Optimal Collusion under Cost Asymmetry

Jeanine Thal*
Gremaq, Toulouse

Abstract

Cost asymmetry is generally thought to hinder collusion because a more efficient firm has both less to gain from collusion and less to fear from retaliation. This paper reexamines the conventional wisdom and characterizes optimal collusion between cost heterogeneous firms without any prior restriction on the class of strategies. We first illustrate how even the most efficient firm can be maximally punished by means of a simple stick-and-carrot punishment. This implies that some collusion is sustainable under cost asymmetry whenever collusion is sustainable under cost symmetry. Efficient collusion, however, is more difficult to sustain when costs are asymmetric. Finally, we show that, in the presence of side payments, cost asymmetry actually facilitates collusion.

keywords: horizontal collusion, cost asymmetry, optimal punishments, side payments
JEL codes: L41, C72

*I am grateful, above all, to my Ph.D. advisor Patrick Rey. This paper has also benefitted from comments by Mohammed Belhaj, Bruno Jullien, Ulrich Kamecke, Andreas Blume, Joe Harrington, Dan Kovenock, and participants of the 2004 Enter Jamboree, the 2004 EARIE meeting, the 2004 WZB/CEPR "Collusion and Cartels" conference, and the Econometric Society World Congress 2005. All remaining errors are my own. The author can be contacted for comments at thal@ensae.fr.

1 Introduction

Economists and policymakers generally agree that cost asymmetry hinders collusion. Checklists of factors hindering collusion, as used by antitrust authorities, typically include cost heterogeneity, and standard industrial organization textbooks teach this conventional wisdom. Scherer (1980), for example, states that "...the more cost functions differ from firm to firm, the more trouble firms will have maintaining a common price policy...". The underlying arguments usually fall within three categories. Firstly, coordination problems are obviously more complex when firms have divergent preferences concerning the collusive price. Secondly, it may be difficult to induce an efficient firm that earns high profits even under competition to participate in a cartel. Thirdly, cost asymmetry may also hinder the *sustainability* of collusion, since (i) it may be more difficult to retaliate against an efficient firm if it deviates from the cartel agreement, and (ii) a more efficient firm may have a higher relative short-term gain from deviation.

This paper focuses on the impact of cost asymmetry on the sustainability of collusion, thus ignoring issues of coordination and participation.[1] Our aim is to analyze the maximum scope for collusion. To this end, we first illustrate that it is possible to design tough punishments even for the most efficient firm: firms can "collude" on punishments that leave the cheating firm with its minmax profits, no matter whether the deviator has high or low costs. One class of such punishments have a stick-and carrot structure à la Abreu (1986, 1988). Thus, cost asymmetry weakens retaliation only if there is some rationale why firms should use standard trigger strategies or other restricted forms of punishments instead of these maximal punishments.

Whether a more efficient firm nonetheless has relatively stronger deviation incentives then depends on the level of the collusive price. Suppose that the cartel members, one low-cost and one high-cost firm who produce perfect substitutes, split the market equally. Then, if the collusive price is equal to the efficient firm's monopoly price, each firm can increase its short-term profits by at most 50% by undercutting the collu-

[1] Schmalensee (1987) applies a variety of selection criteria to model the choice of price and output quotas by an asymmetric cartel, subject to the constraint that each firm is at least as well off as without collusion. His paper, however, does not explicitly examine whether a selected outcome is also sustainable.

We, on the other hand, apply standard repeated game theory, which is well suited to study sustainabaility but not coordination or participation issues. In repeated games, there is generally a large multitude of equilibria and no uncontested method to select one of these. Firms might even be "locked" into a bad equilibrium, in which some firm's profits are lower than in the standard competitive equilibrium.

sive price. Short-term deviation incentives are hence symmetric. When punishments are indeed maximal for both firms, the key implication of this is that collusion is sustainable whenever it is sustainable under cost symmetry.

If the collusive price exceeds the efficient firm's monopoly price, however, the efficient firm can increase its short-term profits by more than 50% by deviating to its monopoly price. We show that efficient collusion, where firms allocate production in some Pareto-optimal way constrained only by the absence of side transfers, is therefore always more difficult to sustain under cost asymmetry than under cost symmetry. This is true both when firms optimally share the market in each period, and when they agree on an intertemporal scheme that grants temporary monopolies.

These conclusions differ from those in the previous literature that has focused on grim trigger straetgies. Bae (1987) as well as Harrington (1991), whose frameworks very closely ressemble ours, determine the set of price and output quotas sustainable by grim trigger strategies. They find that cost asymmetry always hinders collusion, even when allowing for inefficient allocations from the viewpoint of the cartel.

The other key contribution of our paper is to analyze the particular role of monetary side transfers when asymmetric firms collude. While direct transfers are typically prohibited by antitrust law, our position is that indirect forms of transfers between firms are conceivable. Firms may, for example, manipulate transfer prices in inter-firm dealings or disguise side payments in research joint ventures.[2] There is also evidence that agreements included transfers through inter-firm sales in a number of recently detected cartels, e.g. those in citric acid, in lysine, or in sodium gluconate.[3]

As Bain (1948) has argued more than 50 years ago, when firms have asymmetric marginal costs, the objective of maximizing joint industry profits makes sense only when side payments are possible. In the absence of transfers, total industry profits have to be reduced to induce all firms to stick to the collusive outcome. This suggests that side payments are particularly important for cartel stability when firms are heterogeneous.

