
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.

Collusion, firm numbers and asymmetries revisited [Bergen]

PLEASE CITE THE PUBLISHED VERSION

<http://beccl.no/2016-beccl-competition-policy-conference/>

PUBLISHER

© The Authors

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at:
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Garrod, Luke, and Matthew Olczak. 2019. "Collusion, Firm Numbers and Asymmetries Revisited [bergen]".
figshare. <https://hdl.handle.net/2134/21259>.

Collusion, Firm Numbers and Asymmetries Revisited*

Luke Garrod[†] and Matthew Olczak[‡]

January 29, 2016

Abstract

In an infinitely repeated game where market demand is uncertain and where firms with (possibly asymmetric) capacity constraints must monitor the agreement through their privately observed sales and prices, we analyse the firms' incentives to form a cartel when they could alternatively collude tacitly. In this private monitoring setting, tacit collusion involves price wars on the equilibrium path if a firm cannot infer from its low sales whether the realisation of market demand was unluckily low or whether at least one rival has undercut the collusive price. In contrast, explicit collusion involves firms secretly forming an illegal cartel to share their private information to avoid such price wars, but this runs the risk of sanctions. We show, in contrast to the conventional wisdom and consistent with the empirical evidence, that the incentives to form an illegal cartel can be smallest in markets with a few symmetric firms, because tacit collusion is most successful in such markets.

JEL classification: D43, D82, K21, L12, L41

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

*We have benefitted from the comments of seminar participants at the ESRC Centre for Competition Policy. The usual disclaimer applies.

[†]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

Economic theory has played a key role in developing an understanding of the circumstances in which cartels are most likely to form and of what policymakers can do to deter and detect them. A robust finding is that firms have the greatest incentive to form a cartel in markets where there are a few firms with relatively limited asymmetries among them. This result is robust to a large number of possible forms of heterogeneity. For example, collusion is hindered by asymmetries whether they are in terms of firms' capacity constraints (see Compte *et al.*, 2002; Bos and Harrington, 2010 and 2014; Garrod and Olczak, 2016), the number of differentiated products that each firm sells (see Kühn, 2004), the cost structures of firms (Vasconcelos, 2005; and Miklós-Thal, 2011), or the rate at which firms discount the future (see Harrington, 1989). The reason underlying this finding in all of these different settings is that it is difficult to incentivise some members from undercutting the cartel price when there are a large number of asymmetric firms.¹

Nevertheless, this theoretical prediction is in stark contrast to the empirical evidence of detected cartels that shows that many cartels tend to include a large number of generally very asymmetric firms. For example, Davies and Olczak (2008) show that there were, on average, 6 firms involved in detected cartels in Europe between 1990 and 2006, where the extent of the asymmetries between the market shares of the firms “calls into question whether symmetry of market shares is a pervasive feature of real world cartels” (p.198). For the US, Hay & Kelley (1974) found that the 49 detected cartels prosecuted by the US Department of Justice between 1963 and 1972 had, on average, 7 members with a four-firm concentration ratio of 77%. Similarly, Levenstein and Suslow (2006) review of a number of studies analysing US cartels through out the 20th century and find no clear relationship between concentration and the prevalence of cartels. This evidence contrasts with the theoretical predictions but it is also at odds with the evidence on tacit collusion.² Levenstein and Suslow (2006, p.58) suggest three possible explanations, first, because the evidence is based on prosecuted cartels it may be a biased sample. Second, since concentration is endogenously determined, collusion may have allowed an increased number of firms to survive in these markets. Finally, consistent with the conjecture in Harrington

¹A second reason why collusion is more likely in markets with a small number of symmetric firms is due to the fact that reaching an agreement is easier when firms are similar and there are fewer of them.

²Davies et al. (2011) show that the European Commission's interventions on tacit collusion grounds have almost always been confined to cases where there would have been only two very symmetric players post-merger. It is also supported by experimental evidence in Fonseca and Normann (2012) where the gains from communication are much larger with 4 rather than 2 players.

(2010, p39-40), firms tacitly colluding in markets with few symmetric firms rather than colluding explicitly by forming an illegal cartel.

In this paper, we explore this final, previously under-researched, explanation further. We achieve this by extending Garrod and Olczak (2016), where we modelled tacit collusion in a setting where market demand is uncertain and where firms with (possibly asymmetric) capacity constraints never directly observe their rivals' prices or sales. In this private monitoring setting, similar to that first discussed by Stigler (1964) and recently analysed by Harrington and Skrzypacz (2007), the firms monitor a collusive agreement through their privately observed sales, and when they receive low sales they may be unsure whether it is due to low demand or due to a rival undercutting the collusive price. Consequently, to solve this non-trivial signal extraction problem, the firms initiate price wars on the equilibrium path, similar to other models of imperfect monitoring (see Green and Porter, 1984; Tirole, 1988). We extend this model of tacit collusion to allow the firms to instead form an illegal cartel that allows them to share their private information with each other to avoid the non-trivial signal extraction problem. Thus, firms face a tradeoff between these two forms of collusion: on the one hand, a cartel is likely to generate higher profits than tacit collusion, because the cartel can avoid costly price wars; but, on the other hand, forming a cartel is illegal and runs the risk of sanctions.³ We use this model to investigate under which market structures firms have incentives to collude explicitly by forming a cartel when they also have the ability to collude tacitly.

Our paper is related to the literature that analyses explicit collusion in the presence of a competition agency, where collusion requires explicit communication and where cartel members face the threat of sanctions. This literature has focussed mainly on the effects of leniency programmes in markets with symmetric firms (see Motta and Polo, 2003, and Chen and Rey, 2013).⁴

However, in a Bertrand-Edgeworth setting similar to ours, Bos and Harrington (2014) also consider the effect of sanctions on the number and type of firms that join the cartel when firms

³Anticompetitive agreements, such as collusion, are commonly prohibited by competition laws and hard evidence of a cartel (e.g., recorded conversations, minutes of meeting, or emails) is usually required to prove guilt of collusion. Consequently, tacit collusion is not usually considered illegal, despite potentially causing similar effects to explicit collusion. This approach to the law is arguably desirable because it ensures legal certainty over illegal conduct, and it prevents competitive behaviour from being punished erroneously, which could undermine the market mechanism across the economy (see Motta, 2004, p.185–190).

⁴One other exception is Martin (2006) who considers tacit and explicit collusion in a Cournot framework. However, our model allows for a much broader range of comparative statics on the chosen form of collusion, including allowing for asymmetries between firms.

are asymmetrically capacity constrained. All of these papers model collusion under perfect observability yet assume that tacit collusion is not possible.⁵ An implication of this assumption is that the benefit of communication compared to tacit collusion is exogenous. In contrast, we model collusion in a private monitoring setting, where informational problems can mean that tacitly colluding firms are unable to extract the full monopoly profit. Thus, in our model firms can exchange their private information on sales and prices to increase the collusive profits, so that the benefits of communication are endogenous. Furthermore, this exchange of private information captures one of the main functions of a large number of cartels (see, for example, Harrington's (2006b) discussion of the following cartels: carbonless paper, choline chloride, copper plumbing tubes, graphite electrodes, plasterboard, vitamins, and zinc phosphate).

In relation to explicit collusion, we show first that firms do not need to exchange their sales information when market demand is sufficiently high. The reason is that, even in the absence of communication, all firms can always infer when at least one firm's sales are below some firm-specific "trigger level", which is determined by the largest possible sales consistent with them or a rival being undercut on price. Thus, if all firms set a common price, then all firms' sales will exceed their respective trigger levels when the realisation of market demand is high, otherwise they can all fall below the trigger levels. Yet, if all firms do not set a common price, then at least one firm will receive sales below their trigger level. As a consequence, firms need only exchange their sales information to check there has not been a deviation when at least one firm's sales are below their trigger level and this happens in equilibrium when market demand is low. This is consistent with the evidence from the copper plumbing tubes cartel where, as described by Harrington (2006b, p.54): "The European Commission found that from late 1994 onwards, probably due to the increased demand coming from the German construction boom, there were fewer contacts between cartel members and apparently the cartel did not meet at all in 1995. With the weakening of the German economy in 1996, meetings were re-established".

If, consistent with the previous literature, we do not take account of tacit collusion, we need to compare the best equilibrium payoff under explicit collusion with the static Nash equilibrium profits. The static Nash equilibrium profits are increasing in the size of the largest firm, because competition is weaker whereas the per-period profits of explicit collusion are unaffected since if

⁵If tacit collusion were possible, then any sanction would be sufficient to prevent explicit collusion. It is often argued that communication is required in these models to coordinate on price, so these models are most appropriate for price-fixing cartels. However, even this is a bit of a stretch, because this difficulty is not modelled as usually the collusive price is the same in each period.

collusion is sustainable prices are at the monopoly level. Thus, the difference between the static Nash equilibrium profit and the best equilibrium payoff under explicit collusion decreases as the largest firm gets larger. Consequently, if we do not take account of tacit collusion, then we would also conclude that firms have the greatest incentive to form a cartel when they are symmetric.

