MPRA

Munich Personal RePEc Archive

Collusion under Imperfect Monitoring with Asymmetric Firms

Luke, Garrod and Matthew, Olczak

Loughborough University, Aston University

17 March 2016

Online at https://mpra.ub.uni-muenchen.de/70647/
MPRA Paper No. 70647, posted 05 May 2016 16:17 UTC

Collusion under Imperfect Monitoring with Asymmetric Firms*

Luke Garrod† and Matthew Olczak‡

March 17, 2016

Abstract

We explore the effects of asymmetries in capacity constraints on collusion where market demand is uncertain and where firms must monitor the agreement through their privately observed sales and prices. In this private monitoring setting, we show that all firms can infer when at least one firm's sales are below some firm-specific "trigger level". This public information ensures that firms can detect deviations perfectly if fluctuations in market demand are sufficiently small. Otherwise, there can be collusion under imperfect public monitoring where punishment phases occur on the equilibrium path. We find that symmetry faciliates collusion. Yet, we also show that if the fluctuations in market demand are sufficiently large, then the collusive prices of symmetric capacity distributions are actually lower than the competitive prices of asymmetric capacity distributions. We draw conclusions for merger policy.

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

Key words: capacity constraints, mergers, collusion, imperfect monitoring

*This paper builds on Centre for Competition Policy Working Paper 10-3. We are grateful for comments from Steve Davies, Joe Harrington, Morten Hviid, Bruce Lyons, Volker Nocke, Peter Ormosi, Chris M. Wilson, and two anonymous referees. We have also benefitted from the comments of seminar participants at the ESRC Centre for Competition Policy, the International Industrial Organization Conference (IIOC) 2014, the Competition and Regulation European Summer School (CRESSE) conference 2014, the European Association of Research in Industrial Economics (EARIE) conference 2014, and the Bergen Center for Competition Law and Economics (BECCLE) conference 2015. 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

The recent collusion theory literature has developed a clear consensus that asymmetries hinder collusion. For example, this result is robust to whether asymmetries are in terms of firms' capacity constraints (see Compte *et al.*, 2002; Vasconcelos, 2005; and Bos and Harrington, 2010 and 2015) or the number of differentiated products that each firm sells (see Kühn, 2004). These papers in particular have been important for merger policy as they have highlighted which types of mergers can cause coordinated effects, that is, an increased likelihood or sustainability of tacit collusion post-merger. More specifically, with respect to capacity constraints, Compte *et al.* (2002) show that collusion is more difficult as the capacity of the largest firm is increased through a merger, and Vasconcelos (2005) finds that collusion is hindered when the largest firm is larger or when the smallest firm is smaller. Bos and Harrington (2010) show that increasing the capacity of medium-sized firms can facilitate collusion, if only a subset of firms in the market are involved in the collusion.[1]

In practice, the degree to which firms can monitor each other's actions plays an important part in determining whether a merger causes coordinated effects. Yet, all of the papers above assume there is perfect observability of rivals' actions, so deviations from the collusive strategies will be detected immediately. In contrast, many mergers occur in markets in which there is the potential for 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. Consequently, it is inappropriate to consider the effects of such mergers in terms of collusion under perfect observability. Instead, they should be considered in the context of imperfect monitoring, where firms are uncertain over whether their rivals have followed their collusive strategies or not (see Green and Porter, 1984; and Harrington and Skrzypacz, 2007 and 2011). However, while the models in this literature provide many interesting insights into the sustainability of collusion, it is difficult to draw implications for merger policy from them, because they analyse collusion with symmetric firms.

In this paper, we begin to fill this gap in the literature by exploring the effects of asymmetries in capacity constraints on collusion under imperfect monitoring. We achieve this by extending Compte *et al.* (2002) to a setting where there is demand uncertainty and where firms never directly observe their rivals' prices or sales. Thus, similar to the imperfect monitoring setting

[1]Fonseca and Normann (2008, 2012) also find that asymmetries in capacity constraints hinder collusion in laboratory experiments.

first discussed by Stigler (1964), each firm must monitor the collusive agreement using their own privately observed sales. In this regard, our model is related to Tirole's (1988, p.262-264) model of private monitoring that captures the results of Green and Porter (1984) in a Bertrand framework (see also Campbell *et al.*, 2005, and Amelio and Biancini, 2010). Yet, unlike Tirole (1988), where there is a chance in each period that market demand will be zero, in our model market demand is drawn from an interval, where all possible states of demand are positive. We use this model to investigate whether collusion is facilitated or hindered as capacity is reallocated among the firms to draw implications for merger policy.

Using information from their privately observed sales, we show that all firms can always infer when at least one firm's sales are below some firm-specific "trigger level". The trigger level for each firm 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 the highest-priced firms will receive sales below their trigger levels and the lower-priced firm(s) can infer this.[2] Consequently, we restrict attention to equilibria in public strategies, known as perfect public equilibria, where firms condition their play upon this public information (i.e. whether all firms' sales are greater than their trigger levels or not). We show that if fluctuations in market demand are small, then such strategies ensure monitoring is perfect, because firms will only ever receive sales below their trigger levels if they are undercut. However, if fluctuations in market demand are large, then collusive sales will also fall below the trigger levels when the realisation of market demand is low. This implies, in contrast to Compte *et al.* (2002), that there is uncertainty as to whether rivals have followed the collusive strategies or not, so punishment phases must occur on the equilibrium path to provide firms with the correct incentives to collude.

We solve for an optimal perfect public equilibrium.[3] We find that asymmetries hinder collusion whether monitoring is perfect or imperfect. For instance, the critical discount factor is higher when the largest firm is larger or when the smallest firm is smaller. The reason for the

[2]Note that this information is common knowledge in our framework, because a deviation by a firm impacts the sales of all of its rivals. In contrast, this information may not be common knowledge in a framework where firms' products are spatially differentiated, as in the Salop circle, for example, because then firms located next to a deviator would experience lower sales but other firms may not.

[3]In the main paper, we use a strategy profile, similar to Tirole (1988), where firms revert to the static Nash equilibrium for a number of periods when they receive a bad signal. In an appendix, we use the techniques of Abreu *et al.* (1986, 1990) to show that the approach in the main paper generates the maximal equilibrium profits.

former is that the punishment is weaker when the largest firm has more capacity. The latter is due to the fact that 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, decreasing the size of the smallest firm makes monitoring more difficult. Another implication of this is that punishment phases occur more often on the equilibrium path when the smallest firm has less capacity, and the optimal equilibrium profits are lower as a result. The size of the other firms' capacities do not affect the the equilibrium profits or critical discount factor.

After solving the model, we then use it to draw implications for merger policy. In particular, we analyse both the coordinated and unilateral effects of mergers in a unified framework. Unilateral effects arise if any firm is likely to have an individual incentive to raise prices post-merger. It is well understood that unilateral effects are associated with asymmetric post-merger market structures and coordinated effects are associated with symmetric post-merger market structures (see Ivaldi *et al.*, 2003a and 2003b). As discussed by Kühn (2001) and Motta *et al.* (2003), this implies that there is a tradeoff between such effects when the degree of asymmetry in a market is altered by a merger or divestiture remedy: increasing asymmetries reduces the likelihood of coordinated effects but raises the likelihood of unilateral effects, and vice versa. However, in the previous theoretical literature, these effects have been modelled independently of each other. For example, in the framework of Compte *et al.* (2002), either the monopoly price is sustainable in every period, in which case only coordinated effects matter, or collusion is never sustainable at any price, so only unilateral effects matter. Consequently, their focus is solely on the coordinated effects of mergers on the critical discount factor. In contrast, our model allows for a more continuous treatment of unilateral and coordinated effects, because play can alternate between phases of collusion and competition on the equilibrium path.

The conventional wisdom is that coordinated effects are more harmful to welfare than unilateral effects. The reason, as described by Röller and Mano (2006, p.22), is that "it is preferable that any coordination is by only a subset of firms (i.e. the merging parties) rather than all firms (tacitly)". In other words, the fear is that firms will share the monopoly profits in every future period if collusion is sustainable, so only a merger to monopoly would be equally as bad in terms of unilateral effects. This logic also implies that a merger that disrupts collusion, by enhancing the market power of a single firm, should increase consumer surplus post-merger. However, in contrast to this conventional wisdom, we show that unilateral effects can be more harmful than coordinated effects. This is due to the fact that firms are not able to share the monopoly profits under imperfect monitoring, because punishment phases occur on the equilibrium path. Consequently, a merger that facilitates collusion by distributing capacity symmetrically can be less harmful to welfare than one that creates a near monopoly. We demonstrate that if the fluctuations in market demand are sufficiently large, then the collusive prices of symmetric capacity distributions are lower than the competitive prices of asymmetric capacity distributions.

Finally, our model is distinct from the previous literature that analyses collusion with capacity constraints and fluctuations in market demand. The main difference is that our focus is on mergers, which necessarily requires us to model asymmetries in markets with more than two firms. In contrast, the focus of this other literature is on pricing over the business cycle. For instance, Staiger and Wolak (1992) and Knittel and Lepore (2010) endogenise the choice of capacities in an infinitely repeated game. Despite analysing asymmetric games following the capacity choice stage, they restrict attention to duopoly. Other differences are that there is perfect observability and market demand is known when prices are set. In a similar setting to that just described, Fabra (2006) analyses collusion where firms' capacity constraints are exogeneous but symmetric.

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 repeated game. We first show that there is some public information that firms can condition their play on, and find when monitoring is perfect or imperfect. Then we solve the game and analyse the successfulness of collusion for different capacity distributions. In section 4, we consider the implications for merger policy. Section 5 explores the robustness of our results, and section 6 concludes. All proofs are relegated to appendix A. In appendix B, we use the techniques of Abreu *et al.* (1986, 1990) to show that the approach in the main body of the paper is an optimal equilibrium in that it generates the maximal equilibrium profits. This appendix is best read after section 3.2.

2 The Model

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 and they have a common discount factor, $\delta \in (0, 1)$. 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 $i = \{1, \ldots, n\}$ and $\mathbf{p}_{-it}$ is the vector of prices of all of firm i's rivals. Market demand consists of a mass of m_t (infintesimally 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 distribution $G(m)$, with mean $\widehat{m}$ and density $g(m) > 0$ on the interval $[\underline{m}, \overline{m}]$.

Buyers are informed of prices, so they will want to buy from the cheapest firm. However, the maximum that firm i can supply in any period is k_i, where we let $k_n \geq k_{n-1} \geq \ldots \geq k_1 > 0$, 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 contrast to the buyers, firm i never observes firm j's prices, $p_{j\tau}$, or sales, $s_{j\tau}$, $j \neq i$, for all $\tau \in \{0, \ldots, t\}$. Thus, similar to Tirole (1988), our setting has the feature that all buyers are fully aware of prices, yet all firms are only aware of their own prices. Such a 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.[4]

2.2 Demand allocation and sales

Following the literature (for example, see Vasconcelos, 2005; and Bos and Harrington, 2010 and 2015), we make the common assumption that demand is allocated by the following rule:

The proportional allocation rule

Unsupplied buyers want to buy from the firm(s) with the lowest price among those with spare capacity.

- 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.
- 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.

This allocation rule 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). However, this seems inappropriate in our model,

[4]In contrast to Tirole (1988), our main results simply require that enough buyers are informed of prices to hold, if capacity constraints are binding.

because members of such cartels are likely to have some knowledge of market demand and rivals' sales when allocating demand, which is not present in our setup. Instead, we have tacit collusion in mind where firms play no role in the allocation of the market demand, other than through the prices they set. Thus, in our framework it seems more appropriate to consider the demand allocation in terms of buyers allocating themselves to firms according to the proportional allocation rule.

To see that buyers can allocate themselves according to the proportional allocation rule in our framework, note that buyers want to buy from the firm with strictly the lowest price among those with spare capacity. Consequently, it makes sense that demand is allocated to such a firm until either its capacity is exhausted or the market demand is supplied.[5] However, buyers are indifferent between purchasing from firms with tied prices. So, suppose buyers break the tie by randomly selecting such a firm with a probability equal to the firm's proportion of their joint capacity. This requires that the size of the firms is observable to the buyers and it captures the plausible feature that buyers are more likely to be attracted to larger firms when they are indifferent between two or more firms. Thus, if the joint capacity of the firms with tied prices is insufficient to supply the unsupplied buyers, then such firms will supply their full capacities and the remaining unsupplied buyers will then want to purchase from the firm(s) with the next lowest price among those with spare capacity, and so on. However, if the joint capacity of the firms with tied prices suffices to supply the unsupplied buyers, then it follows from the law of large numbers that such firms will each receive demand equal to its proportion of the joint capacity. We consider the robustness of our results with respect to the allocation rule in section 5.

We also place the following plausible yet potentially restrictive assumption on the capacity distribution:

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

This says that the joint capacity of the smallest firm's rivals should not exceed the minimum market demand, and it 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. To understand the generality of Assumption 1, note that it is not restrictive if all firms can only ever collectively supply as much as the minimum market demand, $\underline{m} \geq K$. Otherwise, for a given level of $\underline{m}$, there is a restriction on the size

[5]Note that in the case where this firm's capacity is insufficient to supply all of the buyers, we do not need to specify which buyers are supplied, since all consumers have the same reservation price. In contrast, when consumers' reservation prices differ, the identities of the buyers who are supplied by the firm has important implications for welfare (see for example Vives, 1999, p.124-6).

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, $\underline{m} \geq 0.5K$. This is in contrast with Tirole's (1988) model, which requires the less realistic assumption that the minimum market demand is zero with some positive probability. Nevertheless, it is clear that Assumption 1 comes with some loss of generality, so we discuss the implications of relaxing it under duopoly in section 5. We also argue in section 4 that Assumption 1 is not very restrictive in the context of mergers. No restriction is placed on the level of the maximum market demand, $\overline{m}$.

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 given by (1), where $\Omega(p_{it})$ denotes the set of firms that price strictly below p_{it} and $p_t^{\max} \equiv \max\{\mathbf{p}_t\}$.