We show that side payments do not only facilitate collusion between asymmetric firms, but also that in this case collusion is easier to sustain between cost asymmetric firms than between symmetric firms. The latter result completely contradicts the conventional wisdom on the impact of cost asymmetries. The underlying intuition is that transfers allow the less efficient firm to profit from the other firm's advantage. For-

[2]See Scherer (1980) or Tirole (1988).

[3]These examples are taken from Harrington & Skrzypacz (2005).

mally, the inefficient firm's short-term deviation gain is reduced. The low-cost firm's incentive to deviate, on the other hand, does not depend on whether it makes a side transfer to the inefficient firm or an equivalent market share concession.

To our best knowledge, our paper is the first to use optimal punishments without any restrictions on the class of strategies in a model with cost asymmetric firms. In frameworks different from ours, the authors who have characterized optimal punishments were so far bound to restrict the strategies considered.

Rothschild (1999) and Vasconcelos (2005) both deal with collusion under cost asymmetry when firms compete à la Cournot. Rothschild uses standard grim trigger strategies. Vasconcelos looks for more general punishments in the class of equilibria with proportional market shares on all equilibrium paths; he shows that optimal punishments, with a stick-and-carrot structure as proposed by Abreu (1986, 1988), exist within this restricted class of equilibria. For a limited range of parameters, these punishments are also maximal and would thus be optimal even without any restrictions.

In the related literature on collusion with asymmetric capacity constraints where firms compete in prices, the characterization of optimal punishments is unfortunately quite difficult. While Lambson (1987) shows that optimal punishments exist in models with symmetric capacity constraints, Lambson (1994) provides only a partial characterization in the asymmetric case. The impact of asymmetry in capacities on collusive sustainability was studied by Davidson & Deneckere (1990) in the context of grim trigger strategies. Compte, Jenny & Rey (2000) extend this analysis and allow for harsher punishments, but again restrict attention to a particular class of equilibria where market shares along any punishment are the same as under collusion.

Our analysis proceeds as follows. Section 2 sets out the framework. Section 3 discusses the use of optimal punishments in models of repeated price setting in which firms have asymmetric costs. We construct a class of subgame perfect maximal punishments, which do not involve any weakly dominated strategies, exist even for the low-cost firm. These punishments have a stick-and-carrot structure à la Abreu (1986,1988). Section 4 deals with stationary collusion without side payments. We first derive the sets of all sustainable collusive outcomes as a function of the discount factor. Next, we restrict attention to efficient collusion. Here, we distinguish between stationary collusion on a Pareto-efficient outcome, and fully efficient collusion on a lottery. We also derive the Pareto frontier of sustainable allocations, i.e. the subset of Pareto undominated allocations within the set of all sustainable allocations. In section 5, we allow for unlimited side payments. We again derive the set of all sustainable stationary allocations, and discuss the difference to the previous results. Section 6 concludes.

2 Framework

We consider a simple model of infinitely repeated Bertrand competition between two firms. The firms produce perfect substitutes, but one firm is more cost efficient than the other one. The low-cost firm produces at a constant marginal cost $\underline{c} > 0$, whereas the high-cost firm produces at a constant marginal cost $\overline{c} > \underline{c}$. There are no other relevant costs of production. The demand function for the good, $D(p)$, is continuous and strictly downward sloping with $\lim_{p \longrightarrow 0} D(p) = \infty$.

The monopoly profit functions of the low-cost and the high-cost firm, respectively, are:

$$\underline{\pi}(p) = (p - \underline{c})D(p),$$
$$\overline{\pi}(p) = (p - \overline{c})D(p).$$

These functions are assumed to be twice differentiable and strictly concave. Since $\overline{c} > \underline{c}$, $\underline{\pi}(p) > \overline{\pi}(p)$ for all p for which demand is positive. We denote the (unique) monopoly prices by $\underline{p}^m$ and $\overline{p}^m$. A standard revealed preferences argument ensures that $\underline{p}^m \leq \overline{p}^m$. Unless stated otherwise, it is assumed that the difference in marginal costs is non-drastic so that

$$\underline{p}^m > \overline{c}.$$

The firms compete in prices in the absence of any capacity constraints. Since the firms produce perfect substitutes, the market demand goes entirely to the low price firm if the two firms set different prices. In case of a price tie, total demand can be split between the two firms in any way consistent with the equilibrium.[4] We will, however, assume that consumer rationing is not feasible, so that the firms always have to serve the entire market demand. The market sharing rule can then be denoted by a single variable s, where s is the low-cost firm's market shares, and $(1 - s)$ is the high-cost firm's market share.

In this set-up, we analyze a noncooperative supergame, where in each period, the two firms simultaneously set prices and, if necessary, decide on a market sharing rule. Prices are publicly observable, and firms have perfect memory; they can thus condition their actions on past prices. Each firm aims to maximize its discounted profit stream. The common

[4]In symmetric models, it is often assumed for simplicity that the market is split equally between the two firms. This is, however, an arbitrary restriction, and may rule out interesting equilibria when firms are asymmeric.

and constant discount rate is $\delta \in (0, 1)$. We look for a subgame perfect equilibrium.