When tacit collusion is taken into account, however, we show that firms' incentives to form a cartel are very different. For instance, we show that if monitoring is perfect under tacit collusion, then there is no incentive to form a cartel, because firms can extract the monopoly profit tacitly without the need to form a cartel. This happens when fluctuations in market demand are small, because firms will only ever receive sales below their trigger levels if they are undercut. Furthermore, if fluctuations in market demand are large, such that there is imperfect monitoring because collusive sales can also fall below the trigger levels, then firms have the least incentive to form a cartel when they are symmetric. This is due to the fact that the best equilibrium payoff under tacit collusion decreases as the smaller firm gets smaller. The reason is that punishment phases occur more often on the equilibrium path when the smallest firm has less capacity, since monitoring is more difficult. Thus firms have a greater incentive to form a cartel when they are sufficiently asymmetric. Finally, given both equilibrium payoffs are independent of the largest firm, it actually follows that firms have the smallest incentive to form a cartel in a symmetric duopoly. This is in stark contrast to the conventional wisdom that suggests that cartels have the greatest incentive to form in such a case.

More broadly, our paper is also related to a literature that is interested in explaining how sharing non-verifiable sales reports can help firms to collude. For instance, Compte (1998) and Kandori and Matsushima (1998), characterise equilibria under private monitoring in which firms truthfully report and this enables them to achieve joint profit maximisation. Harrington and Skrzypacz (2011) model a collusive agreement in which firms truthfully reveal non-verifiable sales information to each other and make inter-firm transfers based on this. In a related paper, Aways and Krishna (2014) show that when firms' sales at similar prices are sufficiently correlated such communication can facilitate collusion without the need for transfers between firms. Furthermore, Spector (2015) shows that such information exchange can facilitate collusion even if sales data will eventually be publicly verified. This is because communication allows firms to immediately put in place temporary market share reallocations and removes the need for price wars.⁶ While

⁶This literature focuses on how communication helps to monitor compliance. In a related literature (Athey and Bagwell, 2001 and 2008) sharing of cost information increases cartel profits by ensuring that the firm with the low costs produces more of the industry output.

we assume in the main body of paper that sales are verifiable, in our ongoing work we intend to demonstrate that our analysis still holds if sales are non-verifiable.

The rest of the paper is organised as follows. Section 2 sets out the assumptions of the model and solves for the static Nash equilibrium. In section 3, we analyse the two forms of collusion. We first show that under explicit collusion, if firms are sufficiently patient, then the best equilibrium payoffs have the firms set the monopoly price in any cartel phase and risk the chance of a fine. Then we present a simplified version of the model of tacit collusion in Garrod and Olczak (2016) that generates identical results to our previous analysis. In section 4, we analyse the incentives of firms to collude explicitly or tacitly. Section 5 then makes some concluding remarks on the policy implications of our results. All proofs are relegated to appendix A.

2 The Model

To capture the difference between explicit and tacit collusion, the capacity-constrained price-setting private monitoring repeated game in Garrod and Olczak (2016) is extended to allow firms to form a cartel to exchange information but this involves a risk of being detected and penalised by a competition agency.

2.1 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. Firms' costs are normalised to zero but the maximum that firm i can produce in any period is k_i , where we let $k_n \geq k_{n-1} \geq \dots \geq k_1 > 0$, without loss of generality. We denote total capacity as $K \equiv \sum_i k_i$ and the maximum that firm i 's rivals can supply in each period as $K_{-i} \equiv \sum_{j \neq i} k_j$. In any period t , firms set prices simultaneously where $\mathbf{p}_t = \{p_{it}, \mathbf{p}_{-it}\}$ is the vector of prices set in period t , p_{it} is the price of firm $i = \{1, \dots, n\}$ and \mathbf{p}_{-it} is the vector of prices of all of firm i 's rivals. Firms have a common discount factor, $\delta \in (0, 1)$.

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 $\tau \in \{0, \dots, t\}$, but they know that m_t is independently drawn from a distribution $G(m)$, with mean \hat{m} and density $g(m) > 0$ on the interval $[\underline{m}, \bar{m}]$. Furthermore, firm i never observes firm j 's prices, $p_{j\tau}$, or sales, $s_{j\tau}$, $j \neq i$, for all $\tau \in \{0, \dots, t\}$. 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 in which 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.

2.2 Demand allocation and sales

Demand is allocated according to the following rule:

The proportional allocation rule

- Demand is allocated to the firm with the lowest price first. If this firm's capacity is exhausted, then demand is allocated to the firm with the second lowest price, and so on.
- If two or more firms set the same price and if their joint capacity suffices to supply the (residual) demand, then such firms each receive demand equal to its proportion of the joint capacity.

This allocation rule is an appropriate assumption for a model of explicit collusion because it is commonly considered in the literature in terms of a cartel selecting how much of the market demand each member supplies. Indeed, there are a number of cartels that have allocated demand in proportion to each member's capacity (see the examples in Vasconcelos, 2005, and Bos and Harrington, 2010, who also use this allocation rule). Furthermore, it is also appropriate for our model of tacit collusion because, as we argue in Garrod and Olczak (2016), it is consistent with buyers' optimal search strategies when they must undertake costly search to discover whether firms have spare capacity (see Arnold, 2000; and Arnold and Saliba, 2011).

Following Garrod and Olczak (2015), we also place the following plausible yet potentially restrictive assumption on the capacity distribution:

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 is a necessary condition that ensures 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. Given the smallest firm's capacity can be no larger than for a symmetric duopoly, a necessary condition for Assumption 1 to hold is that the minimum market demand must be greater than 50% of the total capacity, $\underline{m} \geq 0.5K$. We believe that Assumption 1 is not very restrictive in our context, because tacit collusion is most likely to occur in markets with

two or three relatively symmetric firms (see Davies *et al.*, 2011), so the smallest firm is likely to be relatively large. We place no restriction on the level of the maximum market demand, \bar{m} .

Thus, 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:

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

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 $\pi_{it}(p_{it}, \mathbf{p}_{-it}) = p_{it} \int_{\underline{m}}^{\bar{m}} s_{it}(p_{it}, \mathbf{p}_{-it}; m) g(m) dm$, where we drop time subscripts if there is no ambiguity. We write $\pi_i(p) = k_i p S(p)$ if $p_j = p$ for all j , where $S(p)$ is the expected sales per unit of capacity such that:

$$S(p) = \begin{cases} 1 & \text{if } K \leq \underline{m} \\ \int_{\underline{m}}^K \frac{m}{K} g(m) dm + \int_K^{\bar{m}} g(m) dm & \text{if } \underline{m} < K < \bar{m} \\ \frac{\hat{m}}{K} & \text{if } \bar{m} \leq K. \end{cases} \quad (2)$$

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

2.3 Static Nash equilibrium

In this subsection, we present the static Nash equilibrium of the game, which can be in mixed strategies. An important part of the analysis is firm i 's minimax payoff, which is:

$$\underline{\pi}_i \equiv \begin{cases} \hat{m} - K_{-i} & \text{if } \bar{m} \leq K, \\ \int_{\underline{m}}^K (m - K_{-i}) g(m) dm + k_i \int_K^{\bar{m}} g(m) dm & \text{if } \underline{m} < K < \bar{m} \\ k_i & \text{if } K \leq \underline{m} < \bar{m} \end{cases} \quad (3)$$

for all i . The intuition is that if the realisation of market demand is below total capacity, then a firm that sets the monopoly price expects to supply the residual demand, otherwise it expects to supply its full capacity.

Lemma 1. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m}$:*

- i) if $\underline{m} \geq K$, then there exists a unique pure strategy Nash equilibrium, with profits $\pi_i^N = k_i \forall i$;*
- ii) if $\underline{m} < K$, then there exists a mixed strategy Nash equilibrium, with profits, $\pi_i^N(k_n) = k_i \frac{\pi_n}{k_n} \forall i$.*

The proof of Lemma 1 is the same as in Garrod and Olczak (2016), so we only sketch it here. Competition is not effective if market demand is always greater than total capacity, because firms will always supply their full capacity, k_i , for every level of demand and for any prices. Otherwise, firms have incentives to undercut each other. However, by charging $p_i = 1$, firm i can ensure that its expected per-period profit is at least $\underline{\pi}_i$. Assumption 1 is sufficient to ensure that such profits are nonnegative for all i , so competition does not imply price equals marginal cost. Instead, the largest firm will never set a price below $\underline{p} \equiv \underline{\pi}_n/k_n$ in an attempt to be the lowest-priced firm.⁷ This implies that the smaller firms $i < n$ can sell their full capacity with certainty by charging a price slightly below \underline{p} to obtain a profit of $k_i \underline{p} > \underline{\pi}_i$. Consequently, the mixed strategy Nash equilibrium profits are given by $\pi_i^N(k_n) = k_i \underline{p}$. The lower bound of the support is \underline{p} , where it follows from (3) that $\lim_{\underline{m} \rightarrow K} \underline{p} = 1$.

2.4 The repeated game

We now move on to discuss the repeated game. Given firm i 's past prices and sales are private information, it may have an incentive to share this information with its rivals to facilitate collusion. In this subsection, we first consider the public information that is available to all firms in the absence of communication. Then we set out the assumptions regarding how firms can exchange their private information. Henceforth, we impose $\underline{m} < K$, as collusion is unnecessary otherwise from Lemma 1.