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

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 $\pi_{it}\left(p_{it}, \mathbf{p}_{-it}\right) = p_{it} \int_{\underline{m}}^{\overline{m}} s_{it}\left(p_{it}, \mathbf{p}_{-it}; m\right) g(m) dm$, where we drop time subscripts if there is no ambiguity. Furthermore, we write $\pi_i\left(p\right)$ if $p_j = p \leq 1$ for all j, such that:

$$\pi_i\left(p\right) = \begin{cases} pk_i & \text{if} \quad K \leq \underline{m} \\ pk_i\left(\int_{\underline{m}}^{K} \frac{m}{K} g(m)dm + \int_{K}^{\overline{m}} g(m)dm\right) & \text{if} \quad \underline{m} < K < \overline{m} \\ pk_i \frac{\widehat{m}}{K} & \text{if} \quad \overline{m} \leq K, \end{cases}$$

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

Finally, an important implication of our assumptions is that a firm will meet all demand up to its capacity in any given period. This implies that a deviant is not able to attempt to hide its deviation by limiting the units available at the deviation price. Consequently, our analysis is likely to be appropriate for markets where collusive prices are agreed at the senior management level, but the total output of the firm is determined at a lower level by sales representatives who are unaware of the collusion. This is the setting for many, if not most, cartels (see Harrington, 2006, for evidence from Europe). Nevertheless, we have explored this issue beyond the analysis presented below and can show that our main results are robust to this alternative setting if deviants are only able to limit their sales below demand to some extent and if this is more difficult for large firms than smaller firms. This may be the case, for example, if the senior managers of a deviant firm can only imperfectly limit the total output of their sales representatives, who are unaware of the collusion, and larger firms have more sales representatives which makes this task more problematic.[6]

2.3 Static Nash equilibrium

In this subsection, we analyse the stage game. Consistent with the standard Bertrand-Edgeworth setting, the static Nash equilibrium can be in pure strategies or mixed strategies. While the proof of the former is trivial, we extend the equilibrium analysis in Fonseca and Normann (2008) to our setting of demand uncertainty to solve for the latter. This is also equivalent to the equilibrium analysis of Gal-Or (1984) if firms are symmetric.

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, $\forall$ i:

$$\pi_i^N(k_n) = \begin{cases} \frac{k_i}{k_n}\left(\int_{\underline{m}}^{K}(m - K_{-n})\,g(m)dm + k_n\int_{K}^{\overline{m}} g(m)dm\right) & \text{if} \quad K < \overline{m} \\ \frac{k_i}{k_n}\left(\widehat{m} - K_{-n}\right) & \text{if} \quad \overline{m} \leq K. \end{cases} \tag{2}$$

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 market demand can be below total capacity, firms are not guaranteed to supply their full capacity for every level of demand, so they 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:

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

This defines firm i's minimax payoff. The intuition is that the firm with strictly the highest price expects to supply its full capacity if the realisation of market demand exceeds total capacity, but it expects to supply the residual demand otherwise. It follows from this that 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. This implies

[6] This analysis is available from the authors upon request.

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 $k_i\underline{p} \geq \overline{\pi}_i$. Consequently, the mixed strategy Nash equilibrium profits are given by $\pi_i^N(k_n) = k_i\underline{p}$. This is equivalent to (2), where Assumption 1 is sufficient to ensure that these are nonnegative for all i. The lower bound of the support is $\underline{p}$.

3 Monitoring with Asymmetries

In this section, we analyse the repeated game. We first show that there is some public information that firms can condition their play on, and find when monitoring is perfect or imperfect. We then solve the game and compare the results for alternative capacity distributions. Henceforth, we impose $\underline{m} < K$, as collusion is unnecessary otherwise from Lemma 1.

3.1 Information and monitoring

Under our assumptions, repetitions of the stage game generate private and public information histories. For instance, the private history of firm i in period t is the sequence of its past prices and sales, denoted $z_i^t \equiv (p_{i0}, s_{i0}; \ldots; p_{it-1}, s_{it-1})$. In contrast, a public history is the sequence of information that is observed by all firms, regardless of their actions. In this subsection, we show that the fact that each firm observes its own sales implies that all firms will always know when at least one firm's sales are below some firm-specific "trigger level". As we discuss below, firms can then use public strategies in which they condition their play on this public information.

Formally, let $m^*(k_1, \overline{m}) \equiv \frac{K(\overline{m}-k_1)}{K_{-1}}$ where firm i's trigger level is $s_i^* \equiv \min\left\{\frac{k_i}{K}m^*(k_1,\overline{m}), k_i\right\}$ for all i. As we show below, such 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. This then guarantees that at least one firm will always receive sales below their trigger level, if all firms do not set a common price. Now consider the history $h^t = (y_0, y_1, \ldots, y_{t-1})$ where, for all $\tau = \{0, 1, \ldots, t-1\}$:

$$y_\tau = \begin{cases} \overline{y} & \text{if } s_{i\tau}(p_{i\tau}, \mathbf{p}_{-i\tau}; m_\tau) > s_i^* \;\forall\; i \\ \underline{y} & \text{otherwise.} \end{cases} \tag{4}$$

This says that $y_\tau = \overline{y}$ 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.

We wish to establish that h^t is a public history. This requires that y_τ is common knowledge for all τ, for any z_i^t. Clearly, this is the case if 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. This occurs only if the maximum market demand is above the total capacity, $\overline{m} \geq K$, because then a firm is uncertain as to whether a rival has undercut it on price, even if the firm sells its full capacity. So consider $\overline{m} < K$, where it is possible for firms to receive sales above their trigger levels, since $s_i^* < k_i$. In this case, if all firms do not set a common price, then the sales of the firm(s) with the highest price in the market will never exceed their trigger levels. For instance, for any nonempty set of rivals with a price strictly below $p^{\max}$, $\Omega\left(p^{\max}\right)$, the sales of firm i with $p_i = p^{\max} \leq 1$ are:

$$s_i = \frac{k_i}{K - \sum_{j \in \Omega(p^{\max})} k_j} \left(m_t - \sum_{j \in \Omega(p^{\max})} k_j\right) \leq \frac{k_i\left(\overline{m} - k_1\right)}{K_{-1}} = s_i^* < k_i, \tag{5}$$

from (1). This guarantees that h^t is also a public history if $\overline{m} < K$ for the following reasons. 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, as has just been demonstrated, if all firms do not set such a common price, then the sales of the highest-priced firms will not exceed their trigger levels and their lower-priced rivals will supply their full capacities.[7] 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 strictly below the highest in the market.[8]

This public information allows firms to make inferences about the behaviour of their rivals. In particular, each firm knows that all firms' sales will exceed their trigger levels, such that $y = \overline{y}$, if $p_j = p \leq 1$ for all j and if $m > m^*\left(k_1, \overline{m}\right)$; otherwise, at least one firm's sales will not exceed its trigger level, so $y = \underline{y}$. It follows from this that if $\underline{m} > m^*\left(k_1, \overline{m}\right)$, then there is perfect monitoring of a strategy in which all firms set a common collusive price. This is due to the fact that each firm would only receive sales below its trigger level, if it has been undercut. In contrast, there is imperfect monitoring if $\underline{m} \leq m^*\left(k_1, \overline{m}\right)$. The reason can be understood by considering $\Pr\left(\underline{y}|p_i, \mathbf{p}_{-i}\right)$ which denotes the probability of observing $\underline{y}$ if firm i sets p_i and its

[7]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.

[8]Notice that if the trigger levels were below s_i^* for all i, then a firm that supplies its full capacity would be uncertain as to whether at least one rival has received sales below its trigger level. So, any such trigger levels would not generate a public history. In contrast, trigger levels above s_i^* for all i would also ensure that h^t is a public history. However, such trigger levels have the strange feature that firms can receive a bad signal $\underline{y}$, even when all firms know that they have set a common price. Consequently, such alternative trigger levels are inferior to s_i^*: they raise the critical discount factor and lower equilibrium profits compared to the main analysis.

rivals price according to $\mathbf{p}_{-i}$. For the case of $\underline{m} \leq m^*(k_1, \overline{m})$:

$$\Pr\left(\underline{y}|p_i, \mathbf{p}_{-i}\right) = \begin{cases} \int_{\underline{m}}^{\min\{m^*(k_1,\overline{m}),\overline{m}\}} g(m)\, dm \in [0,1] & \text{if } p_j = p\ \forall j \\ 1 & \text{otherwise.} \end{cases} \tag{6}$$

This says that a firm's sales can be below its trigger level if the realisation of market demand is sufficiently low, even when firms set a common price. Thus, for such an outcome 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.

Proposition 1 finds the conditions for perfect and imperfect monitoring in terms of the maximum market demand, holding the minimum market demand constant.

Proposition 1. *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 $\overline{m} \in (\underline{m}, \underline{x}(k_1))$, then monitoring is perfect. Otherwise, there is imperfect monitoring.*

Monitoring is perfect if the fluctuations in market demand are sufficiently small, otherwise there is imperfect monitoring. The critical level is strictly increasing in the capacity of the smallest firm, k_1. The reason is that deviations by the smallest firm are most difficult to detect, from (5). Furthermore, it follows from this logic that detecting a deviation is less difficult when the smallest firm is larger. Consequently, if it is just possible for a firm to infer that the smallest firm has not deviated for a given level of $\overline{m}$, then it is also possible for the same level of $\overline{m}$ if the smallest firm has more capacity. This implies that deviations can be detected perfectly for a wider range of fluctuations in market demand if the smallest firm is larger.

Finally, we have so far considered the public information that firms can infer from their privately observed sales. Before moving on, we should discuss two possible scenarios in which a firm's sales can provide it with private information that is not common knowledge among all firms. In either case though, it should be noted that any such private information is not payoff relevant if rivals follow public strategies. Thus, it will not be possible for a firm to use its private information to gain by deviating from an equilibrium in public strategies. The first case is when a firm knows for sure that it has been undercut. This occurs if firm i's sales are inconsistent with all firms setting a common price, $s_i < \frac{k_i}{K}\underline{m}$ for some i. Such information is not common knowledge if monitoring is imperfect, because the deviants $j \neq i$ would be unaware of the specific levels of its rivals' sales: they simply knows that at least one rival's sales are below its trigger level. The second case is when the smallest firm knows for sure that all firms have set a common price, but its rivals $i > 1$ are uncertain as to whether the smallest firm has undercut them. This may occur if firm 1 is strictly the smallest firm and if fluctuations in market demand are not large, such that $\overline{m} < K$. In such a case, the highest possible sales of the smallest firm if it is undercut are $\frac{k_1}{K_{-2}}(\overline{m} - k_2) < s_1^*$. Thus, if the smallest firm's sales are below its trigger level, s_1^*, yet above $\frac{k_1}{K_{-2}}(\overline{m} - k_2)$, then it knows for sure that all firms have set a common price. Nevertheless, the fact that its sales are below its trigger level will inform the smallest firm that its rivals' sales are also below their trigger levels.

3.2 Optimal collusive equilibrium profits

We now solve the repeated game restricting 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). We solve the model using two seemingly different approaches. First, in the main body of the paper, we restrict attention to a particular class of PPE 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 in a collusive period. We formally describe the strategy profile for this approach below and refer to it as trigger-sales strategies. Second, given that restricting attention to trigger-sales strategies leaves open the question of whether there are other PPE with higher profits, we solve for the set of PPE in appendix B using the techniques of Abreu *et al.* (1986, 1990). This appendix shows that trigger-sales strategies are optimal equilibrium strategies in that they support the maximal PPE payoffs and they generate the lowest critical discount factor.

Trigger-sales strategies are formally defined as follows. There are 'collusive phases' and 'punishment phases'. Suppose period t is in a collusive phase. In any such period, a firm should set the collusive price, $p^c > \underline{p}$. If $y_t = \overline{y}$, 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 collusive phase.

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 $k_i V^p$, if all firms follow trigger-sales strategies, then:

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

for all i, where $\Pr\left(\underline{y}|p^c\right) = G\left(m^*\left(k_1, \overline{m}\right)\right)$ from (6). Substituting $k_i V^p$ into $k_i V^c$ and solving yields:

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

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

The profile of 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 under trigger-sales strategies is not sustainable 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 $\Pr\left(\underline{y}|p_i, p^c\right) = 1$ for any $p_i \neq p^c$ from (6). The incentive compatibility constraint (ICC) for firm i is as follows:

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

This says that firm i will not deviate in any period in a collusive phase if it cannot gain by marginally undercutting $p^c > \underline{p}$ to supply its full capacity k_i.[9] Note that (8) is never satisfied when the maximum market demand is greater than total capacity, $\overline{m} \geq K$, as then $G\left(m^*\left(k_1,\overline{m}\right)\right) = 1$, from (6). Thus, collusion under trigger-sales strategies is not sustainable if $\overline{m} \geq K$, so we can henceforth focus on the case where $\overline{m} < K$.

Substituting $k_i V^p$ and $k_i V^c$ into (8), then rearranging yields:

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

It follows from the fact that (9) 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 the left-hand side of (9) is less than the expression in square brackets on the right-hand side, such that (9) can only hold

[9] It follows from Lemma 1 that firm i's optimal deviation is to undercut $p^c > \underline{p}$ for all i. Furthermore, note that $p^c > \underline{p}$ is a necessary condition for firms to attain collusive profits per-period greater than the static Nash equilibrium profit.

if both are positive, since $\delta^T \in (0, \delta]$ for all $T \in [0, \infty)$. Thus, similar to Green and Porter (1984) and Tirole (1988), it follows from (9) that there are three necessary conditions for the profile of triggers-sales strategies to be a PPE. First, the length of the punishment phase must be sufficiently long, where the critical duration, denoted $T^* (k_1, k_n, p^c)$, is implicitly defined by the level of T where (9) holds with equality.[10] Second, firms must also be sufficiently patient, such that:

$$\delta \geq \frac{(K - \widehat{m})\, p^c}{(1 - G(m^* (k_1, \overline{m})))\, K \left(p^c - \underline{p}\right)} \equiv \delta^* (k_1, k_n, p^c), \tag{10}$$

in which case the left-hand side of (9) is positive, such that the ICC holds as $T \to \infty$. This implies that if firms are not sufficiently patient, then even a punishment phase that lasts an infinite number of periods is insufficient to outweigh the short-term benefit from deviating. Furthermore, the critical punishment phase duration $T^* (k_1, k_n, p^c) < \infty$ for any $\delta > \delta^* (k_1, k_n, p^c)$ and $T^* (k_1, k_n, p^c) \to \infty$ if $\delta = \delta^* (k_1, k_n, p^c)$. Third, the probability of receiving a bad signal must be sufficiently low, where:

$$G (m^* (k_1, \overline{m})) < 1 - \frac{(K - \widehat{m})\, p^c}{K \left(p^c - \underline{p}\right)}, \tag{11}$$

such that the expression in square brackets in (9) is positive. Note that (11) ensures $\delta^* (k_1, k_n, p^c) < 1$, 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.