3 Optimal Punishments

Deterring deviations from collusive (i.e. supracompetitive) prices requires a credible punishment mechanism. We will focus on optimal punishments, i.e. the harshest sustainable punishments. This permits the characterization of the maximum scope for collusion: if firms employed non-optimal punishments, some outcomes sustainable under optimal punishments would no longer be sustainable.

The security level, that is lowest, profit that can be imposed on a firm is its minmax, which is zero for both firms here: in the absence of any restrictions on prices, each firm can indeed drive the other firm's profits down to zero by undercutting its price. We will therefore say that a punishment is maximal whenever the punished firm's continuation value along the punishment path is equal to zero. Obviously, if any maximal punishment is sustainable, then it will constitute an optimal punishment, since no other punishment can possibly be harsher.

A punishment path for a firm consists of a sequence of prices and market sharing rules $(\underline{p}_t, \overline{p}_t, s_t)$, from the first period after deviation, $t = 1$, until infinity. Let us denote the punishment paths of the two firms by $\underline{\tau}$ and $\overline{\tau}$, respectively. The corresponding continuation values for the punished firms are $\underline{v}$ and $\overline{v}$. Starting from the collusive outcome, the two firms start playing punishment τ_i if firm i deviates from the collusive path.[5] If a firm deviates from the prescribed action at any stage of a punishment path, its own punishment is started (or restarted, respectively). A punishment path is credible (i.e. subgame perfect) if neither of the two firms has an incentive to deviate from the prescribed punishment path at any stage $t \geq 1$.

In a Bertrand model with symmetric firms, reversion to the competitive equilibrium in which both firms charge their common marginal cost constitutes a maximal and thus optimal punishment. Reversion drives the deviating firm's (and all other firms') profits down to zero, the minmax value. It is also obvious that the infinite repetition of the static equilibrium is a credible threat. Under cost asymmetry, the standard static Nash equilibrium is such that both firms charge price $\overline{c}$, and the low-cost firm makes all the sales earning $\underline{\pi}(\overline{c}) > 0$.[6] Thus, a grim trig-

[5] If both firms deviate simultaneously, no punishment is played.

[6] There are different ways to formally "justify" this equilibrium without making appeal to the extreme assumption that the low-cost firm obtains the whole market when both firms charge the equilibrium price.

The general idea that is often expressed is that the low-cost firm charges $\overline{c} - \varepsilon$

ger strategy where firms revert to the competitive equilibrium forever, provides a maximal and optimal punishment for the high-cost firm, but not for the low-cost firm.

We now show that it is easy to design equally tough punishments for the low-cost firm, too. We consider two possible types of maximal punishments: (i) reversion to a different static equilibrium, in which the common price is $\underline{c}$ and $s = 1$, and (ii) stick-and carrot punishments as in Abreu (1986, 1988). Both options constitute optimal punishments, but the first one involves a weakly dominated strategy, whereas an appropriate stick-and carrot punishment does not; therefore, a stick-and carrot punishment appears preferable.

Reversion to a Static Equilibrium Consider a situation in which both firms charge the same price $\underline{c}$, and the low-cost firm makes all the sales, so that both firms earn zero profits. This is a static equilibrium: none of the firms has a (strict) incentive to charge a different price given the other firm's price and market shares.[7] Infinite repetition of this static equilibrium thus constitutes a maximal punishment for both firms. However, this punishment path has an undesirable feature: the high-cost firm plays a weakly dominated strategy. Given *any* strategy of the low-cost firm, the high-cost firm could do equally well or strictly better by charging $\overline{c}$ in each period.

Stick-and-Carrot Punishments It is rather straightforward to construct other punishment paths that leave the low-cost firm with a continuation payoff of zero and that do not involve any weakly dominated strategy. These punishments have a stick-and-carrot structure as in Abreu (1986, 1988): the punished firm makes negative profits for a number of periods before both firms return to a collusive outcome. In the following,

---

where ε is an arbitrarily small positive number (see Tirole (1988)). The "openness problem" of this approach can be solved by discretizing the strategy space.

Blume (2003) offers a different very elegant solution. He shows that, for small enough $\eta > 0$, there is an equilibrium where the low-cost firm charges $\overline{c}$ and the high-cost firm randomizes uniformly over $[\overline{c}, \overline{c} + \eta]$. In fact, any price in $(\underline{c}, \overline{c}]$ can be supported by similar strategies. Only the competitive price $\overline{c}$, however, can be supported by an equilibrium in undominated strategies.

[7]Here, we make appeal to the assumption of an extreme market sharing rule in case both firms charge $\underline{c}$. Blume (2003), however, shows that even without this assumption any price in $(\underline{c}, \overline{c}]$ can be supported in equilibrium when the high-cost firms plays a mixed strategy (see the previous footnote). There hence exist equilibria with prices arbitrarily close to $\underline{c}$ in which the low-cost firm is the only seller, but that do not rely on an extreme market sharing rule assumption.