2.4.1 Information and monitoring

As we showed in Garrod and Olczak (2016), firm i not only has a private history of its past prices and sales, denoted $z_i^t \equiv (p_{i0}, s_{i0}; \dots; p_{it-1}, s_{it-1})$, but there is also a public history that is common knowledge, for any z_i . The reason is that all firms can always infer from their own sales when at least one firm's sales are below some firm-specific "trigger level". Thus, there also exists public information in which firms can condition their play on in the absence of communication. Since this is central to our story, we briefly restate these arguments here.

Let firm i 's trigger level be $s_i^* \equiv \min \left\{ \frac{k_i}{K} m^*(k_1, \bar{m}), k_i \right\}$ for all i , where $m^*(k_1, \bar{m}) \equiv \frac{K(\bar{m}-k_1)}{K-1}$. Furthermore, let the history $h^t = (y_0, y_1, \dots, y_{t-1})$ be a sequence of information

⁷The smaller firms $i < n$ all placing positive probability on charging \underline{p} is necessary and sufficient to ensure that \underline{p} is also the lowest price firm n will charge.

in period t where, for all $\tau = \{0, 1, \dots, t-1\}$:

$$y_\tau = \begin{cases} \bar{y} & \text{if } s_{i\tau}(p_{i\tau}, \mathbf{p}_{-i\tau}; m_\tau) > s_i^* \quad \forall i \\ \underline{y} & \text{otherwise.} \end{cases}$$

This says that $y_\tau = \bar{y}$ if all firms' sales in period τ exceed their trigger levels, but $y_\tau = \underline{y}$ if at least one firm's sales does not. It follows that h^t is a public history, because y_τ is common knowledge for all τ , for any z_i^t . To see this, first consider the case where maximum market demand is above the total capacity, $\bar{m} \geq K$. In this case, the trigger levels are so high that all firms' sales can never exceed them for any prices, that is, $s_i^* = k_i$ so $y_\tau = \underline{y}$ for all τ . Next, consider the case of $\bar{m} < K$, where it is possible for firms to receive sales above their trigger levels, since $s_i^* < k_i$. In this case, if all firms set a common price $p \leq 1$, then the sales of all firms will exceed their respective trigger levels if the realisation of market demand is high, otherwise they can all fall below the trigger levels. Yet, if all firms do not set such a common price, then the sales of the firm(s) that set the highest price will not exceed their trigger levels. This follows from the fact that, for any nonempty set of rivals with a price below p^{\max} , $\Omega(p^{\max})$, the sales of firm i with $p_i = p^{\max} \leq 1$ are:

$$s_i = \frac{k_i}{K - \sum_{j \in \Omega(p^{\max})} k_j} \left(m_t - \sum_{j \in \Omega(p^{\max})} k_j \right) \leq \frac{k_i(\bar{m} - k_1)}{K - 1} = s_i^* < k_i, \quad (4)$$

from (1). Furthermore, all of their lower priced rivals will supply their full capacities. Any firm that supplies its full capacity can infer from this that at least one firm's sales are below its trigger level. The reason is that each firm knows, from (1), that it will supply its full capacity only if its price is below the highest in the market.⁸

The above implies that each firm knows that all firms' sales will exceed their trigger levels, such that $y = \bar{y}$, only if $p_j = p \leq 1$ for all j and if $m > m^*(k_1, \bar{m})$; otherwise, at least one firm's sales will not exceed its trigger level, so $y = \underline{y}$. Thus, if $\underline{m} > m^*(k_1, \bar{m})$, then there is perfect monitoring of a strategy in which all firms set a common collusive price, even in the absence of communication, because each firm would only receive sales below its trigger level, if it has been undercut. In contrast, there is imperfect monitoring of such an agreement in the absence of communication, if $\underline{m} \leq m^*(k_1, \bar{m})$. The reason can be understood by considering $\Pr(\underline{y}|p_i, \mathbf{p}_{-i})$, which denotes the probability of observing \underline{y} if firm i sets p_i and its rivals price according to \mathbf{p}_{-i} .

⁸Likewise, if any firms' prices are above 1, then they will receive zero sales, which is below their trigger levels. In this case, only the firms whose prices do not exceed 1 will supply their full capacities.

For the case of $\underline{m} \leq m^*(k_1, \bar{m})$:

$$\Pr(y|p_i, \mathbf{p}_{-i}) = \begin{cases} \int_{\underline{m}}^{\min\{m^*(k_1, \bar{m}), \bar{m}\}} g(m) dm \in [0, 1] & \text{if } p_j = p \forall j \\ 1 & \text{otherwise.} \end{cases} \quad (5)$$

This says that a firm's sales can be below their trigger level if the realisation of market demand is sufficiently low, even when firms set a common price. Thus, without more information, colluding firms face a non-trivial signal extraction problem: each firm does not know whether the realisation of market demand was unluckily low or whether at least one rival has undercut them. Finally, note that $\Pr(y|p_i, \mathbf{p}_{-i}) = 1$ for all \mathbf{p} , such that firms always receive sales below their trigger levels, if $\bar{m} \geq K$.

Lemma 2 states the conditions for perfect and imperfect monitoring in terms of the maximum market demand, holding the minimum market demand constant.

Lemma 2. *For any given $n \geq 2$, $K_{-1} \leq \underline{m} < K$, and $\delta \in (0, 1)$, there exists a unique level of market demand, $\underline{x}(k_1) \in (\underline{m}, K)$, such that if $\bar{m} \in (\underline{m}, \underline{x}(k_1))$, then monitoring is perfect. Otherwise, there is imperfect monitoring.*

Given deviations by the smallest firm are most difficult to detect, it follows that detecting deviations is less difficult when the smallest firm is larger, so the critical level $\underline{x}(k_1)$ is strictly increasing in the capacity of the smallest firm, k_1 .

2.4.2 Timing of the repeated game

We now set out the assumptions regarding the exchange of information. The term cartel will be used to refer to a group of firms that conspire to collude by sharing 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. Thus, the difference between explicit and tacit collusion is that there is an active cartel in period 0 under explicit collusion, but there is never an active cartel in any period under tacit collusion. Consistent with competition law in Europe and the US, explicit collusion is subjected to enforcement but tacit collusion is not, because no communication takes place. If detected, then each of the cartel members i are convicted with probability 1 and fined $k_i F$, where $F \geq 0$ is the fine per unit of the industry's capacity.⁹We allow for the possibility that each firm will inform

⁹This implies that larger firms receive larger fines, which is consistent with most jurisdictions, including Europe and the US, where the fine for each cartel member is (initially) linked to the size of its sales (see ICN, 2008). It is also consistent with Bos and Harrington (2014).

the competition authority of the cartel in return for leniency, in which case the cartel is detected with probability 1 and the informant receives full leniency. Consistent with Europe and the US, we assume that applying for leniency is publically observable. In addition, to leniency, which in practice is the most common way in which cartels are detected (Stephan and Nikpay, 2015), we assume that there is a competition authority that each period can detect active cartels with some probability. This probability is assumed to be independent of both the number of periods the cartel has been active and whether the cartel has needed to exchange its private information that period. The latter assumption would certainly be true for cartels that also need to discuss the fixing of prices each period. It is also consistent with other common ways, in addition to leniency programmes, through which cartels can be detected, including complaints from buyers and behavioural screens (see Harrington, 2008a).

The timing of the game in a given period t is as follows:

- **Stage 0 (pricing stage)** - firms set prices simultaneously. The game continues to stage 1.
- **Stage 1 (communication stage)** - firms privately realise their sales and profits:
 - If there is not an active cartel (i.e. either firms are colluding tacitly or they are in a punishment phase under explicit collusion), then period t ends and period $t+1$ begins.
 - If the cartel is active, then firms choose whether to share their private information among themselves secretly, and whether to inform a competition agency of the cartel publically in return for leniency. The game continues to stage 2.
- **Stage 2 (detection stage)** - the cartel's punishment is realised:
 - If no firm has informed the competition agency of the cartel, then the competition agency detects and convicts the cartel with a probability $\theta \in (0, 1]$, and all firms are each fined $k_i F$ in such an event. Otherwise, the cartel is not detected and no firm is fined.
 - If at least one firm has informed the competition agency 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.

It should be clear from the timing of the game that, consistent with Motta and Polo (2003), Spagnolo (2005) and Chen and Rey (2013), we assume that the cartel only becomes inactive if a firm deviates from the collusive agreement in the main analysis. Thus, it does not become inactive if it is detected and convicted by the competition agency. This seems a natural assumption to us, because there is widespread evidence of recidivism among detected cartels (see Connor, 2010). An alternative approach followed by Aubert *et al.* (2006), Harrington (2008b), and Bos and Harrington (2014) is to assume that the cartel does become inactive after detection, in which case the firms play the static Nash equilibrium in every period thereafter. We have returned to this in our ongoing work where we allow the firms to tacitly collude after detection, and can show that our main results are robust.

3 Cartels and Asymmetries

In this section, we first model explicit collusion where there is an active cartel in period 0. In this setting, firms can overcome a non-trivial signal extraction problem by forming a cartel to share their private information. We then show that our results are consistent with the previous literature when we do not take account of tacit collusion.