Proposition 2 solves for the optimal PPE profits under trigger-sales strategies. We refer to this as 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, $\overline{x} (k_1, k_n) \in (\underline{x} (k_1), K)$, that solves $G (m^* (k_1, \overline{x} (k_1, k_n))) = \frac{K_{-n}}{K} < 1$, such that, if $\overline{m} \in [\underline{x} (k_1), \overline{x} (k_1, k_n))$ and if $\delta \geq \delta^* (k_1, k_n) \equiv \frac{1}{1 - G(m^*(k_1, \overline{m}))} \frac{k_n}{K} \in \left(\frac{k_n}{K}, 1\right)$, then firm i's optimal PPE profits under trigger-sales strategies are:*

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

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

This says that, if the necessary conditions (10) and (11) are satisfied, then the optimal PPE profits under trigger-sales strategies have the firms set the monopoly price during a collusive

[10] Although $T^* (k_1, k_n, p^c)$ may not be an integer, the expected length of the punishment phase could equal this length if there were some optimally set publically observable randomisation device that varied the length of punishment phases.

phase and the optimal punishment phase duration is $T^*(k_1, k_n, p^m)$ such that the ICC (9) is binding with no slack. Despite the fact that firms set the monopoly price during collusive phases, the sum of such equilibrium profits is below the monopoly profit, because punishment phases occur on the equilibrium path. 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 trigger-sales strategies, with firms setting the monopoly price during collusive phases, is a PPE or it is not an equilibrium strategy profile at any collusive price.

Next, we turn our attention to the case of perfect monitoring, where $G(m^*(k_1, \overline{m})) = 0$. In this case, any PPE is also a subgame perfect Nash equilibrium (SPNE) and, as we explain further in a moment, we can easily generate the optimal SPNE profits by letting the punishment phase last an infinite number of periods, such that $T \to \infty$. Thus, the optimal SPNE profits are summarised by the following corollary. We refer to this as collusion under perfect monitoring.

Corrolary 1. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} < \overline{m} \leq \underline{x}(k_1)$, there exists a unique discount factor $\underline{\delta}^*(k_n) \equiv \frac{k_n}{K} \in \left(0, \frac{\underline{m}}{K}\right]$, such that if $\delta \geq \underline{\delta}^*(k_n)$, then firm i's optimal SPNE profits under trigger-sales strategies are:*

$$k_i V^* = \frac{k_i}{K} \widehat{m} > \pi_i^N(k_n) \ \forall \ i.$$

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

The firms divide the monopoly profits between them if they are sufficiently patient. The equilibrium profits are highest and the critical discount factor is lowest when the firms set the monopoly price. The critical discount factor is the same as in Compte *et al.* (2002) and it also coincides with the lowest possible discount factor that sustains collusion given the proportional allocation rule. The reason is that, as showed by Lambson (1994), the optimal punishments under the proportional allocation rule are such that the largest firm receives the stream of profits from its minimax strategy. In our setting, this is the case as $T \to \infty$, because in each period of the punishment phase the firms receive the static Nash equilibrium profits, which for the largest firm is equivalent to its minimax payoff. Thus, it is not possible to lower the critical discount factor below this level, given the proportional allocation rule.

These results are brought together in Figure 1. It highlights that the critical discount factor under imperfect monitoring, $\delta^*(k_1, k_n)$, converges to the critical level under perfect monitoring, $\underline{\delta}^*(k_n)$, at $\overline{m} = \underline{x}(k_1)$, but it is strictly above $\underline{\delta}^*(k_n)$ for any higher maximum market demand. The optimal equilibrium profits under trigger-sales strategies, $k_i V^*$, equal the monopoly level

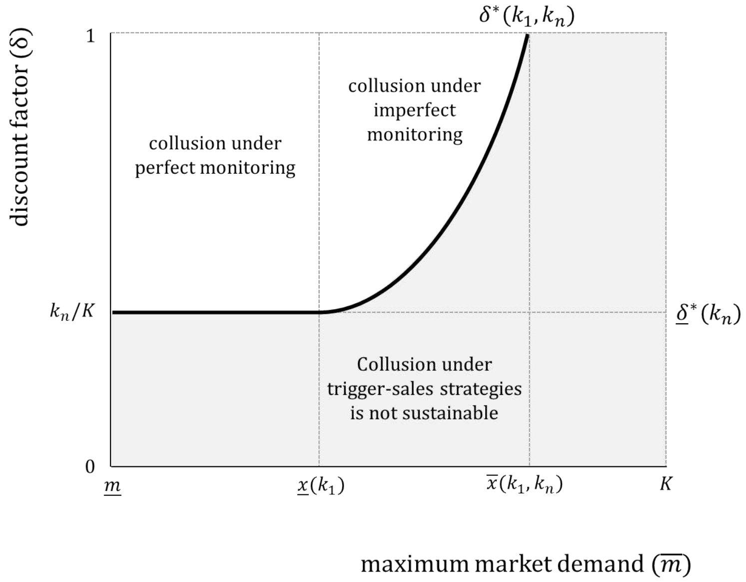

Figure 1: Parameter space of collusion

at or below $\underline{x}(k_1)$ for all $\delta \geq \underline{\delta}^*(k_n)$, and they equal the static Nash equilibrium profits at $\overline{m} = \overline{x}(k_1, k_n)$ where $\delta^*(k_1, k_n) = 1$. Furthermore, assuming a mean-preserving spread, they are strictly decreasing in $\overline{m}$ between $\underline{x}(k_1)$ and $\overline{x}(k_1, k_n)$. Before moving on, the reader may wish to check appendix B, where we use the techniques of Abreu *et al.* (1986, 1990) to show that the optimal equilibrium profits under trigger-sales strategies are the maximal PPE payoffs. Thus, henceforth we refer to them as the optimal equilibrium profits.

3.3 Comparing capacity distributions

We want to analyse the effects of mergers in our setting. Before doing so, it is helpful to get a clear understanding of how the capacity distribution affects collusion by analysing changes in the capacity distribution, when the number of firms and the total capacity are held 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. Thus, throughout this subsection, we assume 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 firm or the largest firm.

Proposition 3 analyses the effects of reallocating capacity among the firms on the critical discount factor, which ensures that the ICC holds as $T \to \infty$.

Proposition 3. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} < K$:*
i) if $\overline{m} \in [\underline{m}, \underline{x}(k_1))$, then $\underline{\delta}^*(k_n)$ is strictly increasing in the capacity of the largest firm, k_n;
ii) if $\overline{m} \in [\underline{x}(k_1), \overline{x}(k_1, k_n))$, then $\delta^*(k_1, k_n)$ is strictly increasing in the capacity of the largest firm, k_n, and strictly decreasing in the capacity of the smallest firm, k_1.

Consistent with Compte *et al.* (2002), increasing the size of the largest firm hinders collusion. The reason is that a punishment that lasts an infinite number of periods is weaker when the largest firm is larger, because the static Nash equilibrium profits increase for each firm, so the critical discount factor rises. In contrast to Compte *et al.* (2002), increasing the size the smallest firm facilitates collusion. This is due to the fact that firms can monitor an agreement more successfully when the smallest firm is larger, because it is less likely that firms' sales will be below their trigger levels when they set a common price. This does not affect the critical discount factor under perfect monitoring but, as we saw in section 3.1, it does imply that monitoring is perfect for a wider range of fluctuations in market demand. Under imperfect monitoring, it is less likely that a collusive phase will switch to a punishment phase on the equilibrium path when the smallest firm is larger. Consequently, the expected future profits from collusion are higher than before, which implies that a punishment that lasts an infinite number of periods is relatively harsher, so the critical discount factor falls.[11]

Next, we analyse the effects of reallocating capacity among the firms on the optimal equilibrium profits. For convenience, we transform such profits to an average price and compare it to the average static Nash equilibrium price, given by $\widehat{p}^N(k_n) \equiv \frac{K}{\widehat{m}} \frac{(\widehat{m} - K_{-n})}{k_n}$ for all $\overline{m} < K$. The average price of the optimal equilibrium profits under perfect monitoring is independent of the capacity distribution, since firms set p^m in each period if they are sufficiently patient. So, Proposition 4 investigates the effect of reallocating capacity on the average price associated with the optimal equilibrium profits under imperfect monitoring. We refer to this as the optimal average price, and this is given by $\widehat{p}^c(k_1, \overline{m}) \equiv \frac{K}{\widehat{m}} V^*$ in expectation, where $\widehat{p}^N(k_n) < \widehat{p}^c(k_1, \overline{m}) < p^m$.

[11]Both results are consistent with the findings of Vasconcelos (2005). The underlying incentives for his results are very different to ours though, as they rely on capacities affecting marginal costs in a setting of perfect observability.

Proposition 4. *For any given $n \geq 2$, $K_{-1} \leq \underline{m} < \underline{x}(k_1) < \overline{m} < \overline{x}(k_1, k_n)$ and $\delta \geq \delta^*(k_1, k_n)$, the optimal average price, $\widehat{p}^c(k_1, \overline{m})$, is strictly increasing in the capacity of the smallest firm, k_1.*

The optimal average price is increasing in the capacity of the smallest firm for two reasons. 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 phases are less likely to switch to punishment phases than before. Second, such an increase in profits also introduces slack into the ICC. Consequently, the optimal punishment phase duration shortens to ensure that the ICC is binding with no slack, which increases equilibrium profits further.

Surprisingly, the optimal average price is independent of the capacity of the largest firm, when the capacity of the smallest firm is held constant. This is due to the fact that there are two effects that perfectly offset each other. The first effect is that an increase in the capacity of the largest firm raises profits on the equilibrium path, other things equal, because the static Nash equilibrium profits of each firm are greater than before. However, this also tightens the ICC, so the second effect is that the optimal punishment phase duration lengthens to ensure that the ICC is binding with no slack. This second effect cancels out the first, implying the size of the largest firm has no effect on the optimal average price.

It follows from the above analysis that asymmetries hinder collusion under perfect and imperfect monitoring. In summary, Proposition 3 implies that the parameter space of collusion is greatest when firms' capacities are symmetric, because the punishment is harshest when the largest firm is as small as possible, and since monitoring is most successful when the smallest firm is as large as possible. The latter also implies that the optimal average price is higher if firms are symmetric from Proposition 4. Furthermore, since the optimal average price is independent of the size of the largest firm, it follows that the optimal average price is highest for a symmetric duopoly and that, for example, it would be higher for a symmetric triopoly than an asymmetric duopoly with $k_1 < K/3$.

Despite the fact that symmetry is ideal for collusion, Proposition 5 next shows that the competitive prices of asymmetric capacity distributions can be higher than the collusive prices of less asymmetric capacity distributions. To prove this result, we compare the optimal average price of one distribution, (k_1, k_n), to the static Nash equilibrium average price of another, denoted (k_1', k_n').

Proposition 5. *For any given $n \geq 2$, $K_{-1} \leq \underline{m} < K$ and $\delta \geq \delta^* (k_1, k_n)$, if $k'_n > k_n$, such that $k'_1 \leq k_1$, then there exists a unique level of market demand, $\overline{x}(k_1, k'_n) \in (\underline{x}(k_1), \overline{x}(k_1, k_n))$, that solves $G(m^*(k_1, \overline{x}(k_1, k'_n))) = 1 - \frac{k'_n}{K} < 1$, where if $\overline{m} \in (\overline{x}(k_1, k'_n), \overline{x}(k_1, k_n))$, then the static Nash equilibrium average price of (k'_1, k'_n) is greater than the optimal average price of (k_1, k_n), $\widehat{p}^N(k'_n) > \widehat{p}^c(k_1, \overline{m})$.*

This says that if fluctuations in market demand are sufficiently large, then the competitive prices of asymmetric capacity distributions are higher than the collusive prices of less asymmetric capacity distributions. The intuition is that an increase in the maximum market demand raises the likelihood that firms' sales will be below their trigger levels when firms set a common price. Thus, punishment periods are expected to occur on the equilibrium path more often than before. As a result, the optimal average price of (k_1, k_n) falls towards its corresponding static Nash equilibrium average price as the maximum market demand increases towards the critical level $\overline{x}(k_1, k_n)$. Yet, the average static Nash equilibrium price is strictly increasing in the capacity of the largest firm, k_n. So, consider an alternative distribution (k'_1, k'_n) that is more asymmetric than the original in the sense that $k'_n > k_n$ (so $k'_1 \leq k_1$). It follows that if the maximum market demand is sufficiently close to $\overline{x}(k_1, k_n)$, then (k'_1, k'_n) will have a higher average static Nash equilibrium price than the optimal average price of (k_1, k_n). The critical level $\overline{x}(k_1, k'_n)$ is the point at which $\widehat{p}^c(k_1, \overline{m}) = \widehat{p}^N(k'_n)$ for all $\delta \geq \delta^*(k_1, k_n)$. Furthermore, the condition that the maximum market demand exceeds $\overline{x}(k_1, k'_n)$ guarantees that collusion under trigger-sales strategies is not sustainable for (k'_1, k'_n). This is due to the fact that collusion under trigger-sales strategies requires that the maximum market demand is below $\overline{x}(k'_1, k'_n)$, but this is contradicted since $\overline{x}(k'_1, k'_n) \leq \overline{x}(k_1, k'_n) < \overline{m}$ for all $k'_1 \leq k_1$.