The infinum of the low-cost firm's punishment payoff achievable by infinite repetition of a static equilibrium is thus zero, even without any assumption on the market sharing rule.

we will characterize such punishments, show that they are credible, and finally prove that they do not involve any weakly dominated strategies.

Lemma 1 *Suppose there exists a collusive path with common price $p^* > \overline{c}$ and market sharing rule s^* in each period that can be supported by maximal punishments. Then, any punishment path that has the following structure*

$$\underline{\tau} : (\underline{p}_t, \overline{p}_t, s_t) = \begin{cases} (p^P, p^P + \varepsilon, 1) \, for \; t \in [1, T] \\ (p^*, p^*, s^*) \qquad for \; t > T \end{cases},$$

where $\varepsilon \in \left[0, \underline{c} - p^P\right)$, and $p^P < \underline{c}$ and T are such that

$$\underline{v} = \sum_{t=1}^{T} \delta^{t-1} \underline{\pi}(p^P) + \frac{\delta^T}{1-\delta} s^* \underline{\pi}(p^*) = 0, \tag{1}$$

provides an optimal (i.e., maximal and subgame perfect) punishment for the low-cost firm. Moreover, such a punishment path does not involve any weakly dominated strategies.

Proof. Note first that, since the demand is continuous with $\lim_{p \longrightarrow 0} D(p) = \infty$, and $\underline{c} > 0$, an exact solution $p^P < \underline{c}$ to (1) exists for any T.

By construction, any punishment such as constructed in lemma 1 is maximal since $\underline{v} = 0$.

We now show that any such punishment is also subgame perfect. To show that neither firm has a strict incentive to deviate at any stage of the punishment path, it is sufficient to consider unilateral one-shot deviations. By construction, no firm has an incentive to deviate for $t > T$, since the collusive equilibrium is played. Also, a firm has no incentive to deviate in any period $t \in [2, T]$ if it has no incentive to deviate at $t = 1$: The short-term gains from a deviation are the same at any stage $t \in [1, T]$, whereas the cost of foregoing the future return to collusion increases with t. It is thus sufficient to show that no firm wants to deviate at $t = 1$.

The low-cost firm's best deviation is to charge a price above $p^P + \varepsilon$ and earn zero instead of negative profits. After this deviation, $\underline{\tau}$ would be restarted yielding zero discounted profits to the low-cost firm. The low-cost firm therefore has no strict incentive to deviate since

$$0 + \delta \underline{v} = 0 \leq \underline{v} = 0.$$

Finally, the high-cost firm cannot benefit from deviating at $t = 1$: it would not generate any gain in that period, but nevertheless trigger the punishment $\overline{\tau}$ instead of eventually returning collusion. Thus, the high-cost firm has a strict incentive *not* to deviate:

$$0 + \delta\overline{v} = 0 < 0 + \frac{\delta^T}{1-\delta}(1-s^*)\overline{\pi}(p^*).$$

It remains to be shown that $\underline{\tau}$ does not involve any weakly dominated strategies. We will use the following simple argument: a strategy is *not* weakly dominated if it is the unique best reply to *some* strategy of the other firm.

The high cost firm's strategy along $\underline{\tau}$ is obviously undominated for $t \leq T$ because it is the strict best reply to the low-cost firm's strategy along $\underline{\tau}$. The low-cost firm however is indifferent between compliance and a one-shot deviation at $t = 1$. Thus, an alternative strategy in which the low-cost firm deviates in the first punishment period could potentially weakly dominate compliance. To show that this is not the case, consider a strategy of the high-cost firm that consists of respecting $\underline{\tau}$ in all periods but the first one. At $t = 1$, the high-cost firm charges a price below the stick price p^P. In the following periods, the high-cost firm conditions its actions on whether the low-cost firm complied with its own punishment in the first period or not. It continues with the second stage of $\underline{\tau}$ after compliance, but restarts the punishment $\underline{\tau}$ from the beginning again after a deviation. Given this strategy of the high-cost firm, the strict best reply of the low-cost firm is compliance in all periods. Hence, no other strategy can weakly dominate compliance.

We still need to show that compliance remains an undominated strategy at later stages of the punishment path ($t > T$), i.e. along the collusive path. To see this, simply note that compliance is the unique best reply to a strategy that consists of charging a price slightly above the collusive price in the current period and then returning to the collusive outcome if and only if the other firm has not deviated.

Consequently, the punishment path $\underline{\tau}$ does not involve any weakly dominated strategy. ■

While our focus was on constructing a strick-and-carrot punishment for the efficient firm, it should be clear that punishments as constructed in lemma 1 also exist for the high-cost firm.

Our analysis of optimal stick-and-carrot punishments is also robust to introducing some product differentiation. Consider a punishment of the low-cost firm as constructed in lemma 1. When products are perfect substitutes, $p^P + \varepsilon < \underline{c}$ is a sufficient condition to rule out deviations by both firms. When products are differentiated, however, the high-cost firm's stick price must lie far enough below $\underline{c}$ such that the low-cost firm could not attract any consumers when deviating to some price above $\underline{c}$.