3.1 Explicit collusion

In this subsection, we consider the following strategy profile, which we refer to as explicit trigger-sales strategies. There are ‘cartel’ phases and ‘punishment phases’. In a pricing stage of a period during an active phase, a firm sets the collusive price p^c . Then in the communication stage, the firm realises its sales. If $y_t = \bar{y}$, such that all firms’ sales are above their trigger levels, then each firm does not secretly share its private information, it does not apply for leniency, and the cartel phase continues into the next period. However, if $y_t = \underline{y}$, such that at least one firm’s sales are below its trigger level, then each firm secretly shares its private information and does not apply for leniency. The active phase continues only if $p_{jt} = p^c$ and $s_{jt} = k_i S(p^c)$ for all j . Otherwise, firms enter an inactive phase. Once in the inactive phase, each firm prices according to the static Nash equilibrium forever and firms never share their private information.

There are three related comments to make regarding this strategy profile. First, reversion to the static Nash equilibrium is the harshest possible punishment under our assumptions because,

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 (see Lemma 1), so it is not possible to implement a harsher punishment. Second, note that it follows from the analysis of section 2.4 that if the realisation of market demand is so high that no firm receives sales below its trigger level, $y_t = \bar{y}$, there is no benefit to firms from exchanging their private information. Third, the fact that the cartel is still active after detection and conviction implies that, under certain conditions, the cartel will want to exploit the leniency programme by applying for leniency in every period in an attempt to reduce its expected fines. To our knowledge, there is no evidence of such game playing in the real world, so we focus on the above strategy profile in the main paper. The equilibrium is then optimal if following this strategy is more profitable than exploiting the leniency system.

One reason why such game playing is unlikely is that it will only happen outside of the empirically important parameter space.¹⁰ In ongoing work we are investigating the alternative game playing strategy. An important determinant for that analysis is which informant is selected for leniency when there is more than one. We've assumed that the deviant with the lowest price is given leniency, which is consistent with Spagnolo (2005), who assumes that a deviant informant will be given leniency over a colluding informant. However, this has no effect on the current analysis because firms do not apply for leniency on the equilibrium path, so there is only ever one deviating informant in this case. Consequently, the assumption that firms cannot apply for leniency before sales are realised is innocuous for our main analysis, because it is more profitable for a deviating firm to wait until sales are realised.

3.1.1 Collusive profits

We first consider the expected profit firms can make if each firm abides by its prescribed strategy. Given conviction does not lead to the breakdown of the cartel, firm i 's expected (normalised) discounted payoff in stage 1 of any period during a cartel phase is:

$$k_i V^e = (1 - \delta) (\pi_i(p^c) - \theta k_i F) + \delta k_i V^e, \quad (6)$$

for any y_t . Where solving yields:

$$k_i V^e = \pi_i(p^c) - \theta k_i F. \quad (7)$$

¹⁰Pre-empting that analysis we show that $\theta < \frac{1}{2}$ is a sufficient condition for this. Given the low detection rates of cartels, this seems likely to hold in most jurisdictions, hence the reason why this analysis is deferred until later.

This says that in stage 0 of any collusive period firm i expects to receive the expected per-period collusive profit minus its expected fine. Note that V^e is independent of k_i and δ and that $V^e < p^c$ for any $F \geq 0$.

3.1.2 Incentive compatibility constraints

We next consider whether firm i has an incentive to deviate from its prescribed strategy. Clearly, firms do not have an incentive to deviate in a punishment phase. So, below we consider each firm's incentives to deviate in a cartel phase. We begin in stage 1 of a given period t (because firms do not act in stage 2) and then we move back to stage 0. Recall that during a cartel phase, firm i 's rivals should set p^c , not apply for leniency, and share their private information only if $y_t = \underline{y}$.

Stage 1

Consider whether firm i has an incentive to deviate from its prescribed strategy in stage 1 of period t . First, suppose firm i has abided by its strategy in stage 0 of period t by setting $p_i = p^c$. This implies that the cartel phase will continue into period $t+1$ if firm i continues to abide by its strategy in stage 1. Otherwise, firms will enter a punishment phase. Firm i can deviate in stage 1 by applying for leniency for any y_t and/or by not sharing its private information if $y_t = \underline{y}$. If firm i applies for leniency in stage 1, its expected (normalised) discounted payoff is:

$$(1 - \delta) \pi_i(p^c) + \delta \pi_i^N(k_n),$$

for any y_t , and regardless of whether firm i shares its information if $y_t = \underline{y}$. Thus, applying for leniency is always firm i 's optimal deviation at this stage, because if it deviates by not sharing its private information in the event that $y_t = \underline{y}$ without applying for leniency, then firm i would receive the above profits minus θF .

It follows that a necessary and sufficient condition that ensures firm i will not deviate in stage 1, for any y_t , is that it will not apply for leniency if firms receive sales below their trigger levels, such that $y_t = \underline{y}$. The corresponding ICC is:

$$(1 - \delta) (\pi_i(p^c) - \theta k_i F) + \delta k_i V^e \geq (1 - \delta) \pi_i(p^c) + \delta \pi_i^N(k_n). \quad (8)$$

The left-hand side of (8) is firm i 's profit from abiding by its strategy, given in (6), and the right-hand side is the profit from firm i 's optimal deviation. We refer to this as the ‘‘communication’’

incentive compatibility constraint (ICC). Rearranging (8) yields:

$$\delta \geq 1 - \frac{V^e - \underline{p}}{V^e - \underline{p} + \theta F} \equiv \delta_e^m(k_1, k_n, p^c, F). \quad (9)$$

This critical discount factor is independent of k_i , such that if the communication ICC holds for firm i , then it holds for firm $j \neq i$. Furthermore, $\delta_e^m(k_1, k_n, p^c, F) < 1$ for any $V^e > \underline{p}$, where $\lim_{F \rightarrow 0} \delta_e^m(k_1, k_n, p^c, F) = 0$ and where $\delta_e^m(k_1, k_n, p^c, F)$ is strictly increasing in F (noting from (7) that V^e is a strictly decreasing function of F). This implies that there is a greater incentive for each firm to deviate from its prescribed strategy in stage 1, if the fine per unit of capacity is larger.

Now, suppose firm i had deviated from its prescribed strategy in stage 0 of period t by setting $p_i \neq p^c$. An implication of this is that $y_t = \underline{y}$, such that at least one firm will receive sales below its trigger levels with certainty, from (5), so firms should share their private information. However, the fact that $p_i \neq p^c$ implies that firms will enter a punishment phase, regardless of whether firm i 's shares its private information or not. Thus, in this case firm i has a dominant strategy to apply for leniency to raise its expected profits by reducing its expected fine (in which case it is indifferent between sharing its private information or not).

Stage 0

Next, consider firm i 's incentive to deviate from its prescribed strategy in stage 0. It follows from the above that if firm i deviates in stage 0, then in stage 1 it will apply for leniency and firms will enter a punishment phase. In contrast, if a firm does not deviate in stage 0, then it will also not deviate in stage 1 if the communication ICC is satisfied. Thus, assuming the communication ICC is satisfied, such that $V^e > \underline{p}$ and $\delta \geq \delta_e^m(k_1, k_n, p^c, F)$, it follows that firm i has no incentive to deviate in the stage 0 if:

$$k_i V^e \geq (1 - \delta) k_i p^c + \delta \pi_i^N(k_n). \quad (10)$$

The left-hand side of (10) is firm i 's profit from abiding by its strategy, given in (7), and the right-hand side is the profit from firm i 's optimal deviation. The first term on the right-hand side follows from $p^c > V^e > \underline{p}$ which implies firm i 's optimal deviation in stage 0 is to undercut p^c marginally to supply its full capacity, k_i , rather than supplying the residual demand at the monopoly price. We refer to this as the ‘‘pricing’’ ICC. Rearranging (10) yields:

$$\delta \geq 1 - \frac{V^e - \underline{p}}{p^c - \underline{p}} \equiv \delta_e^p(k_1, k_n, p^c, F), \quad (11)$$

This critical discount factor is independent of k_i , so if it holds for firm i it holds for firm $j \neq i$. Furthermore, it follows from $V^e < p^c$ that $\delta_e^p(k_1, k_n, p^c, F) < 1$ for any $V^e > \underline{p}$, where $\lim_{F \rightarrow 0} \delta_e^p(k_1, k_n, p^c, F) = 1 - \frac{p^c S(p^c) - \underline{p}}{p^c - \underline{p}} > 0$ and where $\delta_e^p(k_1, k_n, p^c, F)$ is strictly increasing in F (noting from (7) that V^e is a strictly decreasing function of F). This implies that there is a greater incentive for each firm to deviate from its prescribed strategy in stage 0, if the fine per unit of capacity is larger.

3.1.3 Equilibria

It follows from the above that any given $V^e \in (\underline{p}, p^c)$ is supportable as a subgame perfect Nash equilibrium (SPNE) if $\delta \geq \delta_e^*(k_1, k_n, p^c, F) \equiv \max\{\delta_e^m(k_1, k_n, p^c, F), \delta_e^p(k_1, k_n, p^c, F)\}$. Proposition 1 now finds the best SPNE payoffs.

