuts to sell its full capacity. Thus, they are given by $\underline{s}_i \equiv \frac{k_i}{K} m^*(k_1, \overline{m})$ for all i, where $m^*(k_1, \overline{m}) \equiv \min\left\{\frac{K(\overline{m}-k_1)}{K_{-1}}, K\right\}$, from (1). Such trigger levels imply that in every period it is common knowledge whether all firms' realised sales are greater than their trigger levels or not.[14] This provides a binary (possibly noisy) public signal of $\mathbf{p}_t$ that firms can condition their play on, and the past realisations of these signals constitute a public history. Thus, we restrict attention to equilibria in public strategies, known as perfect public equilibria (PPE).

If $\underline{m} > m^*(k_1, \overline{m})$, then tacit monitoring is perfect in the sense that firms will only ever receive sales below their trigger levels if they are undercut. Thus, rearranging $\underline{m} > m^*(k_1, \overline{m})$ in terms of $\overline{m}$, it follows that tacit monitoring is perfect if $\overline{m} < k_1\left(\frac{K-\underline{m}}{K}\right) + \underline{m} \equiv \underline{x}(k_1) \in (\underline{m}, K)$. However, if $\overline{m} \geq \underline{x}(k_1)$, then tacit monitoring is imperfect in the sense that firms' sales will also fall below the trigger levels when firms set a common price and the realisation of market demand is sufficiently low. In this case, colluding firms do not know whether the realisation of market demand was unluckily low or whether at least one rival has undercut them. The probability that at least one firm's sales are below their trigger level is $G(m^*(k_1, \overline{m}))$ when all firms set a common price and is 1 when they all do not.

We analyse the following strategies, which we refer to as tacit monitoring strategies. There can be 'collusive periods' and 'punishment phases'. Period 0 is a collusive period. In any collusive period $t \geq 0$, each firm should set p^m. If all firms' sales are above their trigger levels in period t, then period $t+1$ is a collusive period. If at least one firm's sales are below its trigger level in period t, then firms enter a punishment phase in period $t+1$. In a punishment phase, each firm

[13] We refer the reader to Garrod and Olczak (2017) for a more detailed version of this case.

[14] The reason is that, if all firms set a common price, then either all firms' sales will exceed their respective trigger levels or they all will not. Yet, if all firms do not set a common price, then the highest-priced firms receive sales below their trigger levels and the lower-priced firm(s) can infer this from the fact that they supply their full capacity.

should play the static Nash equilibrium for T periods, and then period $t+T+1$ is a collusive period. The sequence repeats.[15]

Let $k_i V^c$ and $k_i V^p$ denote firm i's expected (normalised) profits in a collusive period and at the start of a punishment phase, respectively, where:

$$\begin{aligned} k_i V^c &= (1-\delta)\,\pi_i\left(p^m\right) + \delta\left[\left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right) k_i V^c + G\left(m^*\left(k_1, \overline{m}\right)\right) k_i V^p\right] \\ k_i V^p &= (1-\delta) \textstyle\sum_{t=0}^{T-1} \delta^t \pi_i^N\left(k_n\right) + \delta^T k_i V^c, \end{aligned} \tag{6}$$

for all i. Solving simultaneously yields:

$$k_i V^c = \pi_i^N\left(k_n\right) + \frac{(1-\delta)}{1-\delta+G\left(m^*\left(k_1,\overline{m}\right)\right)\delta\left(1-\delta^T\right)}\left(\pi_i\left(p^m\right) - \pi_i^N\left(k_n\right)\right), \tag{7}$$

where $\pi_i\left(p^m\right) \geq k_i V^c > k_i V^p$ for any $T > 0$ and $k_i V^p > \pi_i^N\left(k_n\right)$ for any $T < \infty$.

The profile of tacit monitoring strategies is a PPE if, for each period t and any public history, the strategies yield a Nash equilibrium from that date on. Firms play the static Nash equilibrium in each period of a punishment phase, so we need only consider deviations in collusive periods. The ICC for firm i in such periods is:

$$k_i V^c \geq (1-\delta)\, k_i + \delta k_i V^p, \ \forall\ i. \tag{8}$$

This says that firm i will not deviate in a collusive phase if it cannot gain by marginally undercutting p^m to supply its full capacity k_i, which triggers a punishment phase with certainty. This ICC can only ever be satisfied if $\overline{m} < K$.[16] Thus, when $\overline{m} < K$, substituting $k_i V^p$ and $k_i V^c$ into the ICC and rearranging yields:

$$\left(1 - G\left(m^*\left(k_1,\overline{m}\right)\right)\right)\frac{K}{k_n} - \frac{1}{\delta} \geq \delta^T\left[\left(1 - G\left(m^*\left(k_1,\overline{m}\right)\right)\right)\frac{K}{k_n} - 1\right]. \tag{9}$$

This is independent of k_i under the proportional allocation rule, which implies that if the ICC holds for firm i, then it also holds for all other firms $j \neq i$.

Let $\overline{x}\left(k_1, k_n\right)$ denote the unique level of $\overline{m}$ that sets the right hand-side of (9) to zero, where $\overline{x}\left(k_1, k_n\right) \in \left(\underline{x}\left(k_1\right), K\right)$. Furthermore, let $\delta_c^*\left(k_1, k_n\right)$ denote the unique level of δ that sets the left hand-side to zero, such that (9) holds for all $\delta \geq \delta_c^*\left(k_1, k_n\right)$ if $T \to \infty$. We say that collusion

[15] In our companion paper, we used the techniques of Abreu *et al.* (1986, 1990) to show that, in our framework, this profile of strategies generates the maximal PPE payoffs and the lowest critical discount factor.

[16] The reason is that if $\overline{m} \geq K$, then $m^*\left(k_1, \overline{m}\right) = K$ such that $\underline{s}_i = k_i$ for all i, so $G\left(K\right) = 1$. Then, using (6) with $G\left(K\right) = 1$, it is easy to check that (8) cannot hold.

under tacit monitoring is not sustainable if no PPE under tacit monitoring strategies exists with $k_i V_c > \pi_i^N(k_n)$.