4 Monitoring and Mergers

We now use our equilibrium analysis to draw implications for merger policy. A merger in our framework amounts to the merging firms consolidating their capacity, so we draw on the analysis of section 3.3. However, the following analysis differs to section 3.3 in that a merger will reduce the numbers of firms and can increase both the size of the smallest and largest firm at the same time. We also consider the firms' incentives to merge and analyse the welfare effects. Following the terminology of Farrell and Shapiro (1990), we henceforth refer to the merging firms as insiders and those not involved in the merger as outsiders. We say that a merger is privately optimal if the sum of insiders' profits post-merger is *strictly* greater than the sum of their profits pre-merger. Finally, with respect to welfare, we focus on the effects of mergers on consumer surplus for two reasons. First, expected total welfare is independent of the capacity distribution, so mergers do not affect it. Second, ensuring consumer surplus does not fall post-merger is commonly perceived to be the main objective of merger control (see Lyons, 2002).

We are particularly interested in comparing unilateral and coordinated effects in our framework. Such effects have been considered independently of each other in the previous literature. For instance, Compte *et al.* (2002) and Vasconcelos (2005) analyse models where firms can share the monopoly profits if they are sufficiently patient, and as a consequence they both focus solely on the coordinated effects of mergers on the critical discount factor. Bos and Harrington (2010) analyse the coordinated effects of mergers on the price of a cartel that does not encompass all firms in the market. They find that mergers that increase the capacity controlled by the cartel can raise the cartel price towards the monopoly level. However, in contrast to our model, they restrict attention to capacity distributions for which there is a unique pure strategy static Nash equilibrium price equal to marginal cost, so unilateral effects are not an issue. Thus, such papers are consistent with the conventional wisdom that collusive post-merger outcomes are worse than non-collusive outcomes. In our setting, collusion under imperfect monitoring does not enable firms to share the monopoly profits in every period. As a result, the conventional wisdom will not hold, if competition in the noncollusive outcome is weak and hence prices are high. Below we explore for which mergers the conventional wisdom does not hold and discuss the implications.

Before doing so, we return to discuss the generality of Assumption 1 in the context of mergers. Recall that this restricts the smallest firm's capacity from being too small. In this context, the necessary condition for Assumption 1 to hold is stricter than before, due to the fact that we are only interested in mergers with at least 2 firms post-merger. Thus, there must be $n > 2$ firms pre-merger, in which case the minimum market demand must be greater than $\frac{n-1}{n}$ of the total capacity, $\underline{m} \geq \left(\frac{n-1}{n}\right) K$. The reason is that the smallest firm's capacity can be no larger pre-merger than for a symmetric capacity distribution, $k_1 = K/n$. This implies that Assumption 1 is more restrictive when there are more firms in the market pre-merger. However, Assumption 1 is likely to hold for a large number of mergers that raise concerns of collusion pre- and/or post-merger. For example, Davies *et al.* (2011) found that the European Commission was concerned about post-merger collusion between duopolists in all bar one of the cases that were seriously investigated for tacit collusion between 1990 and 2004 (see also Davies and Olczak, 2010). Consequently, there were usually only three main players pre-merger, such that the necessary condition is $\underline{m} \geq \frac{2}{3}K$. Moreover, given collusion is likely to occur when asymmetries are small, Assumption 1 is especially unrestrictive for the cases in which there is collusion pre-merger.

4.1 The competitive effects of mergers

In this subsection, we analyse mergers that change 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 mergers on the parameter space of collusion. A merger that increases the size of both the smallest and the largest firm will have a mix of the effects described here, and this issue is discussed more in the following subsections. All other mergers will not affect the equilibrium analysis.

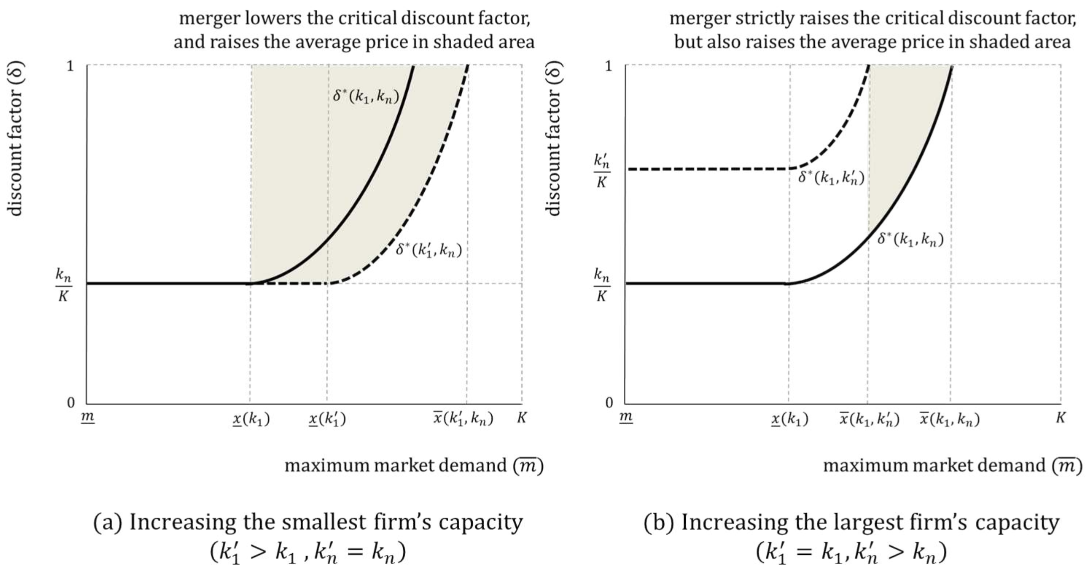

Figure 2: The effects of mergers on the parameter space of collusion

A merger that only increases the size of the smallest firm will facilitate collusion. It follows from Proposition 3 that the parameter space of collusion will expand and Proposition 4 implies that the average price may also rise post-merger. More specifically, the price will rise if collusion is sustainable to some degree post-merger and if either the pre-merger outcome is noncollusive or collusion pre-merger is less than perfect. Furthermore, similar to other models where collusion pre-merger implies firms share the monopoly profits, such a merger has no effect on the average price if there is collusion under perfect monitoring pre- and post-merger. Thus, the complete parameter space for which such a merger raises the average price is illustrated in the shaded area of Figure 2(a). Any such merger that raises the average price is privately optimal and it also strictly increases the profits of the outsiders. This follows since the present discounted value of profits for any set of firms M, given an average price $\widehat{p}$, is $\sum_{i \in M} k_i \frac{\widehat{m}}{K} \left(\frac{\widehat{p}}{1-\delta}\right)$, where this is higher post-merger only if the average price rises. As a consequence, such a merger will also lower consumer surplus, since the expected consumer surplus per unit is $1 - \widehat{p}$. Thus, consistent with the conventional wisdom, any collusive outcome that has been facilitated by a merger that only increases the size of the smallest firm is worse than the pre-merger outcome.[12]

A merger that only increases the size of the largest firm will hinder collusion. Proposition 3 implies that the parameter space of collusion will contract. Yet, in contrast to the conventional wisdom, it follows from Proposition 5 that such a merger may actually increase prices if collusion under imperfect monitoring is destabilised by the merger and if fluctuations in market demand are sufficiently large. The parameter space for which such a merger raises the average price is illustrated in the shaded area of Figure 2(b). Outside of this shaded area, the conventional wisdom holds if collusion under imperfect monitoring is destabilised post-merger. Nevertheless, our model suggests that it is only in the insiders' interests to propose such a merger, if the average price rises post-merger. This follows since such a merger is privately optimal for any set of firms M if $\sum_{i \in M} k_i \frac{\widehat{m}}{K} \left(\frac{\widehat{p}^N\left(k_n'\right)}{1-\delta}\right) > \sum_{i \in M} k_i \frac{\widehat{m}}{K} \left(\frac{\widehat{p}^c(k_1, \overline{m})}{1-\delta}\right)$. Consequently, the condition that guarantees the insiders' profits increase post-merger also ensures that the average price rises post-merger. Moreover, the same condition also guarantees that such a merger increases the profits of the outsiders and lowers consumer surplus. Finally, it should be noted for completeness that if there is no collusion pre- or post-merger, then such a merger would increase the average price post-merger through unilateral effects.

4.2 An example

We now complement our general results by analysing an example. We do this for three reasons. First, we wish to show that average prices can be substantially lower (and hence consumer surplus can be higher) for collusive merger outcomes than for noncollusive outcomes. Second, we want to

[12]It also follows from this that larger firms $i > 1$ can actually increase their profits by divesting capacity to the smallest firm, so that monitoring is easier. Such divestments are not unheard of in actual merger cases (see Compte *et al.*, 2002; and Davies and Olczak, 2010).

analyse the effects of a merger that increases the size of the smallest and largest firm at the same time. Third, we wish to explore the effects of divestment remedies on post-merger consumer surplus.

With these objectives in mind, we consider an example where total capacity is $K = 100$ and suppose that this is divisible into 6 equal sized parts. There is an asymmetric triopoly pre-merger, denoted (1/6, 2/6, 3/6), where firm 1 has 1/6 of this capacity, firm 2 has 2/6 and firm 3 has 3/6. We then consider three alternative merger outcomes: a symmetric duopoly, (3/6, 3/6); an asymmetric duopoly, (2/6, 4/6); and a very asymmetric duopoly, (1/6, 5/6). These three post-merger outcomes can be thought of in two ways. First, each outcome could arise directly from a merger. For example, (3/6, 3/6) can result from a merger between firms 1 and 2; (2/6, 4/6) from a merger between firms 1 and 3; and (1/6, 5/6) from a merger between firms 2 and 3. Alternatively, one outcome could arise directly from a merger and the other two could result from divestment remedies of this merger. For example, if (1/6, 5/6) is created from a merger between firms 2 and 3, then (2/6, 4/6) and (3/6, 3/6) could result from a remedy of that merger, where capacity of the merged entity is divested to the outsider to remedy concerns of unilateral effects.

We analyse the effects of such mergers on the expected consumer surplus per unit of the most profitable equilibrium, denoted $CS\left(\widehat{p}^{*}\right) \equiv 1 - \widehat{p}^{*}$. The preceding analysis implies that $\widehat{p}^{*}$ is the static Nash equilibrium average price if collusion under trigger-sales strategies is not sustainable, otherwise it is either the optimal average price or the monopoly price. Figure 3 plots $CS\left(\widehat{p}^{*}\right)$ as a function of $\Delta m \equiv \frac{\overline{m}-\underline{m}}{\widehat{m}}$ for the various scenarios, assuming demand is drawn from a uniform distribution. Parameter values are chosen such that $\widehat{m} = 92$ for all Δm and that $K_{-1} \leq \frac{5}{6}(100) \leq \underline{m} \leq \overline{m} \leq K = 100$, so Assumption 1 holds. We let $\delta \to 1$ such that collusion under trigger-sales strategies is not sustainable only if $\overline{m} \geq \overline{x}\left(k_1, k_n\right)$. Finally, the analysis above implies that each merger is privately optimal whenever $CS\left(\widehat{p}^{*}\right)$ is strictly lower post-merger than pre-merger.

Each of the plotted lines in Figure 3 has a similar shape. If $CS\left(\widehat{p}^{*}\right) = 0$, then monitoring is perfect and the average price is p^{m}. If $CS\left(\widehat{p}^{*}\right)$ is upward-sloping, then there is imperfect monitoring and the optimal average price is strictly decreasing in Δm. If $CS\left(\widehat{p}^{*}\right)$ is positive and constant, then it is the expected consumer surplus per unit of the static Nash equilibrium, because the outcome is noncollusive. Furthermore, note that comparing (1/6, 2/6, 3/6) with (3/6, 3/6) for increasing values of Δm in Figure 3 is consistent with moving horizontally from left to right in Figure 2(a) for $\delta \to 1$, because only the capacity of the smallest firm changes.

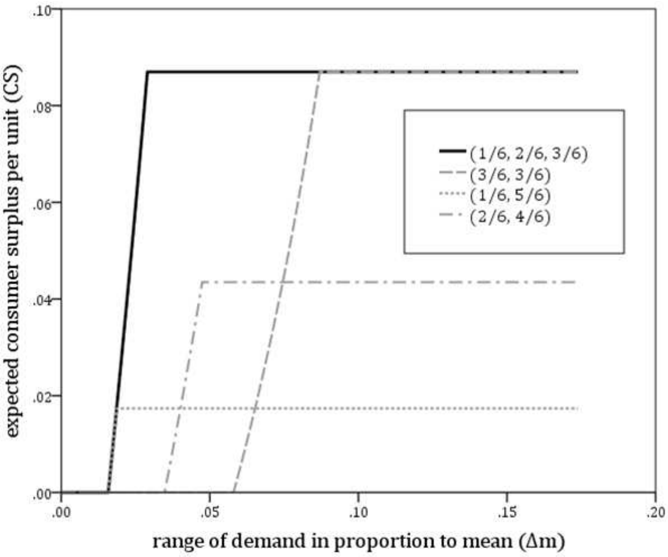

Figure 3: $G(m) = \frac{m-\underline{m}}{\overline{m}-\underline{m}}$, $\widehat{m} = 92$, $K_{-1} \leq \frac{5}{6}(100) < 100 = K$, and $\delta \rightarrow 1$

Likewise, comparing (1/6, 2/6, 3/6) with (1/6, 5/6) for increasing values of Δm is consistent with moving horizontally from left to right in Figure 2(b) for $\delta \rightarrow 1$, because only the capacity of the largest firm changes. For (2/6, 4/6), both the capacities of the smallest and the largest firms are larger than compared with (1/6, 2/6, 3/6).[13]

First, consider the merger that creates the very asymmetric duopoly (1/6, 5/6). This merger only increases the size of the largest firm, so it can cause unilateral effects even when it destabilises collusion. This happens at around $\Delta m = 0.025$ where, in contrast to the conventional wisdom, the pre-merger expected consumer surplus per unit is approximately four times greater than post-merger, despite the fact that there is collusion pre- but not post-merger. Next, consider the merger that creates a symmetric duopoly (3/6, 3/6). This merger only increases the size of the smallest firm, so it can only cause coordinated effects. In the worst cases, and consistent with the conventional wisdom, it can substantially lower consumer surplus per unit from 9% of total welfare per unit to 0%. This happens approximately over the range $0.02 < \Delta m < 0.06$. However, the difference between the pre- and post-merger expected consumer surplus per unit becomes smaller as Δm increases towards 0.09 and collusion post-merger becomes increasingly difficult to

[13] If $\delta < 1$, then there would be a discontinuity in each of the $CS(\widehat{p}^*)$ lines at the threshold of Δm where the outcome is noncollusive. At this thresholds, a line would jump up to the expected consumer surplus per unit of the static Nash equilibrium, such that this level of consumer surplus extends for lower levels of Δm than in Figure 3.

monitor. Finally, consider the merger that creates (2/6, 4/6) where the size of the smallest and the largest firm has increased. This merger has a mix of the two effects just discussed. It causes coordinated effects at approximately $\Delta m < 0.04$, otherwise it causes unilateral effects.