To rule out deviations by the high-cost firm on the other hand, the price difference ε should be sufficiently large to avoid considerable negative profits by the high-cost firm during the stick phase of the punishment. To construct a punishment, we thus simply need to set the high-cost firm's stick price sufficiently far below $\underline{c}$, and then allow for a sufficiently large difference between the two firms' stick prices.[8]

Our analysis can also easily be extended to situations with three or more cost asymmetric firms. In the standard static Nash equilibrium, which is a perfect equilibrium, the most efficient firm then charges a price equal to the second lowest marginal cost and serves the whole market, while all other firms charge their own marginal costs. Reversion to this equilibrium is thus an optimal punishment for all firms but the most efficient one. For the most efficient firm, a stick-and-carrot punishment of the kind introduced in lemma 1 again provides an optimal punishment that does not involve any weakly dominated strategy.

Finally, let us make a few remarks regarding the "plausibility" of optimal stick-and-carrot punishments as compared to standard grim trigger strategies. Stick-and-carrot punishments have a simple structure: upon deviation, firms first vigorously fight against each other for some time before returning to a cooperative equilibrium. Such a scenario seems more natural than an eternal return to a non-cooperative outcome. In particular, stick-and-carrot punishments are not prone to renegotiation anymore once the stick phase is over, while firms would want to renegotiate a grim trigger strategy in each and every period. Moreover, the punishments we constructed are maximal for the deviating firm, but much less costly for all the other firms. In this sense, the stick-and carrot punishments are truly targeted at the deviator, while a grim trigger strategy punishes all firms. Intuitively, this targeting ability is a desirable characteristic of any punishment.

Our punishments are subgame perfect and thus self-enforcing. Nevertheless, one might feel that coordinating on optimal stick-and-carrot punishments (as well as on some collusive price and a market sharing rule), an issue which is not modelled here, requires some explicit communication between the firms.

4 Collusion without Side Payments

We now turn to the question which agreements firms can sustain by optimal punishment. We first derive the set of all sustainable stationary

[8]Note that the low-cost firm's profits in a single stick period constructed this way might already be so low that continuation profits are negative. In this case, it is necessary to introduce some randomisation, such that firms play the stick phase only with a probability below 1.

collusive outcomes. Next, we restrict attention to allocations that are efficient in the sense of being Pareto-optimal for the firms. In this context, it is important to distinguish between collusive agreements such that both firms share the market in each period, and lotteries that grant temporary monopoly positions, since firms are able to achieve higher expected profits with lotteries. We then check when a given efficient agreement, be it deterministic or a lottery, is indeed sustainable. Finally, we analyze what we call the constrained Pareto frontier, i.e. the subset of undominated allocations (for the firms) within the set of sustainable (stationary) outcomes.

4.1 Sustainability

A collusive allocation is defined by a couple (p, s), where $p \in (\overline{c}, \overline{p}^m]$[9] is the common collusive price and s the market sharing rule. Here, we thus focus on the sustainability of stationary equilibria in which the price and market shares are constant over time. In appendix 1, we show that this is not restrictive, since permitting collusive arrangements that vary over time cannot facilitate collusion in the sense of reducing the lowest critical discount factor. Also note that correlated equilibria, where firms collude on a lottery assigning the monopoly to one of the firms in each period, do not facilitate collusion either, because the deviation incentives of the firm that loses the lottery in any given period are generally high. We will study Pareto-optimal correlated equilibria later when looking at fully efficient collusion.

A collusive outcome is sustainable if and only if it can be induced by punishment strategies which form a subgame perfect equilibrium. We will characterize all the collusive outcomes sustainable by maximal punishments ($\overline{v} = \underline{v} = 0$). Since the optimal punishments constructed in the previous section are indeed maximal, our analysis is not restricted to any particular class of strategies.

The high-cost firm could deviate from the collusive outcome by slightly undercutting the collusive price in order to capture the whole market. Thus, it has no incentive to deviate from the collusive outcome if

$$\overline{\pi}(p) \leq \frac{1}{1-\delta}(1-s)\overline{\pi}(p)$$

For the low-cost firm, the optimal deviation is to charge $\underline{p}^m$ if the collusive price lies above $\underline{p}^m$, and to undercut its rival otherwise. The low-cost

[9] The firms will never find it optimal to price above $\overline{p}^m$, since higher prices render collusion more difficult but a Pareto improvement could be achieved by moving to $\overline{p}^m$.

firm has therefore no incentive to deviate if

$$\underline{\pi}\left(\min[p, \underline{p}^m]\right) \leq \frac{1}{1-\delta} s\underline{\pi}(p).$$

These conditions are equivalent to

$$\delta \geq s \tag{$\overline{C}$}$$

and

$$s \geq (1-\delta)\frac{\underline{\pi}\left(\min[p, \underline{p}^m]\right)}{\underline{\pi}(p)}, \tag{$\underline{C}$}$$

respectively.