Proposition 2. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} < K$:*
i) if $\overline{m} < \underline{x}(k_1)$ and $\delta \geq \delta_c^*(k_1, k_n) = \frac{k_n}{K}$, then firm i's optimal equilibrium profits under tacit monitoring strategies are $k_i V_c^* = \pi_i(p^m)\ \forall\ i$;
ii) if $\underline{x}(k_1) \leq \overline{m} < \overline{x}(k_1, k_n)$ and $\delta \geq \delta_c^*(k_1, k_n) = \frac{1}{1-G(m^*(k_1,\overline{m}))}\frac{k_n}{K} \in \left(\frac{k_n}{K}, 1\right)$, then firm i's optimal equilibrium profits under tacit monitoring strategies are $k_i V_c^* = \frac{k_i}{K}\left(\frac{\widehat{m}-G(m^*(k_1,\overline{m}))K}{1-G(m^*(k_1,\overline{m}))}\right) \in \left(\pi_i^N(k_n), \pi_i(p^m)\right)\forall\ i$;
iii) otherwise, collusion under tacit monitoring is not sustainable.

This says that, if tacit monitoring is perfect, such that $\overline{m} < \underline{x}(k_1)$, then the optimal punishment phase duration is $T \to \infty$, because punishment phases do not occur on the equilibrium path. Yet, if tacit monitoring is imperfect, such that $\overline{m} \in [\underline{x}(k_1), \overline{x}(k_1, k_n))$, then the optimal punishment phase duration is at the level where (9) binds with no slack.

4.3 Comparative statics

We want to analyse the effects of asymmetries on the incentive to form a cartel when firms could alternatively collude tacitly. Before doing so, it is helpful to get a clear understanding of how the capacity distribution affects both forms of collusion in isolation. In this subsection, we discuss the effects of changes to the capacity distribution holding the total capacity and the number of firms constant. This implies that when the capacity of firm j changes by a small amount, other things equal, the capacities of the other firms change to the extent that $\frac{\partial k_i}{\partial k_j} \in [-1, 0]$ for all $i \neq j$, where $\sum_{i \neq j} \frac{\partial k_i}{\partial k_j} = -1$. However, in what follows we restrict the discussion to capacity reallocations that directly affect the equilibrium analysis, and this is the case for changes to the capacity of the smallest or the largest firm.

Proposition 3. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} < K$:*
i) raising the capacity of the largest firm, k_n, strictly increases the critical discount factor under explicit monitoring strategies, $\frac{\partial \delta_c^*}{\partial k_n} > 0$, $\forall\ \overline{m} > \underline{m}$, and the critical discount factor under tacit monitoring strategies, $\frac{\partial \delta_c^*}{\partial k_n} > 0$, $\forall\ \overline{m} \in (\underline{m}, \overline{x}(k_1, k_n))$;
ii) raising the capacity of the smallest firm, k_1, strictly decreases the critical discount factor under tacit monitoring strategies and strictly increases the optimal equilibrium profits under tacit monitoring strategies, $\frac{\partial \delta_c^*}{\partial k_1} < 0$ and $\frac{\partial V_c^*}{\partial k_1} > 0$, $\forall\ \overline{m} \in [\underline{x}(k_1), \overline{x}(k_1, k_n))$.

Raising the capacity of the largest firm, k_n, increases the static Nash equilibrium profits under Assumption 1. Consequently, the punishment is weaker than before under both explicit and tacit monitoring, and this increases the critical discount factors, $\delta_e^*(k_n, F)$ and $\delta_c^*(k_1, k_n)$.

Raising the capacity of the smallest firm, k_1, does not affect the optimal equilibrium under explicit monitoring. However, it can affect the optimal equilibrium under tacit monitoring. The reason is that as k_1 gets larger, firms are better able to monitor each other, because each firm's trigger level is lower. This implies that if tacit monitoring is imperfect, then it is less likely that firms will receive sales below their trigger levels when they set a common price.[17] Consequently, punishment phases are less likely to arise on the equilibrium path than before. This implies that profits rise on the equilibrium path, other things equal, and it follows that a punishment phase that lasts for a given number of periods is relatively harsher than before. This introduces slack into the ICC, which has two implications. First, $\delta_c^*(k_1, k_n)$ decreases, implying that collusion under tacit monitoring is easier to sustain than before. Second, firms can reduce the punishment phase duration to the extent that the ICC binds with no slack, and this further increases their profits, $k_i V_c^*$.[18]

In summary, Proposition 3 shows that symmetry is ideal for both forms of collusion. The reason is that when firms are symmetric, such that $k_i = K/n$ for all i, the largest firm is as small as possible and the smallest firm is as large as possible. Differences between the two forms of collusion emerge when we consider the number of symmetric firms, n. For instance, as n increases by 1, each firm's capacity decreases by $\frac{1}{n(n+1)}$ when total capacity is held constant. Thus, using Proposition 3, collusion under explicit monitoring is actually easier to sustain as n increases, because $\delta_e^*\left(\frac{K}{n}, F\right)$ falls. This contrasts with the conventional wisdom but it is consistent with the results of Compte *et al.* (2002) and Kühn (2012). Similarly, when tacit monitoring is perfect, $\delta_c^*\left(\frac{K}{n}, \frac{K}{n}\right)$ falls as n rises for the same reason. However, if tacit monitoring is imperfect, then as n increases, collusion under tacit monitoring is less profitable but there is an ambiguous effect on $\delta_c^*\left(\frac{K}{n}, \frac{K}{n}\right)$. In fact, placing more structure on the fluctuations in market demand, we get the following:

[17] Similarly, tacit monitoring is perfect for a wider range of fluctuations in market demand.