Proposition 1. *For any given $n \geq 2$, $K_{-1} \leq \underline{m} < K$ and $\overline{m} > \underline{m}$, there exists a unique fine per unit of capacity, $\overline{F} > 0$, that solves $V^e = \underline{p}$, such that if $F \in [0, \overline{F}]$ and if $\delta \geq \delta_e^*(k_1, k_n, p^m, F) = \delta_e^p(k_1, k_n, p^c, F) \in (\frac{k_n}{K}, 1)$, then the best equilibrium payoff for firm i under explicit trigger-sales strategies is:*

$$k_i V_e^* = \pi_i(p^m) - \theta k_i F \in (\pi_i^N(k_n), \pi_i(p^m)], \forall i,$$

Otherwise, collusion under explicit trigger-sales strategies is not sustainable.

This says that, if firms are sufficiently patient, then the best equilibrium payoffs have the firms set the monopoly price in any cartel phase and risk the chance of a fine. As is common in collusion models with capacity constraints, any collusive price below the monopoly price not only lower profits but it also raises the critical discount factor. Thus, either the profile of explicit trigger-sales strategies, with firms setting the monopoly price during cartel phases, is a SPNE or collusion under explicit trigger-sales strategies is not sustainable at any collusive price.

3.2 Tacit collusion

We next move on to model tacit collusion where there is no active cartel. In this subsection, we present a simplified version of our analysis in Garrod and Olczak (2016). We restrict attention to sequential equilibria in public strategies, in which firms condition their play only on the public history. Such equilibria are known as perfect public equilibria (PPE) (see Fudenberg and Tirole, 1994, p.187-191). In this simplified version, we restrict attention to a particular class of PPE

where firms follow a strategy profile in which, similar to Green and Porter (1984) and Tirole (1988), firms punish each other by reverting to the static Nash equilibrium for a fixed number of periods, if they receive a bad signal during a collusive phase. We formally describe the strategy profile below and refer to it as tacit trigger-sales strategies. This simplified version generates identical results as in Garrod and Olczak (2016) where we solve for the optimal symmetric PPE using the techniques of Abreu *et al.* (1986, 1990).

Tacit trigger-sales strategies are formally defined as follows. There are ‘collusive phases’ and ‘punishment phases’. Suppose firms are in a collusive phase in period t . In any such period, a firm should set the collusive price, p^c . If $y_t = \bar{y}$, such that all firms received sales above their trigger levels, then the collusive phase continues into the next period $t + 1$. If $y_t = \underline{y}$, such that at least one firm received sales below its trigger level, then firms enter a punishment phase in the next period $t + 1$. In the punishment phase, each firm should play the static Nash equilibrium for T periods, after which a new collusive phase begins. This sequence repeats in any future period in a collusive phase.

3.2.1 Collusive profits

Thus, denoting firm i ’s expected (normalised) profit in a collusive phase as $k_i V^c$ and its expected (normalised) profit at the start of a punishment phase as $k_i V^p$, it follows that if all firms follow trigger-sales strategies, then:

$$\begin{aligned} k_i V^c &= (1 - \delta) \pi_i(p^c) + \delta [(1 - \Pr(y|p^c)) k_i V^c + \Pr(\underline{y}|p^c) k_i V^p] \\ k_i V^p &= (1 - \delta) \sum_{t=0}^{T-1} \delta^t \pi_i^N(k_n) + \delta^T k_i V^c, \end{aligned}$$

for all i , where $\Pr(\underline{y}|p^c) = G(m^*(k_1, \bar{m}))$ from (5). Substituting v_i^p into v_i^c and solving yields:

$$k_i V^c = \pi_i^N(k_n) + \frac{(1 - \delta)}{1 - \delta + G(m^*(k_1, \bar{m})) \delta (1 - \delta^T)} (\pi_i(p^c) - \pi_i^N(k_n)), \quad (12)$$

where it is then easy to check that $\pi_i(p^c) \geq k_i V^c > k_i V^p$ for any $T > 0$ and that $k_i V^p > \pi_i^N(k_n)$ for any $T < \infty$.

3.2.2 Incentive compatibility constraints

The profile of tacit trigger-sales strategies is a PPE if, for each date t and any history h^t , the strategies yield a Nash equilibrium from that date on. We say that collusion is not sustainable under trigger-sales strategies if no such equilibrium strategies exist. Given firms play the static Nash equilibrium during each period of the punishment phase, it is clear that they have no

incentive to deviate in any such periods. Thus, we need only consider deviations during collusive phases, in which case $\Pr(\underline{y}|p_i, p^c) = 1$ for any $p_i \neq p^c$ from (5). The incentive compatibility constraint (ICC) for firm i is as follows:

$$k_i V^c \geq (1 - \delta) k_i p^c + \delta k_i V^p, \forall i. \quad (13)$$

This says that firm i will not deviate in any period in a collusive phase if it cannot gain by marginally undercutting p^c to supply its full capacity k_i .¹¹ Notice that (13) is never satisfied for any $\bar{m} \geq K$, as then $G(m^*(k_1, \bar{m})) = 1$, from (5). Thus, collusion is not sustainable under tacit trigger-sales strategies if $\bar{m} \geq K$, so we can henceforth focus on the case where $\bar{m} < K$.

Substituting $k_i V^p$ and $k_i V^c$ into (13), then rearranging yields:

$$(1 - G(m^*(k_1, \bar{m}))) K (p^c - \underline{p}) - \frac{(K - \hat{m}) p^c}{\delta} \geq \delta^T [(1 - G(m^*(k_1, \bar{m}))) K (p^c - \underline{p}) - (K - \hat{m}) p^c]. \quad (14)$$

It follows from the fact that (14) is independent of k_i that if the ICC holds for firm i , then it also holds for all other firms $j \neq i$. This implies that, despite potential asymmetries, each firm has the same incentive to deviate as its rivals. Furthermore, note that (14) can only hold if the left-hand side and the expression in square brackets on the right-hand side are positive, since $\delta^T \in (0, 1]$ for all $T \in [0, \infty)$. Thus, similar to Green and Porter (1984) and Tirole (1988), it follows from (14) that there are three necessary conditions for the profile of tacit triggers-sales strategies to be a PPE. First, the length of the punishment phase must be sufficiently long, where the critical length, denoted $T^*(k_1, k_n, p^c)$, is implicitly defined by the level of T where (14) holds with equality. Second, firms must also be sufficiently patient, such that:

$$\delta \geq \frac{(K - \hat{m}) p^c}{(1 - G(m^*(k_1, \bar{m}))) K (p^c - \underline{p})} \equiv \delta_c^*(k_1, k_n, p^c), \quad (15)$$

in which case the left-hand side of (14) is positive. Note that $T^*(k_1, k_n, p^c) \rightarrow \infty$ if $\delta = \delta_c^*(k_1, k_n, p^c)$ and $T^*(k_1, k_n, p^c) < \infty$ for any $\delta > \delta_c^*(k_1, k_n, p^c)$. This implies that even a punishment phase that lasts an infinite number of periods is insufficient to outweigh the short-term benefit from deviating, if firms are not sufficiently patient. Third, the probability of receiving a bad signal must be sufficiently low, where:

$$G(m^*(k_1, \bar{m})) < 1 - \frac{(K - \hat{m}) p^c}{K (p^c - \underline{p})}, \quad (16)$$

such that the expression in square brackets in (14) is positive. Note that (16) guarantees $\delta_c^*(k_1, k_n, p^c) < 1$, which implies that if this condition is not met, then the firms are not sufficiently patient for any δ , even if a punishment phase lasts an infinite number of periods.

¹¹This is firm i 's optimal deviation since $p^c > \underline{p}$.

3.2.3 Equilibria

Proposition 5 solves for the best PPE payoffs. We refer to this as tacit collusion under imperfect monitoring.

Proposition 2. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} < K$, there exists a unique level of market demand, $\bar{x}(k_1, k_n) \in (\underline{x}(k_1), K)$, that solves $G(m^*(k_1, \bar{x}(k_1, k_n))) = \frac{K-n}{K} < 1$, such that, if $\underline{m} \leq \bar{m} < \bar{x}(k_1, k_n)$ and if $\delta \geq \delta_c^*(k_1, k_n, p^m) = \frac{1}{1-G(m^*(k_1, \bar{m}))} \frac{k_n}{K} \in (\frac{k_n}{K}, 1)$, then the best equilibrium payoff for firm i under tacit trigger-sales strategies is:*

$$k_i V_c^* = \frac{k_i}{K} \left(\frac{\hat{m} - G(m^*(k_1, \bar{m}))K}{1 - G(m^*(k_1, \bar{m}))} \right) \in \left(\pi_i^N(k_n), \frac{k_i}{K} \hat{m} \right), \forall i.$$

Otherwise, collusion under tacit trigger-sales strategies is not sustainable.

This says that, if the necessary conditions (15) and (16) are satisfied, then the best equilibrium payoffs have the firms set the monopoly price in any period in a collusive phase and the length of the punishment phases is set at the level where the ICC (14) is binding with no slack. Nevertheless, despite the fact that firms set the monopoly price during collusive phases, the sum of the best equilibrium payoffs is below the monopoly profit, because punishment phases occur on the equilibrium path under imperfect monitoring. Finally, if firms set a collusive price below the monopoly price, it not only lowers profits but it also raises the critical discount factor. Thus, either the profile of tacit trigger-sales strategies, with firms setting the monopoly price during collusive phases, is a PPE or collusion under tacit trigger-sales strategies is not sustainable at any collusive price.