A collusive outcome (p, s) is thus sustainable by maximal punishments if and only if it satisfies conditions $\overline{C}$ and $\underline{C}$. In particular, collusion at price p is sustainable for some market share(s) if and only if

$$\delta \geq \widetilde{\delta}(p),$$

where the discount factor threshold $\widetilde{\delta}(p)$ can be found by adding up $\overline{C}$ and $\underline{C}$ when both are binding:

$$\widetilde{\delta}(p) \equiv \frac{\underline{\pi}\left(\min[p, \underline{p}^m]\right)}{\underline{\pi}(p) + \underline{\pi}\left(\min[p, \underline{p}^m]\right)}.$$

For collusive prices $p \in \left(\overline{c}, \underline{p}^m\right]$, the critical discount factor is $\widetilde{\delta}(p) = \frac{1}{2}$, as under cost symmetry. This result arises because in this case each firm's deviation would consist in slightly undercutting its rival, and both firms' punishments impose zero continuation profits. Each firm's deviation incentives then only depend on its market share relative to the discount factor, and the non-deviation constraints are symmetric:

$$1 - s \geq 1 - \delta, \tag{$\overline{C}'$}$$

$$s \geq 1 - \delta. \tag{$\underline{C}'$}$$

The conditions for collusive sustainability are thus exactly the same as under cost symmetry.

For $p \in \left(\underline{p}^m, \overline{p}^m\right]$, on the other hand, the discount factor threshold $\widetilde{\delta}(p)$ strictly increases with p. This result is driven by the wedge between the low-cost firm's stand-alone collusive profits $\underline{\pi}(p)$ and its deviation profits $\underline{\pi}(\underline{p}^m)$. Given any market sharing rule, the low-cost firm's incentive to deviate is clearly higher the larger the difference between the collusive price and its own monopoly price.

The market sharing rules such that collusion at price p is indeed sustainable for some discount factor $\delta \geq \widetilde{\delta}(p)$ are $s \in [\widetilde{s}(p,\delta), \delta]$, where

$$\widetilde{s}(p,\delta) \equiv (1-\delta)\frac{\underline{\pi}\left(\min[p, \underline{p}^m]\right)}{\underline{\pi}(p)}.$$

The lower and upper bounds on s are easily derived from the non-deviation constraints. On the one hand, the low-cost firm only complies with the collusive arrangement if its market share s is sufficiently high; its minimum market share $\widetilde{s}(p,\delta)$ is thus such that its non-deviation constraint $\underline{C}$ is binding. On the other hand, the high-cost firm demands a market share $(1-s)$ at least equal to $1-\delta$, which yields an upper bound of δ on the market sharing rule s. For $p \in \left(\overline{c}, \underline{p}^m\right]$, when the non-deviation constraints are independent of price, the range of sustainable sharing rule is simply $[1-\delta, \delta]$, as under cost symmetry. For prices between the two firms' monopoly prices, the lower bound on the low-cost firm's market share, $\widetilde{s}(p,\delta)$, strictly increases with the price to accommodate the low-cost firm's increasing deviation incentives.

The set of sustainable allocations, which we will denote by $\Delta(\delta)$, includes all allocations $(p,s) \in (\overline{c}, \overline{p}^m] \times [0,1]$ such that $\widetilde{\delta}(p) \leq \delta$ and $s \in [\widetilde{s}(p,\delta), \delta]$. Figure 1 provides a graphical representation of this set for different discount factors. As under cost symmetry, only equal market sharing rules are sustainable for $\delta = \frac{1}{2}$: $\Delta(\frac{1}{2}) = \left\{(p,s) \mid p \in \left(\overline{c}, \underline{p}^m\right], s = \frac{1}{2}\right\}$. For $\delta_1 > \frac{1}{2}$, $\Delta(\delta_1)$ includes all outcomes left of or on the line labelled $\overline{C}(\delta_1)$, along which the high-cost firm is indifferent between complying and deviating, as well as right of or on the line labelled $\underline{C}(\delta_1)$, along which the low-cost firm is indifferent between deviating and complying. As long as the maximum collusive price, p_1, remains below the high-cost firm's monopoly price, both non-deviation constraints must be binding at that price: $p_1 = \widetilde{\delta}^{-1}(\delta_1)$. Clearly, as the discount factor increases, the set of sustainable allocations becomes larger and larger. First, the maximum sustainable price, $\widetilde{\delta}^{-1}(\delta)$, increases with the discount factor. Second, for any given sustainable price, the range of sustainable market sharing rules expands in both directions as the discount factor rises.

Drastic Cost Difference In the preceding analysis, we assumed that the cost difference was sufficiently small so that $\overline{c} < \underline{p}^m$. If instead the low-cost firm enjoys a drastic cost advantage, it charges its monopoly price $\underline{p}^m < \overline{c}$ and earns profits $\underline{\pi}(\underline{p}^m)$ in the static equilibrium. Using grim trigger strategies, there is therefore no sustainable collusive outcome.