[18] The optimal equilibrium profits under tacit monitoring strategies are independent of the capacity of the largest firm, k_n. When tacit monitoring is imperfect, this is the result of two equal offsetting effects. The first effect is that an increase in k_n raises the static Nash equilibrium profits and this increases the profits on the equilibrium path under tacit monitoring, other things equal. The second effect results from the fact that the ICC is now tighter than before, so the optimal punishment phase duration lengthens to ensure that the ICC binds with no slack.

Corollary 1. *Assume* $G(m) = \frac{m-\underline{m}}{\overline{m}-\underline{m}}$ *and* $k_i = \frac{K}{n}\ \forall\ i$. *Then if* $\underline{x}\left(\frac{K}{n}\right) \leq \overline{m} < \overline{x}\left(\frac{K}{n}, \frac{K}{n}\right)$, *the critical discount factor under tacit monitoring strategies is* $\delta_c^*\left(\frac{K}{n}, \frac{K}{n}\right) = \left(\frac{\overline{m}-m}{K-\overline{m}}\right)\left(1-\frac{1}{n}\right)$ *such that it strictly increases with the number of symmetric firms,* n.

This implies that, in contrast to collusion under explicit monitoring (as well as Compte *et al.*, 2002; and Kühn, 2012), as the number of symmetric firms increases, collusion under tacit monitoring is more difficult to sustain and less profitable, when tacit monitoring is imperfect and market demand is drawn from a uniform distribution.

5 Cartels, Firm Numbers and Asymmetries

In this section, we now use our equilibrium analysis to investigate the effects of asymmetries on the incentives to form a cartel when firms can alternatively collude tacitly. We say that it is privately optimal for a firm to be a cartel member if the optimal equilibrium profits from being a member are *strictly* greater than the optimal equilibrium profits from not being a member. Furthermore, define a cartel phase as a sequence of periods that begins the period after detection and ends the period of the next detection. Thus, the expected duration of a cartel phase is $\sum_{t=1}^{\infty} t\theta\left(1-\theta\right)^{t-1} = 1/\theta$, which implies that the lower the probability of detection, the longer the expected duration of a cartel phase.

Before comparing explicit and tacit monitoring, first consider as a benchmark the incentives to be a cartel member when the alternative to explicit monitoring is to not collude and instead play the static Nash equilibrium forever. For this benchmark case, it follows from Proposition 1 that it is privately optimal to be a cartel member for each firm i if $\delta \geq \delta_e^*\left(k_n, F\right)$ for any $F \in \left[0, \overline{F}\left(k_n\right)\right)$. This is due to the fact that $k_i V_e^* > \pi_i^N\left(k_n\right)$ implies:

$$F < \frac{1}{\theta}\left(S\left(p^m\right) - \underline{p}\right) \equiv \overline{F}\left(k_n\right). \tag{10}$$

Note that the left-hand side of (10) is, from the cartel's perspective, the expected marginal cost of a cartel phase per unit of capacity. Moreover, the right-hand side is the expected marginal benefit of a cartel phase per unit of capacity, which is the multiplication of the expected duration of a cartel phase, $1/\theta$, and the expected per-period marginal benefit per unit of capacity, $S\left(p^m\right) - \underline{p}$. Thus, the above implies that the incentive to be a cartel member is reduced as the largest firm gets larger. The reason is that the expected marginal benefit of a cartel is lower, because the static Nash equilibrium profits are higher under Assumption 1.

5.1 Explicit vs tacit monitoring

We now investigate the incentives to be a cartel member when firms could alternatively collude tacitly. Henceforth, we restrict attention to $\overline{m} < \overline{x}(k_1, k_n)$. Otherwise, collusion under tacit monitoring is not sustainable, so the benchmark case above is the appropriate comparison.

Proposition 4 establishes when collusion under tacit monitoring is more profitable than collusion under explicit monitoring, and vice versa.

Proposition 4. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} \leq \overline{m} < \overline{x}(k_1, k_n)$, there exists a unique fine per unit of capacity given by:*

$$F^*(k_1) \equiv \begin{cases} 0 & \text{if} \quad \overline{m} < \underline{x}(k_1) \\ \frac{1}{\theta}\left(\frac{\widehat{m}}{K} - V_c^*\right) \in \left(0, \overline{F}(k_n)\right) & \text{if} \quad \underline{x}(k_1) \leq \overline{m} < \overline{x}(k_1, k_n), \end{cases}$$

such that if and only if $F \geq F^*(k_1)$, the optimal profits that could be sustained as an equilibrium under tacit monitoring strategies are (weakly) greater than under explicit monitoring strategies, $V_c^* \geq V_e^*$.

This implies that it can only be privately optimal for a firm to be a cartel member if the fine per unit of capacity is sufficiently low:

$$F < \frac{1}{\theta}\left(S(p^m) - V_c^*\right) = F^*(k_1), \tag{11}$$

where $S(p^m) = \frac{\widehat{m}}{K}$ for all $\overline{m} < K$ from (2). Similar to (10), the above inequality compares the expected marginal cost of a cartel phase per unit of capacity on the left-hand side against the expected marginal benefit per unit of capacity on the right-hand side. The only difference to (10) is that V_c^* appears on the right-hand side of (11) instead of $\underline{p}$. Thus, if tacit monitoring is perfect, such that $\overline{m} < \underline{x}(k_1)$, then being a cartel member is never privately optimal for any firm. The reason is that firms are able to extract the monopoly profit through tacit monitoring, so $V_c^* = \frac{\widehat{m}}{K}$ such that the expected marginal benefit of a cartel phase is zero. However, if tacit monitoring is imperfect, such that $\overline{m} \in [\underline{x}(k_1), \overline{x}(k_1, k_n))$, then it can be privately optimal for each firm to be a cartel member. In this case, the expected marginal benefit of a cartel phase is positive, but less than the benchmark case, from $\frac{\widehat{m}}{K} > V_c^* > \underline{p}$.

Next, Proposition 5 shows that if collusion under tacit monitoring is more profitable than explicit monitoring, then it is sustainable at lower discount factors, and vice versa.