Figure 3 also shows that our analysis has important consequences for the appropriate divestment remedies. For instance, if the logic of the conventional wisdom were followed, then a collusive post-merger outcome would be remedied by a divestment that increases asymmetries to ensure that collusion is destabilised. However, our model shows that this will not always result in a less harmful outcome. For example, consider the merger that creates (3/6, 3/6). Post-merger asymmetries between the firms can be created by divesting some of the capacity of the merged entity to the outsider to create either (2/6, 4/6) or (1/6, 5/6). Notice that, consistent with the conventional wisdom, at around $\Delta m = 0.06$ such remedies do indeed raise the post-merger expected consumer surplus per unit compared to (3/6, 3/6). However, in contrast to the conventional wisdom, they do the opposite close to $\Delta m = 0.09$. Here, the expected consumer surplus per unit for (3/6, 3/6) is approximately four times greater than for (1/6, 5/6) and two times greater than for (2/6, 4/6), despite the fact that there is collusion only in the symmetric duopoly. Of course, the reverse is also true: for the mergers that create the asymmetric duopolies, (1/6, 5/6) or (2/6, 4/6), divesting capacity from the merged entities to the outsiders to create (3/6, 3/6) would lead to a less harmful post-merger outcome at around $\Delta m = 0.09$, even though it facilitates collusion. Such remedies would not be implemented if the conventional wisdom were followed.

4.3 Relation to European merger decisions

Finally for this section, we briefly relate our analysis to two merger decisions by the European Commission that at first glance seem questionable but can make sense in terms of our model. The first concerns the merger between Linde and BOC where the Commission used both the unilateral and coordinated effects theories of harm simultaneously to justify an intervention.[14] Such interventions have arisen since 2004, when a key change to the European Commission Merger Regulation made clear that interventions for unilateral effects below the level of single dominance were possible (see Röller and Mano, 2006).[15] The intervention in question concerned the global helium wholesale market, where pre-merger there were three relatively symmetric firms

[14] Linde/BOC, Commission decision of 6 June 2006, case number COMP/M.4141, paragraphs 150-192.

[15] For a similar intervention in a different merger, see T.Mobile Austria/Tele.ring, Commission decision of 26 April 2006, case number COMP/M.3916, paragraphs 50-129.

and two smaller firms: the largest firms were Air Products, Praxair and BOC, whose market shares were approximately 25%, and the smaller firms were Air Liquide and Linde, who supplied about 15% and 5% of market, respectively. Thus, the merger between BOC and the recent entrant Linde had the potential to remove the smallest firm from the market and increase the size of the largest firm marginally at the same time.

The Commission's decision that this merger could lead to either unilateral or coordinated effects goes against the view that such effects should be mutually exclusive for a given market. This view is based on the fact that theory has consistently shown that asymmetries increase the likelihood of unilateral effects but decrease the likelihood of coordinated effects (see Kühn, 2001, p.13-15). In contrast to this view but consistent with the approach of the Commission in *Linde/BOC*, it can be justifiable in our framework to intervene in a merger on the grounds of coordinated and unilateral effects simultaneously, if both the smallest and the largest firms increase as a result of the merger. For example, consider the merger that creates (2/6, 4/6) in the scenario of section 4.2, where the capacities of the smallest and largest firms have increased compared to pre-merger (1/6, 2/6, 3/6). As we noted in the previous subsection, such a merger can lead to coordinated effects if the fluctuations are small, due to the fact that the smallest firm is larger post-merger. Yet, if for some reason the firms are unable to coordinate on any collusive equilibrium, then the merger could also lead to unilateral effects over the same range of fluctuations, because the largest firm is larger post-merger. This can be seen in Figure 1 by extending back to the vertical axis the horizontal lines, which relate to the expected consumer surplus of the static Nash equilibrium. Then at around $\Delta m = 0.025$ the merger could cause unilateral or coordinated effects, depending on the equilibrium selection of the firms.

The second decision is the merger between Nestlé and Perrier in the French bottled mineral water market, which has been the subject of much discussion.[16] Pre-merger there were three main players in the market: according to various sources (see Compte *et al.*, 2002, p.19-23, and Motta, 2004, p.279-286), Perrier had a market share of about 30%, BSN had 25% and Nestlé between 20-25%, with the remaining sales made up from a competitive fringe. Nestlé's acquistion of Perrier would have created a firm with a market share over 50%, and the Commission subsequently made clear that they would have been concerned over the merged entity's dominant position. Anticipating this, Nestlé had also agreed that, upon successful completion of the merger, it would transfer some of the acquired assets to BSN. However, this transfer would have effectively created a symmetric duopoly post-merger, where each firm had a 38% share of the market, so it

[16] Nestlé/Perrier, Commission decision of 22 July 1992, case number IV/M.190

was rejected by the Commission on the grounds of coordinated effects. Therefore, an additional divestment was offered by the merging parties in which some more capacity would be sold to a suitable entrant, thereby ensuring the post-merger market had the same number of firms as pre-merger but with reduced asymmetries. The Commission allowed the merger to proceed subject to this additional divestment.

Focussing on only the coordinated effects of the merger, Compte *et al.* (2002) showed that collusion would have been easier to sustain under the accepted asymmetric triopoly market structure than had the merger proceeded unconditionally and without the transfer to BSN, which would have created a very asymmetric duopoly post-merger. They argued that, in its analysis of the appropriate remedy, the Commission had focused too little on the degree of asymmetry of the post-merger market structure and too much on the number of active firms post-merger. While they convincingly show that this was the case, there are two reasons in our framework as to why restoring the number of firms to the pre-merger level can be beneficial, even when this is at the expense of reduced asymmetries. First, if the newly created firm is the smallest, then it will reduce the degree to which firms can monitor each other. Second, if the new firm is created out of the capacity of the largest firm and if the post-merger outcome is noncollusive, then competition will be more intense post-merger.

This second effect is especially important when deciding between alternative outcomes where either unilateral or coordinated effects are possible. To illustrate its importance, consider the distributions (1/6, 5/6) and (1/6, 2/6, 3/6) in the scenario of section 4.2, which resemble the two market structures just discussed in the paragraph above. For these two distributions, the ability of firms to monitor each other is the same, because the size of the smallest firms is constant, so only the second effect is important in this case. As can be seen in Figure 1, the expected consumer surplus per unit for (1/6, 2/6, 3/6) is (weakly) higher than (1/6, 5/6) for all Δm. This implies that it would be worse to have two firms with large asymmetries than three firms with limited asymmetries, even when the latter is collusive. The reason is that, due to the fact that collusion is difficult in (1/6, 5/6), this outcome is noncollusive for most levels of Δm. Consequently, given it is a very asymmetric distribution, the expected competitive prices are very close to the monopoly level. Furthermore, despite the fact that collusion is sustainable for a larger range of fluctuations in market demand in (1/6, 2/6, 3/6) than in (1/6, 5/6), this collusion entails a sufficient number of price wars on the equilibrium path, which contrary to the conventional wisdom leads to higher consumer surplus than the near monopoly distribution of (1/6, 5/6). More generally, our analysis implies that, as $\delta \to 1$ and holding the size of the smallest firm constant, consumer surplus is always higher for outcomes with symmetric firms than outcomes with fewer firms and greater asymmetries, even if the symmetric firms collude.

5 Robustness

In this section, we explore the robustness of our results. We first consider the implications of relaxing Assumption 1, then we analyse an alternative allocation rule. For simplicity, we restrict attention to duopoly throughout this section, such that $K_{-1} = k_2$ and $k_1 \leq \frac{K}{2}$.

5.1 Relaxing Assumption 1

Up to this point, we have assumed that the minimum market demand is not too low, such that $\underline{m} \geq K_{-1}$ (see Assumption 1). One reason for this restriction is that it substantially simplifies the mixed strategy Nash equilibrium analysis when there are more than 2 firms and when capacities are asymmetric. However, as discussed above, this implies a loss generality because it restricts the size of the smallest firm in that it cannot be too small. In this subsection, we demonstrate that this assumption is not a necessary condition for firms to monitor each other through their privately observed sales. Whilst beyond the scope of the current paper, this implies that it would be possible to follow the rest of our analysis to replicate our other results.

Our results rely on the fact that $h^t = (y_0, y_1, \ldots, y_{t-1})$ is a public history, where y_τ is defined in (4). Recall that this implies that all firms can always infer from their sales when at least one firm's sales are below its trigger level. Thus, to demonstrate that Assumption 1 is not a necessary condition for firms to monitor each other through their sales, we establish that h^t is still a public history for some $\underline{m} < K_{-1}$. So, first consider the sales of firm i in period t, for any $p_{it} \leq 1$, if $\underline{m} < K_{-1}$:

$$s_{it}\left(p_{it}, p_{jt}; m_t\right) = \begin{cases} \min\left\{m_t, k_i\right\} & \text{if } p_{it} < p_{jt} \\ \max\left\{0, \min\left\{\frac{k_i}{K - \Phi\left(p_{it}\right) k_j}\left(m_t - \Phi\left(p_{it}\right) k_j\right), k_i\right\}\right\} & \text{if } p_{it} \geq p_{jt}, \end{cases}$$

where $\Phi(p_{it}) = 1$ if p_{it} is strictly above p_{jt} and 0 otherwise. This says that firm i's sales are the same as (1) except that if firm i now sets the highest price, then its sales are no longer guaranteed to be strictly positive for all $m_t > \underline{m}$. The reason is that the low-priced firm j will supply the whole market demand, if the realisation of market demand is sufficiently low, such that $m_t < k_j$.

Next, consider whether y_τ is common knowledge for all τ, for any z_i^t, such that h^t is a public history. Recall that $\overline{m} < K$ is the interesting case where it is possible for firms to receive sales above their trigger levels, since then $s_i^* = \frac{k_i(\overline{m}-k_1)}{K_{-1}} < k_i$. We now establish that h^t is still a public history if $\underline{m} > \frac{\overline{m}}{K}K_{-1}$. The reason is as follows. If both firms set a common price $p \leq 1$, then the sales of both firms will exceed their respective trigger levels if the realisation of market demand is high, otherwise they can both fall below the trigger levels. Yet, if both firms do not set such a common price, then the sales of the high-priced firm will not exceed its trigger level and the low-priced firm will supply either the whole market demand or their full capacity, $\min\{m_t, k_i\}$. Consequently, if the low-priced firm's sales are always strictly greater than its highest possible collusive sales, then it can still infer from its sales that its rival's sales are below its trigger level. Given firm 2's highest possible collusive sales are larger than firm 1's, it follows that both firms can make such inferences if $\underline{m} > \frac{\overline{m}}{K}K_{-1}$.

It follows from the above that h^t is a public history only if $\underline{m} > \frac{\overline{m}}{K}K_{-1}$, where $\frac{\overline{m}}{K}K_{-1} < K_{-1}$ for all $\overline{m} < K$, so Assumption 1 is not a necessary condition for our results. However, the above also implies that there is now a new restriction on the smallest firm that is less strict but still similar to Assumption 1. The reason for this new restriction is that if $\underline{m} \leq \frac{\overline{m}}{K}K_{-1}$ and if firm i receives very high sales, such that $s_{it} \geq \underline{m}$, then it is unsure whether market demand is fortunately high or whether its rival j has overcut it on price. Thus, the firm would not know from such sales whether its rival's sales are below its trigger level or not. Nevertheless, this new restriction is also not a necessary condition for firms to monitor each other through their sales. The reason is that there can be a public history in the case of $\underline{m} \leq \frac{\overline{m}}{K}K_{-1}$ if we also include a second 'upper' trigger level where firms also receive a bad signal, such that $y_\tau = \underline{y}$, if $s_{it} \geq \underline{m}$ for all i. Thus, the upper trigger level is $s_i^u \equiv k_i \frac{\underline{m}}{K_{-1}}$ for all i, which is strictly greater than the standard lower trigger if $\underline{m} > \overline{m} - k_1$ such that $s_i^u > s_i^*$. An implication of this is that if both firms set a common price, then they will only receive a good signal, such that $y_\tau = \overline{y}$, if the realisation of market demand is not too low and not too high. Yet, if both firms do not set a common price, then both firms will always receive a bad signal, because the low-priced firm's sales will be above its new upper trigger level and the high-priced firm's sales will be below its standard lower trigger level. Finally, firms will always receive a bad signal if $\underline{m} \leq \overline{m} - k_1$, in which case collusion under trigger-sales strategies will not be sustainable.

5.2 Allocation rule

Throughout the paper we have assumed that demand is allocated among firms with a common price in proportion to their joint capacity. As explained in section 2.2, this is a common assumption in the literature and it captures the plausible feature that buyers are more likely to be attracted to larger firms when prices are equal. However, it is not the only possible allocation rule. So, in this subsection, we consider the implications for monitoring if demand is instead allocated equally between the duopolists when they set a common price. We demonstrate that monitoring is more difficult than under the proportional allocation rule.