4 Cartels and Asymmetries

We now use our equilibrium analysis to analyse the conditions under which firms have an incentive to form a cartel. To do so, we restrict attention to the case where $F \in [0, \bar{F})$ and $\delta \geq \delta_e^*(k_1, k_n)$, such that explicit collusion can be sustainable as an equilibrium, from Proposition 1. We first analyse such incentives by comparing, from the cartel's perspective, the expected benefits and costs of a cartel phase, which we define as a sequence of periods that begins the period following the detection of the last cartel phase and ends the period of detection. The expected duration of a cartel phase is $\sum_{t=1}^{\infty} t(1-\theta)^{t-1} = 1/\theta$. Then we consider how these incentives changes as the firms' capacities change. We say that a firm has an incentive to be a member of a cartel if the profits under explicit collusion are strictly greater than under any other alternative.

4.1 Incentives to cartelise

We begin by considering the firms' incentives to form a cartel when the counterfactual is the static Nash equilibrium. This comparison is important for two reasons. First, it is a useful benchmark for our analysis, because it is consistent with the analysis of the previous literature, which exogenously assumes that firms cannot collude without forming a cartel. Second, it is also the appropriate comparison in our framework if tacit collusion is not sustainable. Following this analysis, we then consider the firms' incentives to form a cartel when the counterfactual is tacit collusion.

When the counterfactual is the static Nash equilibrium, it follows from Proposition 1 that firm i , for all i , has an incentive to be a member of a cartel for all $F \in [0, \bar{F})$ and $\delta \geq \delta_e^*(k_1, k_n)$. This is due to the fact that $k_i V_e^* > \pi_i^N(k_n)$ for such conditions, where the critical fine \bar{F} is given by rewriting this inequality, such that:

$$F < \frac{1}{\theta} (S(p^m) - \underline{p}) \equiv \bar{F}. \quad (17)$$

The left-hand side of (17) is the expected cost of a cartel phase per unit of capacity from the cartel's perspective. The right-hand side of (17) is the expected benefit of a cartel phase per unit of capacity, and this is composed of the multiplication of two terms. The term in brackets is the per-period difference between the profits per unit of capacity of explicit collusion and of the static Nash equilibrium. The other term is the expected length of a cartel phase.

Proposition 3 next considers the firms' incentives to form a cartel when the counterfactual is tacit collusion. In this case, we must restrict attention to $\underline{m} < \bar{m} < \bar{x}(k_1, k_n)$, such that tacit collusion is sustainable as an equilibrium if $\delta \geq \delta_c^*$, from Proposition 2.

Proposition 3. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} \leq \bar{m} < \bar{x}(k_1, k_n)$, there exists a unique fine per unit of capacity, $F^* \in [0, \bar{F})$, that solves $V_e^* = V_c^*$, such that:*

- i) if $F \in [0, F^*)$, then explicit collusion is strictly more profitable than tacit collusion, $V_e^* > V_c^*$, for all $\delta \geq \delta_c^*$, where $\delta_c^* > \delta_e^*$;*
- ii) if $F \in [F^*, \bar{F})$, then explicit collusion is (weakly) less profitable than tacit collusion, $V_e^* \leq V_c^*$, for all $\delta \geq \delta_e^*$, where $\delta_e^* \geq \delta_c^*$.*

Proposition 3 shows that not only is explicit collusion more (less) profitable than tacit collusion when fines are sufficiently low (high), in addition, the preferred form of collusion is always sustainable whenever the less profitable alternative is. To determine the critical fine level, F^* ,

note that when the counterfactual is tacit collusion, firm i has an incentive to be a member of a cartel, if it increases its profits such that $k_i V_e^* > k_i V_c^*$. Expressing this inequality differently yields:

$$F < \frac{1}{\theta} (S(p^m) - V_c^*) \equiv F^*, \quad (18)$$

where $S(p^m) = \frac{\hat{m}}{K}$ for all $\bar{m} < \bar{x}(k_1, k_n)$ from $\bar{x}(k_1, k_n) < K$ and (2). Similar to (17), the above inequality compares from the cartel's perspective the expected cost of a cartel phase per unit of capacity on the left-hand side against the expected benefit per unit of capacity on the right-hand side. The only difference is that the term in brackets on the right-hand side of (18) is the per-period difference between the profits per unit of capacity of explicit collusion and of tacit collusion, where $V_c^* \in (\underline{p}, \frac{\hat{m}}{K})$ for all $\bar{m} \in [\underline{m}, \bar{x}(k_1, k_n))$ from Proposition 2.

Figure 1 sketches how these two critical fines vary with the maximum market demand, \bar{m} , assuming a mean preserving spread. It shows that if $\bar{m} \leq \underline{x}(k_1)$, such that there is perfect monitoring, then the critical fine F^* equals zero whereas \bar{F} is positive. The reason is that firms are able to extract the monopoly profit through tacit collusion, $V_c^* = \frac{\hat{m}}{K}$, so any positive fine suffices to ensure explicit collusion is less profitable than tacit collusion. In contrast, the static Nash equilibrium is below the monopoly level, so a fine sufficiently close to zero ensures explicit collusion is strictly more profitable than the static Nash equilibrium. Furthermore, if $\bar{m} \in [\underline{x}(k_1), \bar{x}(k_1, k_n))$, such that there is imperfect monitoring, then the critical fine, F^* , is positive but it less than \bar{F} . This is due to the fact that the profits from tacit collusion are below the monopoly level but are above the profits of the static Nash equilibrium, $V_c^* > \underline{p}$, from Proposition 1. The two fines converge at $\bar{m} = \bar{x}(k_1, k_n)$ where $V_e^* = V_c^* = \underline{p}$.

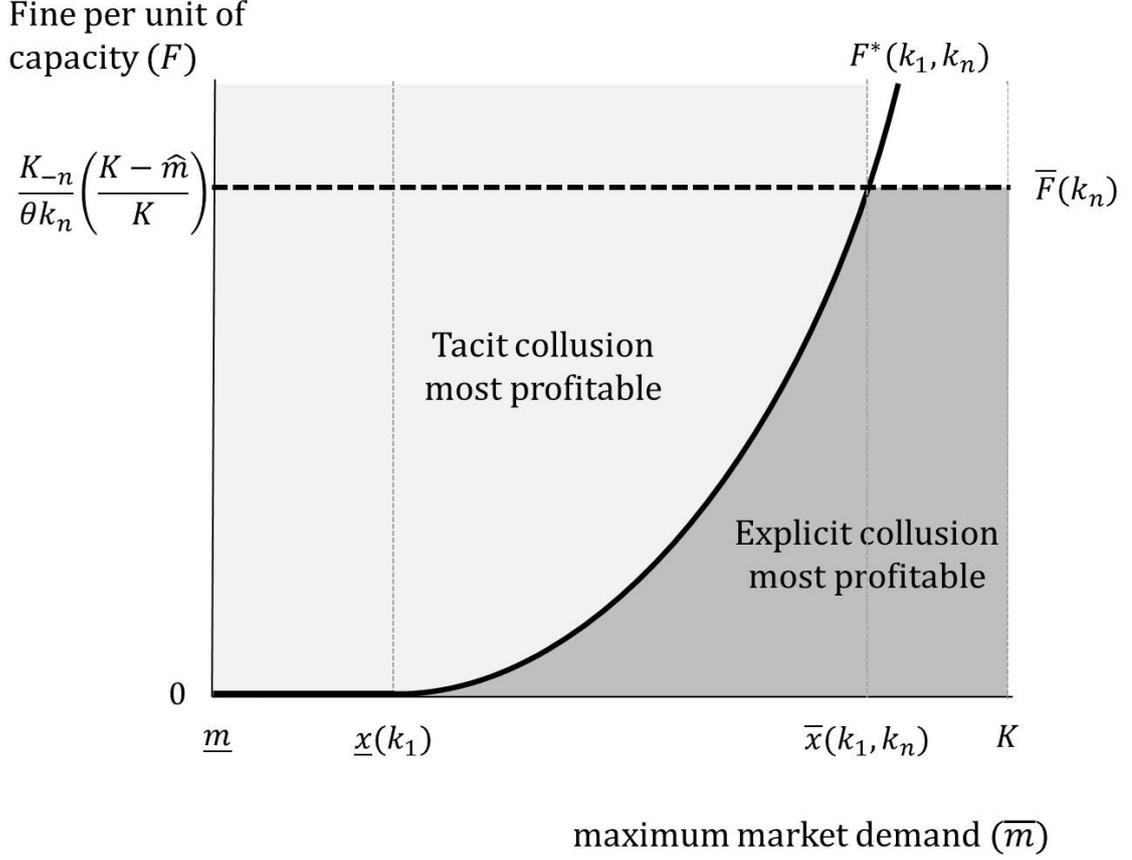


Figure 1: critical fines

4.2 Comparing capacity distributions

In this subsection, we analyse the effects of asymmetries on the incentives to form a cartel. When changing the capacity distribution throughout this section, we hold total capacity and the number of firms constant. This implies that any such changes in the capacity of a given firm will require capacity to be reallocated from a rival. For example, increasing the size of the smallest firm in a duopoly implies that the capacity of the largest firm decreases. Consequently, 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 $\sum_{i \neq j} \frac{\partial k_i}{\partial k_j} = -1$. In what follows, we restrict the discussion to capacity reallocations that directly affect the equilibrium analysis, and this is only the case for changes to the capacity of the smallest firm or the largest firm.