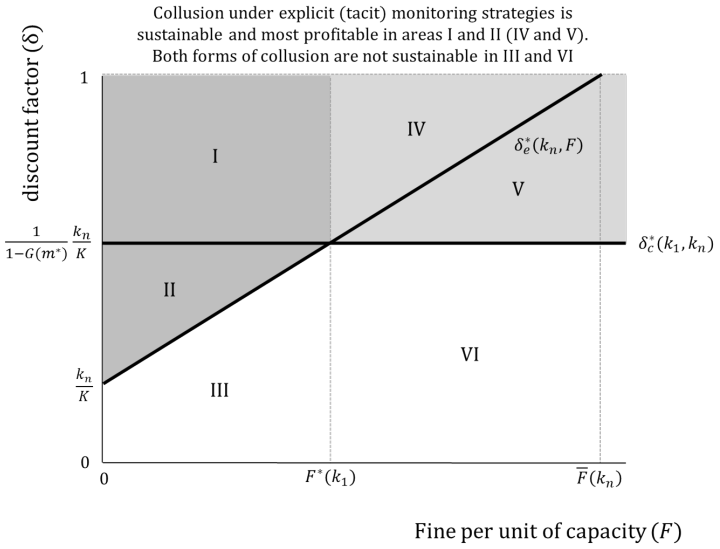


Figure 1: The parameter space of explicit monitoring

Proposition 5. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} \leq \overline{m} < \overline{x}(k_1, k_n)$, if and only if $F \geq F^*(k_1)$, then the critical discount factor under tacit monitoring is (weakly) less than under explicit monitoring, $\delta_c^*(k_1, k_n) \leq \delta_e^*(k_n, F)$.*

Proposition 4 and 5 together imply the following result.

Result 1. *If $F < F^*(k_1)$, then it is privately optimal for each firm to be a cartel member for all $\delta \geq \delta_e^*(k_n, F)$. Whereas, if $F \geq F^*(k_1)$, then it is not privately optimal for any firm to be a cartel member for any δ.*

This result is illustrated in Figure 1, which depicts $\delta_e^*(k_n, F)$ and $\delta_c^*(k_1, k_n)$ for any given fine per unit of capacity, F, assuming tacit monitoring is imperfect, such that $\overline{m} \in [\underline{x}(k_1), \overline{x}(k_1, k_n))$. Note that $\delta_c^*(k_1, k_n)$ is a horizontal line because it is independent of F, and it lies between k_n/K and 1 since it equals the former for all $\overline{m} \in (\underline{m}, \underline{x}(k_1)]$ and it tends to the latter as $\overline{m} \to \overline{x}(k_1, k_n)$. In contrast, $\delta_e^*(k_n, F)$ is linear and strictly increasing in F. These two critical discount factors are equal at $F^*(k_1)$. If $F < F^*(k_1)$, then it it is privately optimal for each firm to be a cartel member in areas I and II, where $\delta \geq \delta_e^*(k_n, F)$. In area I, both forms of collusion are sustainable, but collusion under explicit monitoring yields the highest profits. In area II, only collusion under explicit monitoring is sustainable and it is more profitable than not colluding. In area III, both forms of collusion are not sustainable. In contrast, if $F > F^*(k_1)$, it is never privately optimal

for a firm to be a cartel member for any δ. Collusion under explicit monitoring is only sustainable in area IV, but here collusion under tacit monitoring is also sustainable and yields higher profits. Consequently, we refer to areas I and II as the parameter space of explicit monitoring. For completeness, note that in area V, collusion under tacit monitoring is sustainable and is more profitable than not colluding, and in area VI, both forms of collusion are not sustainable. Finally, assuming a mean preserving spread, notice that the parameter space of explicit collusion will expand as $\overline{m}$ increases towards $\overline{x}(k_1, k_n)$, due to the fact that $\delta_c^*(k_1, k_n)$ and $F^*(k_1)$ strictly increase because monitoring is more difficult.

5.2 Comparing capacity distributions

We now analyse the effects of asymmetries in capacities on the incentives to be a cartel member when firms can alternatively collude tacitly. Holding total capacity constant, we analyse changes to the capacity distribution that affect the equilibrium analysis either by only increasing the size of the smallest firm or by only increasing the size of the largest firm. We discuss each in turn. Figure 2 builds on the illustration in Figure 1 to depict the effects of such changes on the parameter space of explicit monitoring. Changes to the capacities of the smallest and the largest firms at the same time will have a mix of the two effects discussed here.

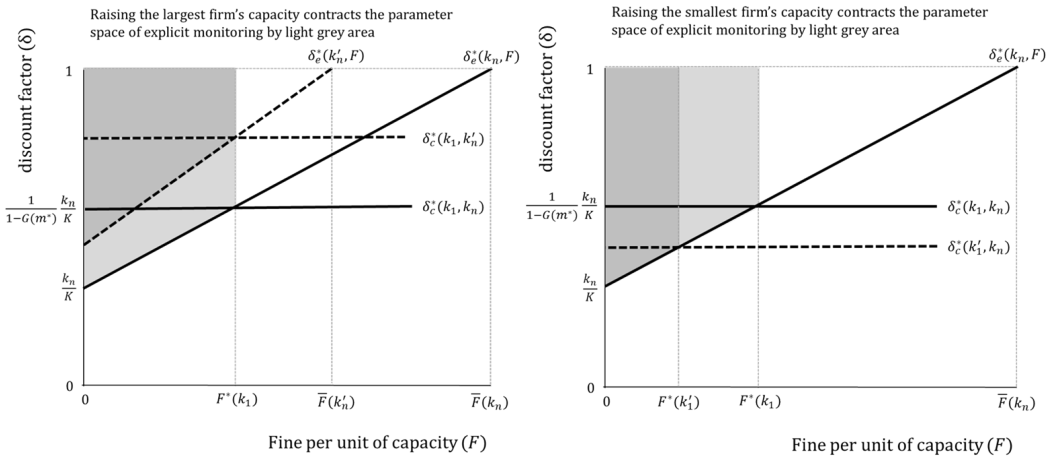


Figure 2: Changes to the capacity distribution

Increasing asymmetries between the smallest and largest firms by raising the capacity of the largest firm, k_n, other things equal, contracts the parameter space of explicit monitoring. This