Under Assumption 1, $K_{-1} \leq \underline{m}$, and assuming $\overline{m} \leq 2k_1 \leq K$, the sales of firm i in period t, for any $p_{it} \leq 1$, are now:

$$s_{it}\left(p_{it}, p_{jt}; m_t\right) = \begin{cases} k_i & \text{if } p_{it} < p_{jt} \\ \frac{1}{2-\Phi(p_{it})}\left(m_t - \Phi\left(p_{it}\right) k_j\right) \in [0, k_i] & \text{if } p_{it} \geq p_{jt}. \end{cases} \tag{12}$$

This says that firm i's sales are the same as (1) except that if firms set a common price, then firm i's sales are now $\frac{1}{2}m_t$, where $\overline{m} \leq 2k_1$ guarantees that $\frac{1}{2}m_t \leq k_i$ for all t and all i. Trigger levels are again determined by the largest possible sales that firm 2 can make if firm 1 undercuts and supplies its full capacity, such that at least one firm will receive sales below its trigger level if a rival undercuts. It follows from (12) that each firm's trigger level is $s^e \equiv \frac{1}{2}m^e\left(k_1, \overline{m}\right) \leq k_i$, where $m^e\left(k_1, \overline{m}\right) \equiv 2\left(\overline{m} - k_1\right)$. This implies, in contrast to the main analysis, that the trigger level is the same for both firms. Firm 2's trigger level is the same as under the proportional allocation rule since the smallest firm still supplies its full capacity if it undercuts, leaving unchanged the largest possible sales of firm 2 in such an event. However, for any strictly asymmetric capacity distribution, such that $k_1 < K/2$, firm 1's trigger level has increased to the level of firm 2's. The reason is that firm 1's share of market demand when firms set a common price has now increased and it equals firm 2's share.[17]

Both firms can receive sales above the trigger level, if $s^e = \overline{m} - k_1 < k_i$ for all i, which is the case if $\overline{m} < 2k_1$. Recall that the equivalent conditon in the main text is $\overline{m} < K$. Consequently, for any strictly asymmetric capacity distribution, this necessary condition is now more stringent than before. The reason is that the smallest firm's share of the market demand at equal prices is now higher, so it supplies its full capacity for smaller fluctuations in market demand. As a result, it cannot infer for fluctuations in market demand as large as before whether it has set the same price as or undercut its rival. Furthermore, it follows from $m^e\left(k_1, \overline{m}\right) > m^*\left(k_1, \overline{m}\right)$

[17]Similar to the main analysis, it can be the case that the smallest firm knows for sure that both firms have set a common price, but its rival is uncertain as to whether it has been undercut. This can occur in this case if the highest possible sales of the smallest firm when it is undercut are less than the trigger level, $\overline{m} - K_{-1} < s^e$, which requires $k_1 < K/2$. Nevertheless, if its sales are above $\overline{m} - K_{-1}$ but below s^e, then it can infer that the largest firm's sales are also below the trigger level.

for any $k_1 < K/2$ that both firms' sales when they set a common price can now be below the trigger level for higher realisations of market demand than under the proportional allocation rule. This is due to the fact that, whilst the largest firm's trigger level remains the same, its share of market demand is now lower when firms set a common price. Thus, the largest firm's collusive sales are now consistent with a deviation by the smallest firm for higher levels of market demand. Consequently, monitoring is more difficult than in the main text. Finally, replicating the steps of section 3.2, it is easy to establish that collusive profits under imperfect monitoring are lower than under the proportional allocation rule, because punishment phases occur on the equilibrium path more often than before.

6 Concluding remarks

We have explored the effects of asymmetries in capacity constraints on collusion in a setting where there is demand uncertainty and where firms never directly observe their rivals' prices and sales. Despite the fact that each firm must monitor the collusive agreement using their privately observed prices and sales, we have shown that firms can perfectly detect deviations if demand fluctuations are sufficiently small, and that the critical level is determined by the capacity of the smallest firm. Otherwise, monitoring is imperfect and punishment phases occur on the equilibrium path. Consistent with the previous literature, we found that asymmetries hinder collusion. Yet, we also analysed both the unilateral and coordinated effects of mergers in a unified framework. We showed that if demand fluctuations are sufficiently large, then the competitive prices of asymmetric capacity distributions are actually higher than the collusive prices of less asymmetric capacity distributions.

Our results have three main implications for merger policy. First, although market transparency is rightly an important criterion in the assessment of coordinated effects in practice, our model re-emphasises the fact that a lack of transparency about rivals' prices and sales is not a sufficient condition to rule out such effects: firms may be able to monitor a collusive agreement using their private information. Second, while the possible effects of imperfect monitoring are explicitly mentioned in general terms in the most recent US and European horizontal merger guidelines, our model suggests that such monitoring will be difficult in markets where firm asymmetries are large. Third, and most importantly, collusive merger outcomes should not be presumed to be more harmful than more asymmetric noncollusive merger outcomes. A collusive agreement may require sufficiently frequent price wars that actually lead to higher consumer surplus than a more asymmetric outcome in which one firm's market power is strengthened unilaterally. Consequently, it can be inappropriate for a competition authority to remedy a collusive post-merger outcome by imposing a divestment that creates a more asymmetric market structure. Likewise, it can be appropriate to remedy a merger outcome with a singularly dominant firm by imposing a divestment that creates a symmetric market structure, even if this faciliates collusion under imperfect monitoring.

Finally, an important avenue for future research is to develop techniques that can assess accurately in practice whether a collusive merger outcome is better than a noncollusive outcome. We believe that there are two existing methodologies that are a good place to start to address this issue. First, if collusion is expected to be facilitated post-merger, then more advanced merger simulation techniques may enable the likely effects to be simulated. While these techniques have been criticised in the past for not allowing the conduct of firms post-merger to differ from that assumed pre-merger (see Whinston, 2007), recent developments have have gone some way to address this. For example, Davis and Huse (2010) and Ivaldi and Lagos (2015) use the standard simulation methodology of unilateral effects to develop empirical techniques that simulate coordinated effects on the critical discount factor in differentiated goods market under perfect observability. Thus, to be applicable to the current context, these techniques would have to be extended to the case of imperfect monitoring. Second, if the pre-merger status quo is thought to be collusive, then it may be possible to use the screening devices for collusion on pre-merger data to estimate the extent to which such collusion is imperfect (see Harrington, 2008, for a review of these devices). For example, one such technique, developed by Porter (1983) and Ellison (1994), discovers collusion under imperfect monitoring by finding evidence of price wars (that is, abrupt changes in prices that cannot be explained by fluctuations in cost or demand). Such price wars are inconsistent with models of competition and with perfect collusion, so if there is no evidence of them pre-merger (and price is at or near the monopoly level), then it is likely that the coventional wisdom will hold if collusion pre-merger is expected to be destabilised post-merger. On the other hand, evidence of frequent or long price wars pre-merger could suggest that collusion is sufficiently imperfect that the conventional wisdom will not hold.

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] Amelio, A., and Biancini, S. (2010) "Alternating Monopoly and Tacit Collusion," *Journal of Industrial Economics,* 58(2), 402-423

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

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

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

[7] Campbell, C., Ray, G., and Muhanna, W. (2005) "Search and Collusion in Electronic Markets." *Management Science*, 51, 497–507

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

[9] Dasgupta, P. and Maskin, E. (1986) "The Existence of Equilibrium in Discontinuous Economic Games, II: Applications" *Review of Economic Studies*, 53(1), 27-41

[10] Davies, S. and Olczak, M. (2010) "Assessing the Efficacy of Structural Merger Remedies: Choosing Between Theories of Harm?" *Review of Industrial Organization*, 37(2), 83-99

[11] 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

[12] Davis, P. and Huse, C. (2010) "Estimating the 'Coordinated Effects' of Mergers," Competition Commission Working Paper

[13] Ellison, G. (1994) "Theories of Cartel Stability and the Joint Executive Committee," *RAND Journal of Economics*, 25(1), 37-57

[14] Fabra, N. (2006) "Collusion with Capacity Constraints over the Business Cycle," *International Journal of Industrial Organization*, 24(1), 69-81

[15] Farrell, J. and Shapiro, C. (1990) "Horizontal Mergers: An Equilibrium Analysis," *American Economic Review*, 80(1), 107-126

[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] Fudenberg, D. and Tirole, J. (1994), *Game Theory*, MIT Press

[19] Gal-Or, E. (1984) "Price Dispersion with Uncertain Demand," *International Economic Review*, 25(2), 441-457

[20] Green, E. and Porter, R. (1984) "Non-Cooperative Collusion under Imperfect Price Information," *Econometrica*, 52(1), 87-100

[21] Hanazono, M. and Yang, H. (2007) "Collusion, Fluctuating Demand, and Price Rigidity," *International Economic Review*, 48(2), 483-515

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

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

[24] Harrington, J. and Skrzypacz, A. (2007) "Collusion under Monitoring of Sales," *RAND Journal of Economics*, 38(2), 314-331

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

[27] Ivaldi, M., Jullien, B., Rey, P., Seabright, P., and Tirole, J. (2003b) "The Economics of Unilateral Effects," Report for DG Comp, European Commission

[28] Ivaldi, M. and Lagos, V. (2015) "Post-merger coordinated effects: Characterization and Assessment by Simulations," mimeo, University of Toulouse

[29] Knittel, C. and Lepore, J. (2010) "Tacit Collusion in the Presence of Cyclical Demand and Endogenous Capacity Levels," *International Journal of Industrial Organization*, 28(2), 131-144

[30] Kühn, K.-U., (2001) "An Economist's Guide Through the Joint Dominance Jungle," John M. Olin Centre for Law and Economics Working Paper 02-014

[31] 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

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

[33] Lyons, B. (2002) "Could Politicians be More Right Than Economists? A Theory of Merger Standards," Centre for Competition and Regulation Working Paper 02-1

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

[35] Motta, M., Polo, M., and Vasconcelos, H. (2003) "Merger Remedies in the European Union: An Overview," in Lévêque, F. and Shelanski, H., (eds.), *Merger Remedies in American and European Union Competition Law*, Edward Elgar

[36] Porter, R. (1983). "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886," *Bell Journal of Economics*, 14(2), 301-314

[37] Röller, L.-H. and Mano, M. (2006) "The Impact of the New Substantive Test in European Merger Control," *European Competition Journal*, 2(1), 9-28

[38] Staiger, R. and Wolak, F. (1992) "Collusive Pricing with Capacity Constraints in the Presence of Demand Uncertainty," *RAND Journal of Economics*, 23(2), 203-220

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

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

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

[42] Vives, X. (1999). *Oligopoly Pricing: Old Ideas and New Tools*, MIT Press

[43] Whinston, M. (2007) "Antitrust Policy toward Horizontal Mergers" in Armstrong, M. and Porter, R., (ed.), *Handbook of Industrial Organization*, North-Holland

Appendix A

Proof of Lemma 1. There exists a unique pure strategy Nash equilibrium if $\underline{m} \geq K$, where $\pi_i^N = k_i \; \forall \; i$. This follows from $\pi_i(p_i, \mathbf{p}_{-i}) = p_i k_i \; \forall \; p_i \leq 1$, so the best reply of firm i is $p_i = 1$ for any $\mathbf{p}_{-i}$, $\forall \; i$. There is no pure strategy Nash equilibrium if $\underline{m} < K$. To see this, note that any such candidate equilibrium requires $p_j = p \; \forall \; j$. Otherwise, firm $i \in \Omega\left(p^{\max}\right)$ has an incentive to increase its price towards $p^{\max}$, from $\pi_i\left(p_i, \mathbf{p}_{-i}\right) = p_i k_i \; \forall \; p_i < p^{\max}$. However, for any $p \in (0, 1]$, firm i has an incentive to lower its price, since $\pi_i(p - \epsilon, p) > \pi_i(p)$ if $\underline{m} < K$, where $\epsilon > 0$ but small. Moreover, for $p = 0$, firm i has an incentive to raise its price, since Assumption 1 ensures $\pi_i(\epsilon, 0) > 0 \; \forall \; i$.

Nevertheless, if $K > \underline{m} \geq K_{-1}$, the existence of a mixed strategy Nash equilibrium is guaranteed by Thereom 1 of Dasgupta and Maksin (1986). To characterise this equilibrium, let $H_i(p)$ denote the probability that firm i charges a price less than or equal to p. Below we demonstrate that the equilibrium profits are given by (2) for all i and that:

$$H_i(p) = \frac{1}{k_i}\left[\frac{\left(\overline{\pi}_n - pk_n\right)}{pk_n\left(\int_{\underline{m}}^{\min\{K,\overline{m}\}}\left(m-K\right)g\left(m\right)dm\right)}\prod_{j=1}^{n} k_j\right]^{1/(n-1)}, \tag{13}$$

where firm i's expected profits are given by $\overline{\pi}_i$ in (3), if it is strictly the highest-priced firm with $p_i = 1$. This converges to the analysis in Fonseca and Normann (2008) as $\overline{m} \to \underline{m}$.