Proposition 4 analyses the effects of reallocating capacity on the critical fines, \bar{F} . It says, consistent with the conventional wisdom, that higher fines are needed to deter cartels when firms have symmetric capacities in our framework, if tacit collusion is not taken into account or indeed if it is not sustainable.

Proposition 4. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} < K$, the critical fine, \bar{F} , is strictly decreasing in the capacity of the largest firm, k_n .*

Increasing the capacity of the largest firm, k_n , reduces the critical fine. The reason is that competition is less intense as the largest firm gets larger, so the static Nash equilibrium profits are higher. Consequently, the expected benefit of a cartel phase falls, because the per-period difference between the profits per unit of capacity of explicit collusion and of the static Nash equilibrium is smaller. This implies that the critical fine, \bar{F} , decreases.¹²

Proposition 4 implies that the critical fine is largest when firms' capacities are symmetric, $k_1 = k_n = K/n$. The reason is that the per-period difference between the collusive profits and the static Nash equilibrium is greatest when the largest firm is as small as possible. Consequently, it follows that if the fine is set below the critical level of symmetric capacities, then the fine would only deter cartels in sufficiently asymmetric markets. Thus, consistent with the conventional wisdom, if tacit collusion is not taken into account or indeed if it is not sustainable, then cartels would arise in markets with least asymmetries. However, we next show that this result does not always hold, in contrast to the conventional wisdom and consistent with the empirical evidence, if the possibility of tacit collusion is taken into account.

Proposition 5 analyses the effects of reallocating capacity among the firms on the critical level, F^* .

Proposition 5. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} \leq K$, the critical fine, F^* , is (weakly) decreasing in the capacity of the smallest firm, k_1 .*

Increasing the capacity of the smallest firm, k_1 , lowers the critical fine if $\bar{m} \in (\underline{x}(k_1), K)$, otherwise the fine is independent of k_1 . This is because the expected benefit of a cartel phase decreases because the difference between the per-period profits of explicit collusion and tacit

¹²Furthermore, the higher static Nash equilibrium profit also makes both the communication and pricing ICCs tighter than before, so the critical discount factor rises.

collusion is smaller. This decreases the incentive to collude explicitly relative to tacit collusion. The intuition is that the per-period profits of explicit collusion, $\frac{\hat{m}}{K}$, are unaffected but, if there is imperfect monitoring, then the per-period profit of tacit collusion, V_c^* , is higher when the smallest firm is larger. This in itself is caused by two effects. First, as the capacity of the smallest firm increases, it is less likely that firms' sales will be below their trigger levels when they set a common price. Thus, profits rise on the equilibrium path, other things equal, because collusive periods are less likely to switch to punishment periods than before. Second, such an increase in profits also introduces slack into the ICC for tacit collusion, so the length of the punishment phase is reduced to ensure that it is binding with no slack. Both effects imply that firms expect there to be more collusive periods on the equilibrium path than when the smallest firm has more capacity, so the best tacitly collusive payoffs rise.

Proposition 5 implies, in contrast with Propostion 4, that the critical fine is smallest when firms' capacities are symmetric, $k_1 = k_n = K/n$. The reason is that the per-period difference between the collusive profits of explicit and of tacit collusion are closest when the smallest firm is as large as possible. It follows that the smallest firm is as large as possible in a symmetric duopoly, so the critical fine is lowest for this case and that, for example, it would be lower for a symmetric triopoly than an asymmetric duopoly with $k_1 < K/3$. Thus, it follows that if the fine is set above the critical level of a symmetric triopoly but below the level of an asymmetric duopoly then fine would only deter cartels in sufficiently symmetric markets. Consequently, in contrast with the conventional wisdom and with the previous literature but consistent with the evidence, it follows that lower fines are required to deter cartels in less asymmetric markets, when tacit collusion is the counterfactual.

5 Concluding remarks

The conventional wisdom is that cartels should happen in symmetric market structures, so competition agencies should actively seek them there. Our model suggests that once tacit collusion is take into account, firms have a greater incentive to form cartels in more asymmetric market structures. This fits well with the previously puzzling empirical evidence from detected cartels which, as described in the introduction, shows that many tend to include a large number of very asymmetric firms.

Our paper emphasises the need to recognise tacit collusion as a potential substitute for explicit collusion. This raises a number of further policy implications. First, it strengthens concerns that

have previously been raised about the use of structural screens based on market characteristics to detect cartel activity. For example, Harrington (2006a) argues that, whilst structural indicators may suggest that cartel activity is more likely, there is still a high chance of false positives because, based on the available evidence, cartels are relatively rare and a collusive outcome is just one of the possible equilibria. In addition, our paper demonstrates that if the structural indicators are identified according to the conventional wisdom, it will not even be the case that they indicate that cartel activity is more likely. Instead, the screen may simply pick-up markets where explicit collusion is unnecessary as tacit collusion is sustainable and more profitable.

Second, as has been recognised by Harrington (2011), tacit collusion following cartel detected, has implications for using post-detection prices to estimate the but-for price had the cartel not been in place for damage calculations. Furthermore, once tacit collusion is recognised as an alternative to explicit collusion, further questions are raised about the appropriate counterfactual that should be used and whether the appropriate but-for price is at the competitive level. More generally, our analysis has taken a normative stance by considering when, for a given level of fines, cartel behaviour occurs. However, a related body of theoretical literature considers how alternative penalty regimes affect cartel formation and pricing (see for example Katsoulacos *et al.* 2014). As yet, to the best of our knowledge, this literature has not taken into account that tacit collusion may be a substitute for cartel formation.

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] Arnold, M. (2000) "Costly Search, Capacity Constraints, and Bertrand Equilibrium Price Dispersion," *International Economic Review*, 41(1), 117-132
- [4] Arnold, M. and Saliba, C. (2011) "Asymmetric Capacity Constraints and Equilibrium Price Dispersion," *Economics Letters*, 111(2), 158-160
- [5] Athey, S. and Bagwell, K. (2001) "Optimal Collusion with Private Information," *RAND Journal of Economics*, 32(3), 428-465

- [6] Athey, S. and Bagwell, K. (2008) "Collusion with Persistent Cost Shocks," *Econometrica*, 76(3), 493-540
- [7] Aubert, C., Rey, P. and Kovacic, B. (2006) "The Impact of Leniency and Whistle-Blowing Programs on Cartels," *International Journal of Industrial Organization*, 24(6), 1241-1266
- [8] Awaya, Y. and Krishna, V. (2014) "On Tacit versus Explicit Collusion," mimeo, Penn State University
- [9] Bos, I. and Harrington, J. (2010) "Endogenous Cartel Formation with Heterogeneous Firms," *RAND Journal of Economics*, 41(1), 92-117
- [10] Bos, I. and Harrington, J. (2014) "Competition Policy and Cartel Size," *International Economic Review*, forthcoming
- [11] Chen, Z. and Rey, P. (2013) "On the Design of Leniency Programs," *Journal of Law and Economics*, 56(4), 917 - 957
- [12] Compte, O. (1998) "Communication in Repeated Games with Imperfect Monitoring," *Econometrica*, 66(3), 597-626
- [13] Compte, O., Jenny, F. and Rey, P. (2002) "Capacity Constraints, Mergers and Collusion," *European Economic Review*, 46(1), 1-29
- [14] Connor, J. (2010) "Recidivism Revealed: Private International Cartels 1990-2009," *Competition Policy International*, 6(2)
- [15] 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
- [16] 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
- [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] Fudenberg, D. and Tirole, J. (1994), *Game Theory*, MIT Press
- [19] Garrod, L. and Olczak, M. (2016) "Collusion under Imperfect Monitoring with Asymmetric Firms," mimeo

- [20] Green, E. and Porter, R. (1984) “Non-Cooperative Collusion under Imperfect Price Information,” *Econometrica*, 52, 87-100
- [21] Harrington, J. (1989) “Collusion Among Asymmetric Firms: The Case of Different Discount Factors,” *International Journal of Industrial Organization*, 7, 289-307
- [22] Harrington, J. (2006a) “Behavioural Screening and the Detection of Cartels” in *European Competition Law Annual 2006: Enforcement of Prohibition of Cartels*, Hart Publishing, Oxford/Portland, Oregon
- [23] Harrington, J. (2006b) “How do Cartels Operate?” *Foundations and Trends in Microeconomics*, 2(1), 1-105
- [24] Harrington, J. (2008a) “Detecting Cartels,” in Buccirosi, P., (ed.), *Handbook in Antitrust Economics*, MIT Press
- [25] Harrington, J. (2008b) “Optimal Corporate Leniency Programs,” *Journal of Industrial Economics*, 56(2), 215–246
- [26] Harrington, J. (2011) “Post-Cartel Pricing during Litigation,” *Global Competition Litigation Review*, forthcoming
- [27] Harrington, J. (2010) “Lectures on Collusive Practices,” Presentation at CRESSE, July 1st-2nd
- [28] Harrington, J. and Skrzypacz, A. (2007) “Collusion under Monitoring of Sales,” *RAND Journal of Economics*, 38(2), 314-331
- [29] Harrington, J. and Skrzypacz, A. (2011) “Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices,” *American Economic Review*, 101(6), 2425-2449
- [30] Hay, G. and Kelley, D. (1974) “An Empirical Survey of Price Fixing Conspiracies” *Journal of Law and Economics*, 17(1), 13-38
- [31] International Competition Network (2008) “Setting of Fines for Cartels in ICN Jurisdictions,” Report to the 7th ICN Annual Conference, Kyoto
- [32] Kandori, M. and Matsushima, H. (1998) “Private Observation, Communication and Collusion,” *Econometrica*, 66(3), 627-652