contraction is illustrated by the light grey area of Figure 2(a). It follows from Proposition 3 that both critical discount factors rise. Yet, in contrast to the benchmark case, it follows from (11) that there is no effect on $F^*(k_1)$. This is due to the fact that V_c^* is independent of k_n for the reasons explained in footnote 18. Thus, all of the contraction in the parameter space of explicit monitoring results from explicit monitoring being replaced by no collusion. Consequently, this effect is similar to that in Compte *et al.* (2002), and it exists even when collusion under tacit monitoring is not sustainable.

Decreasing asymmetries between the smallest and largest firms by raising the capacity of the smallest firm, k_1, other things equal, can also contract the parameter space of explicit monitoring. This occurs only if tacit monitoring is imperfect, such that $\overline{m} \in [\underline{x}(k_1), \overline{x}(k_1, k_n))$, and the contraction is illustrated by the light grey area of Figure 2(b). Under such conditions, it follows from Proposition 3 that $\delta_c^*(k_1, k_n)$ falls and $\delta_e^*(k_n, F)$ is unaffected. Furthermore, in contrast to the benchmark case, $F^*(k_1)$ falls because V_c^* is strictly increasing in k_1, from Proposition 3. Thus, all of the contraction in the parameter space of explicit monitoring results in explicit monitoring being replaced by tacit monitoring.

Finally, we bring together the preceding analysis to consider the effects of firm numbers and asymmetries on cartel formation. Assuming that the parameter space of explicit monitoring is a good proxy for the likelihood of cartel formation and noting that decreasing the number of symmetric firms raises each firm's capacity when the total capacity is held constant, then we get the following result.

Result 2. *In contrast to the conventional wisdom, cartel formation can be more likely in markets where i) the smallest firm has less capacity, holding the capacity of the largest firm constant or ii) there are a larger number of symmetric firms.*

While i) follows immediately from the effect in Figure 2(b), the reason for ii) is that decreasing the number of symmetric firms, where $k_1 = k_n = K/n$, causes the parameter space of explicit collusion to contract through both effects in Figure 2(a) and (b). Consequently, Result 2 implies that, compared with a symmetric duopoly, where $k_1 = k_2 = K/2$, a cartel can be more likely to form in a symmetric triopoly, where $k_1 = k_3 = K/3$, or in any asymmetric triopoly with $k_1 < K/2$ and $k_3 = K/2$. Finally, due to the effect in Figure 2(a), we should note that cartels can be more likely to form in markets with a smaller largest firm, holding the capacity of the smallest firm constant. This implies that, compared to a symmetric triopoly, a cartel will be less likely to form in an asymmetric duopoly with $k_1 = K/3$ and $k_2 = 2K/3$.

6 Discussion

Our analysis provides an explanation for the puzzling evidence that prosecuted cartels do not tend to form in markets with a few, relatively symmetric firms. Of most relevance to our work is Davies and Olczak (2008), who analysed the asymmetries in market shares of members of cartels prosecuted by the European Commission between 1990 and 2006. They found that only 5% of cartels were "tacit-collusive compatible", in that their market shares were consistent with market structures that would likely lead to concerns of tacit collusion in a merger investigation. Our results suggest that cartel activity may be low in such markets, because tacit collusion is relatively more appealing when firms are symmetric. Furthermore, our analysis also suggests that we may expect greater cartel activity in markets where both the largest and the smallest firm are small. Again, this is not inconsistent with the results of Davies and Olczak (2008) as they find that 61% of their sample fell within a group of "unconcentrated" cartels, where the largest firm was relatively small and the number of firms in the cartel was relatively high.

Our results have implications for structural screens that use market characteristics to search proactively for as yet undetected cartels. Despite being credited with some success in practice, scholars are often sceptical over their effectiveness. For example, Harrington (2008) argues that, even if they correctly indicate the markets in which cartel activity is likely to occur, there is still a high chance of false positives because a collusive outcome is just one of the possible multiple equilibria. Our analysis provides a further reason to be sceptical: structural indicators may not even consistently flag up the markets in which cartel activity is more likely to occur. Instead, if the structural indicators are constructed according to the conventional wisdom, then highlighted markets may include ones in which firms prefer tacit over explicit collusion, and markets where cartels are most likely to form may be overlooked. This problem would be compounded further if proactive investigations were to rely on such methods to the extent that the probability of detection is perceived to be greater in symmetric markets, making explicit collusion even less appealing in such markets.

Our analysis also has two implications for deterrence effects of anti-cartel enforcement. First, if the counterfactual to a cartel is assumed to be competition (as in, for example, the estimates of the effectiveness of anti-cartel enforcement by Katsoulacos *et al.,* 2015a) when in fact firms could alternatively collude tacitly, then the fines necessary to deter cartels and the benefits from deterrence will be overestimated. This point is illustrated most starkly when tacit monitoring is perfect such that firms can share the monopoly profit through tacit collusion. Here, any positive

fine is sufficient to deter a cartel, but the benefit of deterrence to society is zero. Second, it follows from our model that a given fine will only deter the cartels that cause the least marginal detriment to consumer surplus. To understand the intuition, consider the example where tacit monitoring is the counterfactual. Furthermore, note that (11) implies that the level of the fine required to deter a cartel from forming increases with the expected marginal benefit of a cartel phase to its members. Given market demand is perfectly inelastic, this expected marginal benefit is equivalent to the expected marginal detriment to consumer surplus. Therefore, it follows that if a given total fine will only deter one of two different cartels, then it will deter the one that benefits its members the least and hence causes the smallest marginal detriment to consumer surplus. It follows from the preceding analysis that, if these two cartels only differ either in the size of the smallest firm or in the number of symmetric firms, then the deterred cartel will have the smallest asymmetries or the fewest symmetric firms, respectively.