In equilibrium, firm i must receive the following expected profit from charging $p \leq 1$:

$$p\left(\prod_{j\neq i} H_j(p)\overline{\pi}_i + \left(1 - \prod_{j\neq i} H_j(p)\right)k_i\right) = \frac{k_i}{k_n}\overline{\pi}_n, \; \forall \; i \tag{14}$$

where $\prod_{j\neq i} H_j(p)$ is the probability that firm i is the highest-priced firm. To solve for the right-hand side of (14), notice firm i has no incentive to price below $\overline{\pi}_i/k_i \equiv \underline{p}_i$, where $\underline{p}_n \geq \underline{p}_{n-1} \geq \ldots \geq \underline{p}_1$. Moreover, any firm $j < n$ can guarantee profits of $\frac{k_j}{k_n}\overline{\pi}_n \geq \overline{\pi}_j$ by charging a price marginally below $\underline{p}_n$, so all firms have no incentive to price below $\underline{p}_n$. Finally, the fact that all firms $j < n$ place positive probability on charging $\underline{p}_n$ is necessary and sufficient to ensure $\underline{p}_n$ is also the lowest price that firm n will charge. Thus, the lower bound of $H_i(p)$ is $\underline{p} = \underline{p}_n = \overline{\pi}_n/k_n$ $\forall \; i$. Manipulating (14) yields:

$$H_i(p) = \frac{pk_n\left(\overline{\pi}_i - k_i\right)}{\overline{\pi}_n - pk_n}\prod_{j} H_j(p)\frac{1}{k_i}. \tag{15}$$

Noting that $\overline{\pi}_i - k_i = \int_{\underline{m}}^{\min\{K,\overline{m}\}}\left(m-K\right)g\left(m\right)dm \; \forall \; i$ for any $K > \underline{m}$ from (3), it follows from

(15) that:

$$\prod_j H_j(p) = \left[\frac{pk_n \left(\int_{\underline{m}}^{\min\{K,\overline{m}\}} (m-K)\, g(m)\, dm \right)}{\overline{\pi}_n - pk_n} \prod_j H_j(p) \right]^n \prod_{l=1}^{n} \left(\frac{1}{k_l} \right).$$

Thus, solving for $\prod_j H_j(p)$ and substituting into (15) shows that $H_i(p)$ is as claimed in (13).

It follows from (13) that $H_i(1) \lesseqgtr 1$ if $\frac{k_i^{n-1}}{\prod_{j\neq n} k_j} \gtreqless 1$. This has two implications. First, if $\frac{k_i^{n-1}}{\prod_{j\neq n} k_j} \geq 1$, then firm i randomises over $\left[\underline{p}, 1\right]$ and puts mass of $1 - H_i(1)$ on a price of 1 when the inequality is strict. Note that $\frac{k_i^{n-1}}{\prod_{j\neq n} k_j} > 1$ never holds if $k_i = k \;\forall\; i$ but always holds for firm n if $k_n > k_1$. Second, if $\frac{k_i^{n-1}}{\prod_{j\neq n} k_j} < 1$ for some $i < n$, then firm i randomises over $\left[\underline{p}, \overline{p}_i\right]$ where $\overline{p}_i < 1$ solves $H_i(\overline{p}_i) = 1$. Consequently, the probability distributions of the larger firms with higher upper bounds must be adjusted accordingly. For example, if $\overline{p}_i < 1$ only for firm 1 (which is the case for any triopoly with $k_1 < k_2$), then the largest $n-1$ firms play with the $H_i(p)$ adjusted so that $n-1$ replaces n over $[\overline{p}_1, 1]$. Note that $\frac{k_i^{n-1}}{\prod_{j\neq n} k_j} < 1$ never holds if $n = 2$ or if $k_i = k \;\forall\; i$ for any $n \geq 2$. ■

Proof of Proposition 1. There is perfect monitoring if $\underline{m} > m^*(k_1, \overline{m})$ and imperfect monitoring otherwise. Given $\frac{\partial m^*}{\partial \overline{m}} > 0$, it follows that there is a unique level of $\overline{m}$ that solves $m^*(k_1, \overline{m}) = \underline{m}$. Substituting in for $m^*(k_1, \overline{m})$ and rearranging yields $\overline{m} = k_1 \left(\frac{K-\underline{m}}{K} \right) + \underline{m} \equiv \underline{x}(k_1)$, where $\underline{x}(k_1) \in (\underline{m}, K)$ for any $\underline{m} < K$. Thus, monitoring is perfect if $\overline{m} < \underline{x}(k_1)$, as this implies $\underline{m} > m^*(k_1, \overline{m})$. Otherwise, there is imperfect monitoring. ■

Proof of Proposition 2. Given $k_i V^c$ in (7) is strictly decreasing in T, the optimal equilibrium profits for firm i can be found by evaluating it at $T^*(k_1, k_n, p^c)$. Thus, it follows from (9) that:

$$1 - \delta^{T^*} = \frac{(1-\delta)(K - \widehat{m})\, p^c}{\delta \left[(1 - G(m^*(k_1, \overline{m}))\, K \left(p^c - \underline{p}\right) - (K - \widehat{m})\, p^c) \right]}. \tag{16}$$

Then, substituting this into (7) yields:

$$k_i V^c = \frac{k_i}{K} \left(\frac{\widehat{m} - G(m^*(k_1, \overline{m}))\, K}{1 - G(m^*(k_1, \overline{m}))} \right) p^c, \;\forall\; i.$$

This is strictly increasing in p^c, so $p^c = 1$ and $k_i V^*$ is as claimed. Substituting $p^c = 1$ into (10) and (11) yields:

$$\delta \geq \frac{1}{1 - G(m^*(k_1, \overline{m}))} \frac{k_n}{K} \equiv \delta^*(k_1, k_n)$$

and

$$G\left(m^*\left(k_1, \overline{m}\right)\right) < \frac{K_{-n}}{K},$$

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

Proof of Proposition 3. Differentiating $\delta^*\left(k_1, k_n\right) = \frac{1}{\left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right)} \frac{k_n}{K}$ with respect to k_j yields:

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

Thus, $\frac{\partial \delta^*}{\partial k_1} < 0$ from $\frac{\partial k_n}{\partial k_1} \in [-1, 0]$, $\frac{\partial m^*}{\partial k_1} < 0$ and $\frac{\partial k_1}{\partial k_1} = 1$. Furthermore, $\frac{\partial \delta^*}{\partial k_n} > 0$ 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, $\underline{\delta}^*\left(k_n\right) = \frac{k_n}{K}$ implies $\frac{\partial \delta^*}{\partial k_n} > 0$. ■

Proof of Proposition 4. Differentiating $\widehat{p}^c\left(k_1, \overline{m}\right) = \frac{K}{\widehat{m}} V^*$ with respect to k_j yields:

$$\frac{\partial \widehat{p}^c}{\partial k_j} = \frac{K}{\widehat{m}} \frac{\partial V^*}{\partial k_j} = -\frac{\left(K - \widehat{m}\right) g\left(m^*\right)}{\widehat{m}\left(1 - G(m^*)\right)^2} \frac{\partial m^*}{\partial k_1} \frac{\partial k_1}{\partial k_j}.$$

Thus, $\frac{\partial \widehat{p}^c}{\partial k_1} > 0$ from $0 < \widehat{m} < \overline{m} < K$, $\frac{\partial m^*}{\partial k_1} < 0$ and $\frac{\partial k_1}{\partial k_1} = 1$. ■

Proof of Proposition 5. We first show that $\widehat{p}^N\left(k_n', \widehat{m}\right) > \widehat{p}^c\left(k_1, \overline{m}\right)$ if $\overline{m} > \overline{x}\left(k_1, k_n'\right)$. This follows since $\widehat{p}^N\left(k_n', \widehat{m}\right) > \widehat{p}^c\left(k_1, \overline{m}\right)$ if $G\left(m^*\left(k_1, \overline{m}\right)\right) > 1 - \frac{k_n'}{K}$. Furthermore, in Proposition 2, $\overline{x}\left(k_1, k_n\right)$ is defined as the level of $\overline{m}$ that solves $G\left(m^*\left(k_1, \overline{x}\left(k_1, k_n\right)\right)\right) = 1 - \frac{k_n}{K}$. Thus, $\overline{x}\left(k_1, k_n'\right)$ is the level that solves $G\left(m^*\left(k_1, \overline{x}\left(k_1, k_n'\right)\right)\right) = 1 - \frac{k_n'}{K}$ where $\overline{x}\left(k_1, k_n'\right) > \underline{x}\left(k_1\right)$. This and $\frac{\partial G(m^*)}{\partial \overline{m}} > 0$ implies that if $\overline{m} > \overline{x}\left(k_1, k_n'\right)$, such that $G\left(m^*\left(k_1, \overline{x}\left(k_1, k_n\right)\right)\right) > 1 - \frac{k_n'}{K}$, then $\widehat{p}^N\left(k_n', \widehat{m}\right) > \widehat{p}^c\left(k_1, \overline{m}\right)$.

Next, note that this comparison is only meaningful if $\overline{m} < \overline{x}\left(k_1, k_n\right)$ and if $\delta \geq \delta^*\left(k_1, k_n\right)$ such that $\widehat{p}^c\left(k_1, \overline{m}\right)$ is an equilibrium average price. So, we next show that $\overline{x}\left(k_1, k_n'\right) < \overline{x}\left(k_1, k_n\right)$ if $k_n' > k_n$ such that $k_1' \leq k_1$. Using the implicit function theorem on $Z \equiv 1 - \frac{k_n}{K} - G\left(m^*\left(k_1, \overline{m}\right)\right) = 0$ yields:

$$\frac{\partial \overline{x}}{\partial k_j} = -\frac{\frac{\partial Z}{\partial k_j}}{\frac{\partial Z}{\partial \overline{m}}} = -\frac{1}{\left(g\left(m^*\right) \frac{\partial m^*}{\partial \overline{m}}\right)}\left(\frac{1}{K} \frac{\partial k_n}{\partial k_j} + g\left(m^*\right) \frac{\partial m^*}{\partial k_1} \frac{\partial k_1}{\partial k_j}\right).$$

It follows from this that $\frac{\partial \overline{x}}{\partial k_n} < 0$, since $\frac{\partial m^*}{\partial \overline{m}} > 0$, $\frac{\partial k_n}{\partial k_n} = 1$, $\frac{\partial m^*}{\partial k_1} < 0$ and $\frac{\partial k_1}{\partial k_n} \in [-1, 0]$. So, if $k_n' > k_n$, then $\overline{x}(k_1, k_n') < \overline{x}(k_1, k_n)$.

Thus, the above implies that, for any $\delta \geq \delta^*(k_1, k_n)$, if $k_n' > k_n$ (so $k_1' \leq k_1$), then there exists a unique $\overline{x}(k_1, k_n') \in (\underline{x}(k_1), \overline{x}(k_1, k_n))$ such that if $\overline{x}(k_1, k_n') < \overline{m} < \overline{x}(k_1, k_n)$, then $\widehat{p}^N(k_n') > \widehat{p}^c(k_1, \overline{m})$. Finally, notice that if $\overline{x}(k_1, k_n') < \overline{m} < \overline{x}(k_1, k_n)$, then $\overline{x}(k_1', k_n') < \overline{m}$ from $k_1' \leq k_1$ and $\frac{\partial \overline{x}}{\partial k_1} > 0$. Consequently, it follows from Proposition 2 that collusion under trigger-sales strategies is not sustainable for (k_1', k_n'). ■

Appendix B

One possible limitation of the analysis in the main paper is that restricting attention to trigger-sales strategies leaves open the question of whether there are other PPE with higher payoffs. Consequently, to check the robustness of our results, we now use the techniques of Abreu *et al.* (1986, 1990) to find the set of perfect public equilibria. This analysis shows that trigger-sales strategies are a strategy profile that supports the maximal PPE payoffs, such that there are no such equilibria with higher payoffs, and it also shows that trigger-sales strategies generate the lowest critical discount factor. Following the main text, we impose $K_{-1} \leq \underline{m} < K$, such that Assumption 1 is satisfied and the static Nash equilibrium is in mixed strategies. We also let $\overline{m} \in [\underline{x}(k_1), K)$ to restrict attention to the case where $G(m^*(k_1, \overline{m})) \in (0, 1)$ such that there is imperfect monitoring.[18] Finally, we make the common assumption that, after observing y_t, the firms observe the realisation of a publically observable randomisation device, which allows them to select among the continuation equilibria. This ensures that the set of PPE payoffs is convex. For a similar analysis to the below, see Athey and Bagwell (2001) and Hanazono and Yang (2007).[19]

Following Abreu *et al.* (1986, 1990), we define an operator $B(W)$ that, for any set $W \subseteq \mathbf{R}$, yields the set of PPE payoffs per unit of capacity as the largest invariant set. The operator is defined as follows:

$$\begin{aligned} B(W) \equiv \{V : \exists \quad & p \in [0,1] \text{ and } V^p, V^c \in co(W), \text{ such that} \\ & V = (1-\delta)\tfrac{\widehat{m}}{K}p + \delta\left[G(m^*(k_1,\overline{m}))V^p + (1 - G(m^*(k_1,\overline{m})))V^c\right] \text{ and} \\ & k_iV \geq (1-\delta)\pi_i(p_i^*, p) + \delta k_i V^p \ \forall\ i\} \cup \underline{p}, \end{aligned}$$

[18] The set of PPE payoffs coincide with the set of SPNE payoffs when there is perfect monitoring, so the latter set can be easily generated for $\underline{m} < \overline{m} < \underline{x}(k_1)$ by setting $G(m^*(k_1, \overline{m})) = 0$ in the below.

[19] In contrast to our analysis below and Athey and Bagwell (2001), Hanazono and Yang (2007) restrict attention to PPE that are strongly symmetric in the sense that each firm uses an identical strategy after every public history.

where $\underline{p}$ is the expected static (mixed strategy) Nash equilibrium profits per unit of capacity, which may be used as an off-the-equilibrium path punishment. This operator decomposes play on the equilibrium path into a current period price, p, and continuation payoffs per unit of capacity, V^p and V^c, that are drawn from the convex hull of the set W. The inequality is the ICC that ensures that no firm is able to gain by a (one-stage) deviation from the public strategy, where firm i's optimal deviation profit from any price $p \in [0, 1]$ is:

$$\pi_i(p_i^*, p) \equiv \begin{cases} k_i p & \text{if} \quad p \geq \overline{\pi}_i / k_i \\ \overline{\pi}_i & \text{if} \quad p < \overline{\pi}_i / k_i. \end{cases}$$

This says that firm i's optimal deviation from p is to undercut p marginally, if p is sufficiently high, otherwise firm i should supply the residual demand at the monopoly price. Recall that $\overline{\pi}_i = \widehat{m} - K_{-i}$ for any $\overline{m} < K$ from (3) such that $\pi_i(p_i^*, p)$ is (weakly) increasing in p.

Next, we establish that the operator $B(W)$ maps compact sets to compact sets, which is the critical property for applying the techniques of Abreu *et al.* (1990).