Once we consider optimal punishments, however, collusion is sustainable even in the presence of a drastic cost difference. First note that

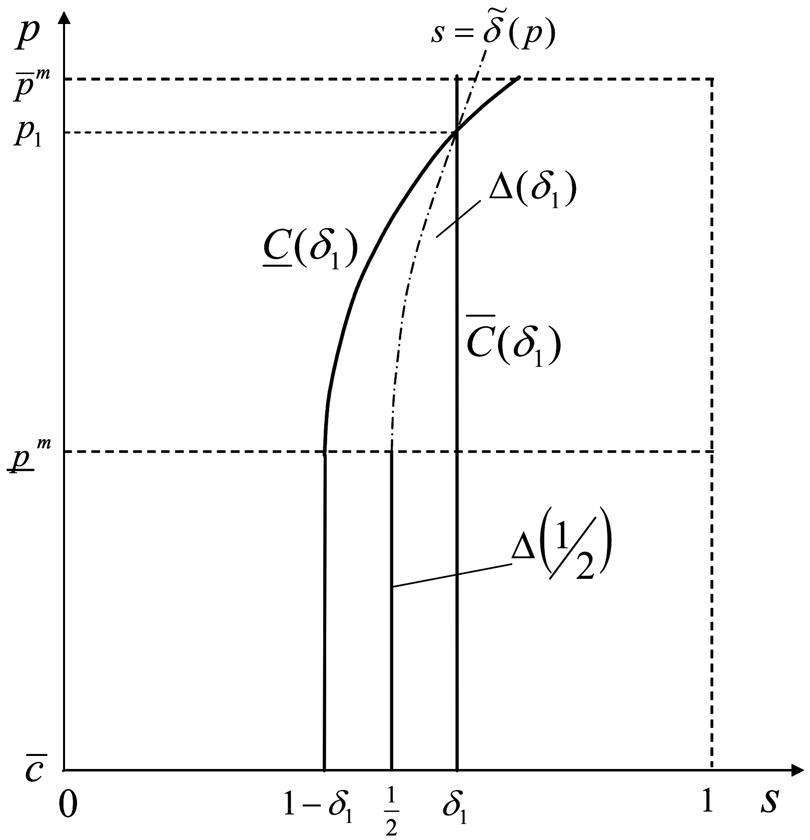

Figure 1: Sustainable Collusive Allocations without Side Payments

lemma 1, which guarantees the existence of a subgame perfect undominated maximal punishment for the low-cost firm, holds whenever the collusive price exceeds $\overline{c}$, irrespective of the size of the cost difference. The characterization of the set of sustainable collusive allocations then remains unchanged, except that the range of collusive price, $(\overline{c}, \overline{p}^m]$, is strictly included in $[\underline{p}^m, \overline{p}^m]$ now. The low-cost firm's optimal deviation is therefore always to charge its monopoly price, and collusion is only possible for discount factors at least equal to $\frac{\underline{\pi}(\underline{p}^m)}{\underline{\pi}(\overline{c})+\underline{\pi}(\underline{p}^m)} > \frac{1}{2}$.

Using optimal punishments, collusion is thus sustainable even if the cost difference is drastic. However, collusion is more difficult to sustain than when the cost difference is non-drastic, because the collusive price must be relatively high (strictly above $\underline{p}^m$), which requires higher discount factors. Note also that whenever collusion takes place in the presence of a drastic cost difference, this means that the efficient firm is locked into a "bad" equilibrium where its profits are lower than in the competitive equilibrium.

4.2 Efficient Collusion

In this section, we consider the sustainability of allocations that are Pareto-optimal for the firms. We first restrict attention to stationary allocations, and show that it is more difficult to sustain efficient allocations under cost asymmetry than under cost symmetry. Then we drop the restriction of stationarity, which is a serious one here because firms can improve Pareto efficiency by taking turns being the monopolist. We therefore also analyze the sustainability of collusion on a lottery for the monopoly position, and show that efficient collusion remains more difficult than under cost symmetry.

4.2.1 Pareto-Efficient Production

Let us first analyze the efficient allocation of production in the absence of side payments, neglecting the issue of collusive sustainability. We consider the following problem:

$$\max_{\{p,s\}} \left[s\underline{\pi}(p)\right]^{\alpha} \left[(1-s)\overline{\pi}(p)\right]^{1-\alpha}, \ \alpha \in [0,1]. \qquad (P1)$$

Solving $(P1)$ for every $\alpha \in [0,1]$ yields a simple characterization of all Pareto-efficient outcomes for the firms.[10] The solution for each α is characterized by:

$$s = \alpha, \qquad (2)$$

and

$$-\alpha\frac{\underline{\pi}'(p)}{\underline{\pi}(p)} = (1-\alpha)\frac{\overline{\pi}'(p)}{\overline{\pi}(p)}. \qquad (3)$$

[10] See exercise 6.1 in Tirole (1988) for a detailed treatment of an equivalent problem.

As α varies between 0 and 1, the optimal market sharing rule varies between 0 and 1, and the optimal price between the two firms' monopoly prices. The two firms' isoprofit lines are tangent at any Pareto optimum.

Combining (2) and (3) yields the following one-to-one relationship between the low-cost firm's market share and the price:

$$s^O(p) = \frac{(\overline{c} - \underline{c})D(p) + (p - \overline{c})\underline{\pi}'(p)}{(\overline{c} - \underline{c})D(p)}, p \in [\underline{p}^m, \overline{p}^m]. \tag{4}$$

As can be checked easily, $s^O(\cdot)$ is strictly downward-sloping.[11] The inverse of $s^O(\cdot)$ will be denoted by $p^O(\cdot)$.

4.2.2 Stationary Collusion on Pareto-Efficient Outcomes

Let us now analyze the sustainability of Pareto-efficient outcomes as characterized by condition (4).