Finally, the conclusion that the cartels that cause the greatest marginal harm may not be deterred is particularly pertinent for jurisdictions, such as the EU and US, where fines are initially based on the revenue of the cartel members (see Bageri *et al.*, 2013). The reason is that in our setup the expected industry revenue ($\widehat{m}$ per period) is independent of the capacity distribution, so it seems reasonable to model such an approach to fines, as we did above, by a constant total fine for all capacity distributions.[19] In contrast, according to our model, alternative approaches where fines reflect the marginal harm caused by the cartel could be more effective at deterring cartels that cause the greatest detriment to consumer surplus. Then, such cartels could receive a higher fine than other less harmful cartels, even when the total revenue of such cartels is the same. This alternative approach may involve estimating the 'overcharge', which is the difference between the cartel price and a counterfactual price absent the cartel. Commonly, the counterfactual price is assumed to be from a competitive benchmark (for example, see Katsoulacos *et al.,* 2015b). However, our analysis implies that, in some cases, a counterfactual price that takes into account that firms will collude tacitly absent the cartel may generate a more accurate (yet smaller) overcharge than a competitive counterfactual price.

7 Conclusion

We have analysed the effects of asymmetries in capacity constraints on the incentives of firms to collude explicitly when they can alternatively collude tacitly. In our private monitoring setting,

[19]Recall that how the total fine is distributed among the cartel members does vary with each firm's revenue, but this is also consistent with the current approach in the EU and US.

firms have the potential to make secret price cuts. Therefore, tacit monitoring involves each firm using its sales to monitor rivals, so price wars can occur on the equilibrium path. In contrast, explicit monitoring involves firms secretly sharing their privately observed prices and sales information to improve their ability to monitor each other. Consequently, explicit monitoring can raise profits relative to tacit monitoring by avoiding price wars, but this runs the risk of sanctions. Consistent with other models of collusion with asymmetric capacity constraints, we showed that both forms of collusion are easiest to sustain when capacities are distributed symmetrically. However, in contrast to the conventional wisdom and consistent with some evidence, we found that the incentives for firms to form an illegal cartel can be smallest in markets with a few symmetric firms. This questions whether structural screening devices that proactively search for cartels will reliably highlight the markets where cartels are most likely to occur.

References

[1] Abreu, D., Pearce, D. and Stacchetti, E. (1986) "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory*, 39(1), 251-269

[2] Abreu, D., Pearce, D. and Stacchetti, E. (1990) "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58(5), 1041-1063

[3] Athey, S. and Bagwell, K. (2001) "Optimal Collusion with Private Information," *RAND Journal of Economics*, 32(3), 428–465

[4] Athey, S. and Bagwell, K. (2008) "Collusion with Persistent Cost Shocks," *Econometrica*, 76(3), 493–540

[5] Awaya, Y. and Krishna, V. (2014) "On Tacit versus Explicit Collusion," mimeo, Penn State University

[6] Bageri, V., Katsoulacos, Y. and Spagnolo, G. (2013) "The Distortive Effects of Antitrust Fines Based on Revenue," *The Economic Journal*, 123 (572), 545-557

[7] Bergemann, D. and Välimäki, J. (2002) "Strategic Buyers and Privately Observed Prices," *Journal of Economic Theory,* 105 (2), 469-482

[8] Bos, I. and Harrington, J. (2010) "Endogenous Cartel Formation with Heterogeneous Firms," *RAND Journal of Economics*, 41(1), 92–117

[9] Bos, I. and Harrington, J. (2015) "Competition Policy and Cartel Size," *International Economic Review*, 56(1), 133-153

[10] Chen, Z. and Rey, P. (2013) "On the Design of Leniency Programs," *Journal of Law and Economics*, 56(4), 917 - 957

[11] Compte, O., Jenny, F. and Rey, P. (2002) "Capacity Constraints, Mergers and Collusion," *European Economic Review*, 46(1), 1-29

[12] Connor, J. (2010) "Recidivism Revealed: Private International Cartels 1991-2009" *Competition Policy International*, 6(2), 101-127

[13] Davies, S. and Olczak, M. (2008) "Tacit versus Overt Collusion Firm Asymmetries and Numbers: What's the Evidence?" *Competition Policy International*, 4(2), 175-200

[14] Davies, S., Olczak, M., and Coles, H. (2011) "Tacit Collusion, Firm Asymmetries and Numbers: Evidence from EC Merger Cases," *International Journal of Industrial Organization*, 29(2), 221-31

[15] Dick, A. (1996) "When Are Cartels Stable Contracts?" *Journal of Law and Economics*, 39(1), 241–8

[16] Fonseca, M. and Normann, H.-T. (2008) "Mergers, Asymmetries and Collusion: Experimental Evidence," *Economic Journal*, 118(1), 287-400

[17] Fonseca, M. and Normann, H.-T. (2012) "Explicit vs. Tacit Collusion – The Impact of Communication in Oligopoly Experiments," *European Economic Review*, 56(8), 1759-1772

[18] Fraas, A., and Greer, D. (1977) "Market Structure and Price Collusion: An Empirical Analysis," *Journal of Industrial Economics*, 26(1), 21–44

[19] Garrod, L. and Olczak, M. (2016) "Collusion, Firm Numbers and Asymmetries Revisited," CCP Working Paper 16-11

[20] Garrod, L. and Olczak, M. (2017) "Collusion under Imperfect Monitoring with Asymmetric Firms," *Journal of Industrial Economics*, forthcoming (most recent working paper at: https://mpra.ub.uni-muenchen.de/70647/)

[21] Green, E. and Porter, R. (1984) "Non-Cooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87-100

[22] Grout, P. and Sonderegger, S. (2005) "Predicting Cartels," A report for the Office of Fair Trading, Economic Discussion Paper 773