Lemma 2. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m}$, $B(W)$ maps compact sets to compact sets.*

Proof. Notice that the feasible set of payoffs per unit of capacity is real-valued, bounded and closed, e.g. $\left[0, \frac{\widehat{m}}{K}\right]$, hence it is compact. Then $B(W)$ is bounded and closed, because the constraints entail weak inequalities and each component of the value function and the constraints is real-valued, continuous, and bounded. Thus, $B(W)$ is compact. ■

Given the operator is compact, we can use the following algorithm to compute the set of PPE. Let $W_0 = \left[0, \frac{\widehat{m}}{K}\right]$ such that it is compact and it contains all feasible payoffs per unit of capacity, and let $W_{z+1} = B(W_z)$, for any $z > 0$. Then $B(W)$ implies that $B(W_z) = W_{z+1} \subseteq W_z$ such that we have a monotonic sequence. It follows from this, and the fact that W_z is non-empty (since $\underline{p}$ is in every W_z), that $W^* = \lim_{z\to\infty} W_z$ is a non-empty, compact set. Following the arguments in Abreu *et al.* (1990), we can conclude that W^* is the largest invariant set of $B(W)$ and hence it is the set of PPE payoffs per unit of capacity. Thus, it can be represented by the interval $[\underline{V}, \overline{V}]$ and solving for this set reduces to the problem of finding the minimal $\underline{V}$ and the maximal $\overline{V}$ that satisfy $[\underline{V}, \overline{V}] = B\left([\underline{V}, \overline{V}]\right)$. Proposition 6 solves for $\overline{V}$ and $\underline{V}$.

Proposition 6. *For any given $n \geq 2$ and $K_{-1} \leq \underline{m} < K$, if $\overline{m} \in [\underline{x}(k_1), \overline{x}(k_1, k_n))$, such that $G(m^*(k_1, \overline{m})) \in \left(0, \frac{K_{-n}}{K}\right)$, and if $\delta \geq \delta^*(k_1, k_n) = \frac{1}{1-G(m^*)}\frac{k_n}{K}$, then $\overline{V} = \frac{1}{K}\left(\frac{\widehat{m} - G(m^*)K}{1-G(m^*)}\right)$ and $\underline{V} = \underline{p}$.*

Proof. First, we find $\overline{V}$ by solving the following constrained maximisation problem:

$$\overline{V} = \max_{p,V^p} V^c$$

subject to:

$$V^c = (1-\delta)\frac{\widehat{m}}{K}p + \delta\left[G\left(m^*\left(k_1,\overline{m}\right)\right)V^p + \left(1 - G\left(m^*\left(k_1,\overline{m}\right)\right)\right)V^c\right] \tag{17}$$

$$k_n V^c \geq (1-\delta)\,\pi_n\left(p_n^*, p\right) + \delta k_n V^p \tag{18}$$

$$p \in [0,1] \text{ and } V^p \in \left[\underline{V}, \overline{V}\right]$$

The first constraint (17) is just an identity that says that the target continuation payoff per unit of capacity, V^c, can be decomposed into the profit per unit of capacity of the stage game and a continuation payoff function, where if $\underline{y}$, the each firm gets V^p per unit of capacity, but if $\overline{y}$, they each get V^c per unit of capacity. The second constraint (18) is the ICC for firm n, which is a necessary and sufficient condition for the ICC to be satisfied for all i and all $p \in [0,1]$.

The Lagrangian function for this constrained maximisation problem is:

$$L = V^c + \lambda^c \xi_n^c,$$

where ξ_n^c denotes the slack in the ICC for firm n, such that $\xi_n^c = k_n V^c - (1-\delta)\,\pi_n\left(p_n^*, p\right) - \delta k_n V^p$, λ^c represents the Lagrange multiplier, and:

$$V^c = \frac{(1-\delta)\frac{\widehat{m}}{K}p + \delta G\left(m^*\left(k_1,\overline{m}\right)\right)V^p}{1 - \delta\left(1 - G\left(m^*\left(k_1,\overline{m}\right)\right)\right)},$$

from rearranging (17). It is helpful to solve the constrained maximum for a given $p \in [0,1]$, and then take this solution to its maximum with respect to p. Thus, the necessary Kuhn-Tucker conditions for a maximum are:

$$\frac{\partial L}{\partial V^p} = \frac{\partial V^c}{\partial V^p} + \lambda^c\left(k_n\frac{\partial V^c}{\partial V^p} - \delta k_n\right) = 0$$

$$\frac{\partial L}{\partial \lambda^c} = \xi_n^c \geq 0, \quad \lambda^c \geq 0, \quad \lambda^c \xi_n^c = 0.$$

These conditions are necessary and sufficient to determine a maximum, because the Lagrangian is concave.

We begin the solution by noting that the Kuhn-Tucker conditions are not satisfied if $\xi_n^c > 0$ such that $\lambda^c = 0$, because then $\frac{\partial L}{\partial V^p} > 0$, which is a contradiction. So, consider the case of $\xi_n^c = 0$ such that $\lambda^c \geq 0$. Rearranging $\frac{\partial L}{\partial V^p} = 0$ yields $\lambda^c = -\frac{\partial V^c}{\partial V^p}\left(\frac{1}{k_n\frac{\partial V^c}{\partial V^p} - \delta k_n}\right) > 0$ for any $\overline{m} \geq \underline{x}\left(k_1\right)$. Thus, the solution to the problem sets $\xi_n^c = 0$, where:

$$V^p = V^c - \frac{1-\delta}{\delta}\left(\frac{\pi_n\left(p_n^*, p\right)}{k_n} - V^c\right).$$

Substituting the above into (17) yields $V^c = \left(\frac{1}{1-G(m^*)}\right)\left(\frac{\widehat{m}}{K}p - G(m^*)\frac{\pi_n(p_n^*,p)}{k_n}\right)$. Given this is strictly increasing in p, for all $p \in [0,1]$, it follows that $\overline{V}$ has $p = 1$ such that $\pi_n(p_n^*, p) = k_n$ and $\overline{V} = \frac{\widehat{m}-G(m^*)K}{K(1-G(m^*))}$. The above conditions imply that the strategy profiles that support the maximal PPE payoffs can be chosen in the following form: for $t = 1$, each firm sets $p = 1$; for every $t > 1$, each firm sets $p = 1$ if $y_{t-1} = \overline{y}$; otherwise firms adopt any PPE strategy profile that supports $V^p = \overline{V} - \frac{1-\delta}{\delta}\left(1 - \overline{V}\right)$ per unit of capacity. Thus, it remains to construct such a PPE profile. One way to achieve this is to use the public randomisation device, where firms choose some probability $\alpha \in [0,1]$ such that $V^p = (1-\alpha)\overline{V} + \alpha\underline{V}$. Substituting in for V^p, $\overline{V}$ and $\underline{V} = \underline{p}$ (which we prove below) yields:

$$\alpha = \frac{(1-\delta)\,k_n}{\delta\left[K\left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right) - k_n\right]}, \tag{19}$$

where $\alpha > 0$ if $\overline{m} \in \left[\underline{x}\left(k_1\right), \overline{x}\left(k_1, k_n\right)\right)$, such that $G\left(m^*\left(k_1, \overline{m}\right)\right) \in \left(0, \frac{K_{-n}}{K}\right)$, and where $\alpha \leq 1$ if $\delta \geq \delta^*\left(k_1, k_n\right)$.

Next, we find $\underline{V}$ by solving the following constrained minimisation problem:

$$\underline{V} = \min_{p, V^c} V^p$$

subject to:

$$V^p = (1-\delta)\,\frac{\widehat{m}}{K}p + \delta\left[G\left(m^*\left(k_1, \overline{m}\right)\right)V^p + \left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right)V^c\right] \tag{20}$$

$$k_n V^p \geq (1-\delta)\,\pi_n\left(p_n^*, p\right) + \delta k_n V^p \tag{21}$$

$$p \in [0,1] \text{ and } V^c \in \left[\underline{V}, \overline{V}\right]$$

Similar to the problem above, the first constraint (20) is just an identity and the second constraint (21) is the ICC for firm n.

The Lagrangian function for this constrained minimisation problem is:

$$L = V^p - \lambda^p\left(k_n V^p - \pi_n\left(p_n^*, p\right)\right),$$

where λ^p represents the Lagrange multiplier, the term in brackets represents the slack in the ICC for firm n, and:

$$V^p = \frac{(1-\delta)\,\frac{\widehat{m}}{K}p + \delta\left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right)V^c}{1 - \delta G\left(m^*\left(k_1, \overline{m}\right)\right)}, \tag{22}$$

from rearranging (20). We again solve the constrained minimum for a given $p \in [0,1]$, and then take this solution to its minimum with respect to p. Thus, the necessary Kuhn-Tucker conditions for a minimum are:

$$\frac{\partial L}{\partial V^c} = \frac{\partial V^p}{\partial V^c} - \lambda^p k_n \frac{\delta\left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right)}{1 - \delta G\left(m^*\left(k_1, \overline{m}\right)\right)} = 0$$

$$\frac{\partial L}{\partial \lambda^p} = -\left(k_n V^p - \pi_n\left(p_n^*, p\right)\right) \leq 0, \quad \lambda^p \geq 0, \quad \lambda^p\left(k_n V^p - \pi_n\left(p_n^*, p\right)\right) = 0.$$

The Lagrangian is concave so these conditions are necessary and sufficient to determine a minimum.

Note that the Kuhn-Tucker conditions are not satisfied if $k_n V^p > \pi_n\left(p_n^*, p\right)$ such that $\lambda^p = 0$, because then $\frac{\partial L}{\partial V^c} > 0$, which is a contradiction. So, consider the case of $k_n V^p = \pi_n\left(p_n^*, p\right)$ such that $\lambda^p \geq 0$. Rearranging $\frac{\partial L}{\partial V^c} = 0$ yields $\lambda^p = 1/k_n > 0$, so the solution to the problem is $V^p = \frac{\pi_n\left(p_n^*, p\right)}{k_n}$ such that:

$$V^c = \frac{\pi_n\left(p_n^*, p\right)}{k_n} + \left(\frac{1-\delta}{\delta\left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right)}\right)\left(\frac{\pi_n\left(p_n^*, p\right)}{k_n} - \frac{\widehat{m}}{K}p\right),$$

from (22). Note that V^p is (weakly) increasing in p, for all $p \in [0,1]$, so V^p is at its minimum if $p \leq \underline{p}$ such that $\underline{V} = \underline{p}$ and $V^c = \underline{p} + (1-\delta)\left(\underline{p} - \frac{\widehat{m}}{K}p\right) > \underline{p}$ for all $p \leq \underline{p}$. Finally, note that the above conditions imply that a strategy profile that supports the minimal PPE payoffs can be chosen in the following form: for $t = 1$, each firm sets some $p \in \left[0, \underline{p}\right]$; for $t > 1$, each firm sets p if $y_{t-1} = \underline{y}$; otherwise, firms adopt any PPE strategy profile that supports V^c. Thus, it remains to show that we can construst such a PPE profile. Again, using the randomisation device, suppose firms choose some probability $\beta \in [0,1]$ such that $V^c = (1-\beta)\overline{V} + \beta\underline{V}$. Substituting in for V^c and $\underline{V}$ yields:

$$\beta = 1 - \left(\frac{1-\delta}{\delta\left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right)}\right)\left(\frac{\underline{p} - \frac{\widehat{m}}{K}p}{\overline{V} - \underline{p}}\right),$$

where if $\overline{m} \in \left[\underline{x}\left(k_1\right), \overline{x}\left(k_1, k_n\right)\right)$, such that $G\left(m^*\left(k_1, \overline{m}\right)\right) \in \left(0, \frac{K_{-n}}{K}\right)$, then $\beta < 1$ for all $p \leq \underline{p}$. Finally, $\beta \geq 0$ if:

$$p \geq \frac{K}{\widehat{m}}\left(\underline{p} - \frac{\delta\left(1 - G\left(m^*\left(k_1, \overline{m}\right)\right)\right)\left(\overline{V} - \underline{p}\right)}{1-\delta}\right) \equiv p^*,$$

where $\delta \geq \delta^*\left(k_1, k_n\right)$ suffices for $p^* < \underline{p}$.

Thus, the above implies that if $\overline{m} \in \left[\underline{x}\left(k_1\right), \overline{x}\left(k_1, k_n\right)\right)$, such that $G\left(m^*\left(k_1, \overline{m}\right)\right) \in \left(0, \frac{K_{-n}}{K}\right)$, and if $\delta \geq \delta^*\left(k_1, k_n\right)$, then $\overline{V} = \frac{1}{K}\left(\frac{\widehat{m} - G(m^*)K}{1 - G(m^*)}\right)$ and $\underline{V} = \underline{p}$. ■

This and Proposition 2 imply that the maximal PPE payoffs for all i are the same as the optimal PPE payoffs under trigger-sales strategies, $k_i\overline{V} = k_i V^* \ \forall\ i$. Furthermore, the critical discount factor and the necessary condition on $G\left(m^*\left(k_1, \overline{m}\right)\right)$ are also the same. To understand the reason, note that the minimal PPE payoffs for all i are the same as the static Nash equilibrium profits, $k_i\underline{V} = \pi_i^N\left(k_n\right) \ \forall\ i$. Thus, in place of the randomisation device to construct the PPE strategy profile that supports $V^p = \overline{V} - \frac{1-\delta}{\delta}\left(1 - \overline{V}\right)$ in the proof of Proposition 6, we could instead have required firms to play the static Nash equilibrium for T periods, and then switch to the maximal PPE. In such a case, T must satisfy $k_i V^p = \left(1-\delta^T\right) \pi_i^N \left(k_n\right) + \delta^T k_i \overline{V}$, where after substiting in for V^p, $\pi_i^N \left(k_n\right)$ and $\overline{V}$, we find that the condition required for V^p to be a PPE is the same as in (19), except that $1-\delta^T$ replaces α. Then (19) is the same as (16) with $p^c = 1$, so the profile of trigger-sales strategies with $p^c = 1$ and $T = T^*$ is equivalent to the one that supports the maximal PPE payoffs in Proposition 6.