- [33] Katsoulacos, Y. Motchenkova, E. and Ulph, D. (2014) “Penalizing Cartels: The Case for Basing Penalties on Price Overcharge,” mimeo
- [34] Kühn, K.-U., (2004) “The Coordinated Effects of Mergers in Differentiated Product Markets,” John M. Olin Centre for Law and Economics Working Paper 04-020
- [35] Lambson, E. (1994) “Some Results on Optimal Penal Codes in Asymmetric Bertrand Supergames,” *Journal of Economic Theory*, 62(2), 444-468
- [36] Levenstein, M. and Suslow, V. (2006) “What Determines Cartel Success?” *Journal of Economic Literature*, 44(1), 43-95
- [37] Martin, S. (2006) “Competition Policy, Collusion and Tacit Collusion,” *International Journal of Industrial Organization*, 24(6), 1299-1332
- [38] Miklós-Thal, J. (2011) “Optimal Collusion under Cost Asymmetry,” *Economic Theory*, 46(1), 99-125
- [39] Motta, M. (2004) *Competition Policy: Theory and Practice*, Cambridge University Press
- [40] Motta, M. and Polo, M. (2003) “Leniency Programs and Cartel Prosecution,” *International Journal of Industrial Organization*, 21(3), 347-379
- [41] Spagnolo, G. (2005) “Divide et Impera: Optimal Leniency Programs,” CEPR. Discussion Paper No. 4840
- [42] Spector, D. (2015) “Facilitating Collusion by Exchanging Non-Verifiable Sales Reports,” PSE Working Papers no. 2015-07
- [43] Stephan, A. and Nikpay, A. (2015) “Leniency Decision-Making from a Corporate Perspective: Complex Realities” in Beaton-Wells, C. and Tran, C. (editors), *Anti-Cartel Enforcement in a Contemporary Age: The Leniency Religion*, Hart Publishing, Oxford
- [44] Stigler, G. (1964) “A Theory of Oligopoly,” *Journal of Political Economy*, 72, 44–61
- [45] Tirole, J. (1988) *The Theory of Industrial Organisation*, MIT Press
- [46] Vasconcelos, H. (2005) “Tacit Collusion, Cost Asymmetries, and Mergers,” *RAND Journal of Economics*, 36(1), 39-62

Appendix A

Proof of Proposition 1. Any given price p^c is supportable if $\delta \geq \delta_e^* = \max\{\delta_e^m, \delta_e^p\}$, such that both the pricing and communication ICCs are satisfied. Given V^e is strictly increasing in p^c and that the critical discount factors δ_e^m and δ_e^p are strictly decreasing in p^c , it follows that the equilibrium payoff is highest and the critical discount factor is lowest if $p^c = 1$. Thus, V_e^* is as claimed. Furthermore, (9) and (11) imply that $\delta_e^p > \delta_e^m$ for all F . This implies that there exists a unique $\delta_e^*(k_1, k_n, p^m, F) = \delta_e^p(k_1, k_n, p^m, F)$ such that firm i 's highest equilibrium payoff is $k_i V_e^* = \pi_i(p^m) - \hat{\theta}F$ if $\delta \geq \delta_e^*$. Otherwise, explicit collusion is not sustainable. ■

Proof of Proportion 2. Given v_i^c is strictly decreasing in T , the best equilibrium payoffs for firm i can be found by evaluating v_i^c at $T^*(k_1, k_n, p^c)$. Thus, it follows from (14) that:

$$1 - \delta^{T^*} = \frac{(1 - \delta)(K - \hat{m})p^c}{\delta[(1 - G(m^*(k_1, \bar{m})))K(p^c - p) - (K - \hat{m})p^c]}.$$

Then, by substituting this into (12), yields:

$$v_i^c = \frac{k_i}{K} \left(\frac{\hat{m} - G(m^*(k_1, \bar{m}))K}{1 - G(m^*(k_1, \bar{m}))} \right) p^c, \quad \forall i.$$

This is strictly increasing in p^c , so $p^c = 1$ and v_i^* is as claimed. Substituting $p^c = 1$ into (15) and (16) yields:

$$\delta \geq \delta_c^*(k_1, k_n, p^m) = \frac{1}{1 - G(m^*(k_1, \bar{m}))} \frac{k_n}{K}$$

and

$$G(m^*(k_1, \bar{m})) < \frac{K-n}{K}. \quad (19)$$

Furthermore, note that $\delta^*(k_1, k_n, p^c)$ is strictly increasing in p^c , so collusion under trigger-sales strategies is not sustainable for any $\delta < \delta^*(k_1, k_n, p^m)$. Finally, it follows from $\frac{\partial G(m^*)}{\partial \bar{m}} > 0$ that there is a unique level of \bar{m} , denoted $\bar{x}(k_1, k_n)$, that sets $G(m^*(k_1, \bar{m})) = \frac{K-n}{K} < 1$, where $\bar{x}(k_1, k_n) < K$ and where $G(m^*(k_1, \bar{m})) \in \left[0, \frac{K-n}{K}\right)$ for all $\bar{m} \in [\underline{x}(k_1), \bar{x}(k_1, k_n))$. This implies $\delta_c^*(k_1, k_n, p^m) \in \left(\frac{k_n}{K}, 1\right)$ and $v_i^* \in \left(\pi_i^N(k_i, k_n, \hat{m}), \frac{k_i}{K}\hat{m}\right)$ for all $\bar{m} \in [\underline{x}(k_1), \bar{x}(k_1, k_n))$. ■

Proof of Proportion 3. First, note that $k_i V_e^* > k_i V_c^*$ if and only if:

$$F < \frac{1}{\theta} \left(\frac{\hat{m}}{K} - V_c^* \right) \equiv F^*.$$

For any $\bar{m} \leq \underline{x}(k_1)$, such that $G(m^*) = 0$, then $F^* = 0$ as $V_c^* = \frac{\hat{m}}{K}$ from Proposition 2. However, if $\underline{x}(k_1) < \bar{m} < \bar{x}(k_1, k_n)$, then $F^* > 0$, and given $k_i V_c^* \in \left(\pi_i^N(k_n), \pi_i(p^m)\right)$ for all

$\underline{m} < \bar{m} < \bar{x}(k_1, k_n)$, it follows that $F^* < \bar{F}$. It also follows from $\lim_{\bar{m} \rightarrow \bar{x}(k_1, k_n)} k_i V_c^* = \pi_i^N(k_n)$ that $\lim_{\bar{m} \rightarrow \bar{x}(k_1, k_n)} F^* = \bar{F}$. Finally note that $\delta_c^* > \delta_e^p$ if and only if $F < F^*$. This implies that if $F \in [0, F^*)$, then $V_c^* < V_e^*$, for all $\delta \geq \delta_c^*$, where $\delta_c^* > \delta_e^* = \delta_e^p$ from the proof of Proposition 1. Whereas, if $F \in [F^*, \bar{F})$, then $V_e^* \leq V_c^*$, for all $\delta \geq \delta_e^*$, where $\delta_e^* = \delta_e^p \geq \delta_c^*$. ■

Proof of Proposition 4. Differentiating $\bar{F} = \frac{S(p^c) - p}{\theta}$ with respect to k_j yields:

$$\frac{\partial \bar{F}}{\partial k_j} = \frac{1}{\theta} \left(-\frac{\partial k_n}{\partial k_j} \frac{\partial p}{\partial k_n} \right).$$

Thus, $\frac{\partial \bar{F}}{\partial k_n} < 0$ from $\frac{\partial k_n}{\partial k_n} = 1$ and $\frac{\partial p}{\partial k_n} > 0$. ■

Proof of Proposition 5. Differentiating $F^* = \frac{S(p^c) - V_c^*}{\theta}$ with respect to k_j yields:

$$\frac{\partial F^*}{\partial k_j} = \frac{1}{\theta} \frac{\partial k_1}{\partial k_j} \left(-\frac{\partial V_c^*}{\partial k_1} \right).$$

Differentiating V_c^* with respect to k_1 yields:

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

where $\frac{\partial V_c^*}{\partial k_1} > 0$ from $\frac{\partial m^*}{\partial k_1} < 0$ and $K > \hat{m}$. Thus, $\frac{\partial F^*}{\partial k_1} < 0$ from $\frac{\partial k_1}{\partial k_1} = 1$. ■