[23] Harrington, J. (2006) "How do Cartels Operate?" *Foundations and Trends in Microeconomics*, 2(1), 1-105

[24] Harrington, J. (2008) "Detecting Cartels," in Buccirossi, P., (ed.), *Handbook of Antitrust Economics,* MIT Press

[25] Harrington, J. and Skrzypacz, A. (2011) "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices," *American Economic Review*, 101(6), 2425-2449

[26] Hay, G. and Kelley, D. (1974) "An Empirical Survey of Price Fixing Conspiracies" *Journal of Law and Economics*, 17(1), 13-38

[27] Hirata, D., 2009, "Asymmetric Bertrand-Edgeworth Oligopoly and Mergers," *The B.E. Journal of Theoretical Economics*, 9(1), 1-25

[28] Huck, S., Normann, H.-T. and Oechssler, J. (2004) "Two are Few and Four are Many - on Number Effects in Cournot Oligopoly," *Journal of Economic Behaviour and Organization*, 53, 435-466

[29] International Competition Network (2008) "Setting of Fines for Cartels in ICN Jurisdictions," Report to the 7th ICN Annual Conference, Kyoto

[30] Ivaldi, M., Jullien, B., Rey, P., Seabright, P., and Tirole, J. (2003) "The Economics of Tacit Collusion," Report for DG Comp, European Commission

[31] Y. Katsoulacos, Y., Motchenkova, E. and Ulph, D. (2015a) "Measuring the Effectiveness of Anti-cartel Interventions: A Conceptual Framework," University of St Andrews, School of Economics and Finance Discussion Paper No. 1602

[32] Y. Katsoulacos, Y., Motchenkova, E. and Ulph, D. (2015b) "Penalizing Cartels: The Case for Basing Penalties on Price Overcharge," *International Journal of Industrial Organization*, 42, 70-80

[33] Kühn, K.-U., (2012) "How Market Fragmentation Can Facilitate Collusion," *Journal of the European Economic Association,* 10(5), 1116-1140

[34] Lambson, E. (1994) "Some Results on Optimal Penal Codes in Asymmetric Bertrand Supergames," *Journal of Economic Theory*, 62(2), 444-468

[35] Levenstein, M. and Suslow, V. (2006) "What Determines Cartel Success?" *Journal of Economic Literature*, 44(1), 43-95

[36] Martin, S. (2006) "Competition Policy, Collusion and Tacit Collusion," *International Journal of Industrial Organization,* 24(6), 1299-1332

[37] Motta, M. (2004) *Competition Policy: Theory and Practice*, Cambridge University Press

[38] Mouraviev, I. (2013) "Explicit Collusion under Antitrust Enforcement," Center for Mathematical Economics Working Paper, no. 494

[39] Motta, M. and Polo, M. (2003) "Leniency Programs and Cartel Prosecution," *International Journal of Industrial Organization*, 21(3), 347-379

[40] Petit, L. (2012) "The Economic Detection Instrument of the Netherlands Competition Authority: The Competition Index," NMa Working Papers, No. 6

[41] Posner, R. (1970) "A Statistical Study of Antitrust Enforcement," *Journal of Law and Economics*, 13(2), 365–419

[42] Spagnolo, G. (2005) "Divide et Impera: Optimal Leniency Programs," CEPR. Discussion Paper No. 4840

[43] Spector, D. (2015) "Facilitating Collusion by Exchanging Non-Verifiable Sales Reports," PSE Working Papers no. 2015-07

[44] Stigler, G. (1964) "A Theory of Oligopoly," *Journal of Political Economy*, 72, 44–61

[45] Symeonidis, G. (2003) "In Which Industries is Collusion More Likely? Evidence from the UK," *Journal of Industrial Economics*, 51(1), 45–74

[46] Tirole, J. (1988) *The Theory of Industrial Organisation*, MIT Press

[47] Vasconcelos, H. (2005) "Tacit Collusion, Cost Asymmetries, and Mergers," *RAND Journal of Economics*, 36(1), 39-62

Appendix

Proof of Proposition 1. Rearranging (5) yields: $\delta \geq \frac{1-S(p^m)+\theta F}{1-\underline{p}} \equiv \delta_e^*(k_n, F)$. Thus, if $\delta \geq \delta_e^*(k_n, F)$, then $k_i V_e^*$ is as claimed from (3). Finally, if $F \in \left[0, \overline{F}(k_n)\right)$, then $\delta_e^*(k_n, F) \in \left[\frac{k_n}{K}, 1\right)$ and $k_i V_e^* \in \left(\pi_i^N(k_n), \pi_i(p^m)\right]$.

Proof of Proposition 2. Note that the left-hand side of (9) is strictly less than the expression in square brackets on the right-hand side for all $\delta < 1$. Consequently, (9) can only hold if both are nonnegative, since $\delta^T \in (0,1]$ for all $T \in [0,\infty)$. This implies that a PPE in tacit monitoring strategies with $k_i V^c > \pi_i^N(k_n)$ must satisfy three conditions. First, the right-hand side of (9) must be non-negative, such that $G(m^*(k_1, \overline{m})) \leq \frac{K_{-n}}{K}$. Second, the left-hand side of (9) must be non-negative, where for any $G(m^*(k_1, \overline{m})) \in \left[0, \frac{K_{-n}}{K}\right)$, $\delta \geq \frac{1}{(1-G(m^*(k_1,\overline{m})))}\frac{k_n}{K} \equiv \delta_c^*(k_1, k_n) \in \left[\frac{k_n}{K}, 1\right)$. Third, $T \geq T^*(k_1, k_n)$ where $T^*(k_1, k_n)$ is implicitly defined by the level of T that ensures (9) holds with equality, so $T^*(k_1, k_n) \to \infty$ if $\delta = \delta_c^*(k_1, k_n)$ and $T^*(k_1, k_n) < \infty \ \forall$ $\delta > \delta_c^*(k_1, k_n)$.

Finally, if $\overline{m} < \underline{x}(k_1)$, such that $G(m^*(k_1, \overline{m})) = 0$, and if $\delta \geq \frac{k_n}{K}$, then $k_i V_c^* = \pi_i(p^m)$ for all i from (7). In contrast, if $\underline{x}(k_1) \leq \overline{m} < \overline{x}(k_1, k_n)$ such that $G(m^*(k_1, \overline{m})) \in \left(0, \frac{K_{-n}}{K}\right)$ and if $\delta \geq \delta_c^*(k_1, k_n)$, then $k_i V_c^*$ can be found by evaluating $k_i V^c$ at $T^*(k_1, k_n)$ such that $k_i V_c^* = \frac{k_i}{K}\left(\frac{\widehat{m} - G(m^*(k_1,\overline{m}))K}{1-G(m^*(k_1,\overline{m}))}\right) \in \left(\pi_i^N(k_n), \pi_i(p^m)\right)$, for all i. Otherwise, for any $\overline{m} \geq \overline{x}(k_1, k_n)$, such that $G(m^*(k_1, \overline{m})) \geq \frac{K_{-n}}{K}$, collusion under tacit monitoring is not sustainable, because (9) cannot be satisfied for any $\delta < 1$. ■

Proof of Proposition 3. Differentiating $\delta_e^*(k_n, F) = \frac{1-S(p^m)+\theta F}{1-\underline{p}}$ with respect to k_j yields:

$$\frac{\partial \delta_e^*}{\partial k_j} = \frac{1 - S(p^m) + \theta F}{\left(1-\underline{p}\right)^2}\frac{\partial \underline{p}}{\partial k_j}.$$

Thus, $\frac{\partial \delta_e^*}{\partial k_n} > 0$ from $\frac{\partial \underline{p}}{\partial k_n} = \frac{1}{k_n^2}\int_{\underline{m}}^{\max\{\overline{m}, K\}}(K-m)\, g(m)\, dm > 0$ and $S(p^m) \in (0,1)$ for all $\underline{m} < K$ from (2).

For any $\underline{x}(k_1) \leq \overline{m} < \overline{x}(k_1, k_n)$, differentiating V_c^* and $\delta_c^*(k_1, k_n)$ with respect to k_j yields:

$$\frac{\partial V_c^*}{\partial k_j} = -\frac{(K - \widehat{m})\, g(m^*)}{K\left(1 - G(m^*)\right)^2}\frac{\partial m^*}{\partial k_1}\frac{\partial k_1}{\partial k_j}$$

and

$$\frac{\partial \delta_c^*}{\partial k_j} = \frac{1}{K\left[1 - G(m^*)\right]}\left[\frac{\partial k_n}{\partial k_j} + k_n \frac{g(m^*)}{1 - G(m^*)}\frac{\partial m^*}{\partial k_1}\frac{\partial k_1}{\partial k_j}\right],$$

respectively. Thus, $\frac{\partial V_c^*}{\partial k_1} > 0$ and $\frac{\partial \delta_c^*}{\partial k_1} < 0$ from $\frac{\partial m^*}{\partial k_1} < 0$, $\widehat{m} < \overline{m} < K$, $\frac{\partial k_1}{\partial k_1} = 1$, and $\frac{\partial k_n}{\partial k_1} \in [-1, 0]$. Furthermore, $\frac{\partial \delta_c^*}{\partial k_n} > 0$ from $\frac{\partial k_n}{\partial k_n} = 1$, $\frac{\partial m^*}{\partial k_1} < 0$ and $\frac{\partial k_1}{\partial k_n} \in [-1, 0]$. Finally, for any $\overline{m} \leq \underline{x}(k_1)$, $\delta_c^*(k_1, k_n) = \frac{k_n}{K}$ implies $\frac{\partial \delta_c^*}{\partial k_n} > 0$. ■

Proof of Proposition 4. Note that $k_i V_e^* > k_i V_c^*$ if and only if $F < \frac{1}{\theta}\left(\frac{\widehat{m}}{K} - V_c^*\right) \equiv F^*(k_1)$, where $V_c^* = \frac{1}{K}\left(\frac{\widehat{m} - G(m^*)K}{1 - G(m^*)}\right)$ from Proposition 2. For any $\overline{m} < \underline{x}(k_1)$, such that $G(m^*) = 0$, then $F^* = 0$. It also follows from $\lim_{\overline{m} \to \overline{x}(k_1, k_n)} V_c^* = \underline{p}$ that $\lim_{\overline{m} \to \overline{x}(k_1, k_n)} F^*(k_1) = \overline{F}(k_n)$. Finally, if $\underline{x}(k_1) \leq \overline{m} < \overline{x}(k_1, k_n)$, then $F^* \in \left(0, \overline{F}(k_n)\right)$ from $V_c^* \in \left(\underline{p}, \frac{\widehat{m}}{K}\right)$. ■

Proof of Proposition 5. For any $\overline{m} < K$, $\delta_e^* < \delta_c^*$ if and only if:

$$\frac{1 - \frac{\widehat{m}}{K} + \theta F}{1 - \underline{p}} < \frac{1}{1 - G(m^*(k_1, \overline{m}))} \frac{k_n}{K}.$$

Rearranging the above yields:

$$\begin{aligned} F &< \tfrac{1}{\theta}\left[\tfrac{1-\underline{p}}{1-G(m^*(k_1,\overline{m}))}\tfrac{k_n}{K} - \left(1 - \tfrac{\widehat{m}}{K}\right)\right] \\ &< \tfrac{1}{\theta}\left(\tfrac{\widehat{m}}{K} - \tfrac{1}{K}\left(\tfrac{\widehat{m} - G(m^*)K}{1 - G(m^*)}\right)\right), \end{aligned}$$

where the right-hand side can be rewritten as $\frac{1}{\theta}\left(\frac{\widehat{m}}{K} - V_c^*\right) = F^*(k_1)$. ■