It is easy to check that the minimum discount factor for which *some* Pareto-efficient outcome is sustainable must be such that both firms are indifferent between colluding and deviating. To see this, suppose first that some allocation $(p^O(s), s)$ is sustainable at discount factor δ, but (at least) one firm strictly prefers compliance. Then, by continuity, there exists another Pareto-efficient allocation sustainable at lower discount factors. First, firms can move along the Pareto frontier, into the direction preferred by the firm with the binding non-deviation constraint, to an allocation at which both firms strictly prefer compliance. Second, if both firms strictly prefer compliance at δ, they will also be willing to collude at a slightly lower discount factor. Alternatively, suppose that both firms are indeed indifferent between deviating and complying from $(p^O(s), s)$ for discount factor δ. By definition, any Pareto-efficient allocation different from $(p^O(s), s)$ is strictly better for one of the firms but strictly worse for the other firm. Thus, for δ, one of the firms would deviate from any Pareto-optimal allocation other than $(p^O(s), s)$. We can conclude that no other Pareto-optimal allocation is sustainable for the same (or a lower) discount factor.

Figure 2 pictures the minimum discount factor $\widehat{\delta}$ for which some efficient collusion is sustainable. The corresponding allocation is denoted by $(\widehat{p}, \widehat{s})$. As just explained, both firms must be indifferent between colluding and deviating. Since both non-deviation constraints are binding at price $\widehat{p}$ only if $\delta = \widetilde{\delta}(\widehat{p})$, and for this discount factor only the market sharing rule $\widetilde{\delta}(\widehat{p})$ is sustainable, it must be that $\widehat{s} = \widetilde{\delta}(\widehat{p})$. Moreover, for $(\widehat{p}, \widehat{s})$ to be Pareto-efficient, it must lie on $s^O(p)$. Since $\widetilde{\delta}(p)$ strictly

[11]It is also strictly concave if $\underline{\pi}'''(p) < 0$ and $Q''(p) < 0$ or these derivatives are positive but not "too" large. However, concavity of $s^O(p)$ is not needed for the results of the following analysis.

increases in p, while $s^O(p)$ strictly decreases in p, there indeed exists a unique allocation $(\widehat{p}, \widehat{s})$ which is both Pareto-efficient and such that both firms are indifferent between colluding and deviating. Graphically, the allocation $(\widehat{p}, \widehat{s})$ can thus be found where $s^O(p)$ cuts $\widetilde{\delta}(p)$, and the corresponding discount factor threshold is $\widehat{\delta} = \widehat{s}$. For discount factors above $\widehat{\delta}$, a whole range of efficient allocations is sustainable. This is illustrated in the figure for $\delta_1 > \widehat{\delta}$.

The following proposition summarizes these results (a formal proof is delegated to appendix 2):

Proposition 1 *Let* $\widehat{p} \in (\underline{p}^m, \overline{p}^m)$ *be uniquely defined by* $s^O(\widehat{p}) = \widetilde{\delta}(\widehat{p})$, *and let* $\widehat{\delta} \equiv \widetilde{\delta}(\widehat{p})$. *Moreover, let* $\widetilde{p}^O(\delta)$ *be the function implicitly defined by* $s^O(\widetilde{p}^O(\delta)) = \widetilde{s}(\widetilde{p}^O(\delta), \delta)$, *and note that* $\widehat{p} = \widetilde{p}^O(\widehat{\delta})$. *Then,*

(i) *for* $\delta < \widehat{\delta}$, *no Pareto-efficient allocation is sustainable.*

(ii) *for* $\delta \geq \widehat{\delta}$, *all Pareto-efficient allocations* $\left(p^O(s), s\right)$ *with market sharing rules* s *in the range* $\left[s^O(\widetilde{p}^O(\delta)), \delta\right]$ *are sustainable. In particular, for* $\delta = \widehat{\delta}$, *the unique sustainable Pareto-efficient allocation is* $\left(\widehat{p}, \widehat{\delta}\right)$.

Under cost symmetry, the discount factor for some efficient collusion is $\frac{1}{2}$: collusion on the common monopoly price is possible for any discount factor above this threshold if firms split the market equally. Since $\widehat{\delta} > \frac{1}{2}$, cost asymmetry thus hinders the sustainability of efficient collusion.

4.2.3 Fully Efficient Collusion on a Lottery

So far, we have analyzed collusion on efficient allocations under the restriction of stationarity. It is however straightforward to see that the firms could achieve a Pareto improvement by taking turns being the monopolist. Take any efficient stationary allocation $\left(p^O(s), s\right)$. Then, the firms' per-period profits are $s\underline{\pi}\left(p^O(s)\right)$ and $(1-s)\overline{\pi}\left(p^O(s)\right)$. Alternatively, let the low-cost firm be a monopolist with probability s, and the high-cost firm with probability $(1-s)$. Now, the low-cost firm's expected per-period profits are $s\underline{\pi}\left(\underline{p}^m\right)$, and the high-cost firm's expected per-period profits are $(1-s)\overline{\pi}\left(\overline{p}^m\right)$. Thus, both firms are better-off in terms of expected profits.[12] In other words, the Pareto profit frontier of problem $(P1)$ is convex, and fully efficient collusion must allow for alternating or random monopolies.

[12] We suppose that firms are risk-neutral, and thus only care about expected profits.

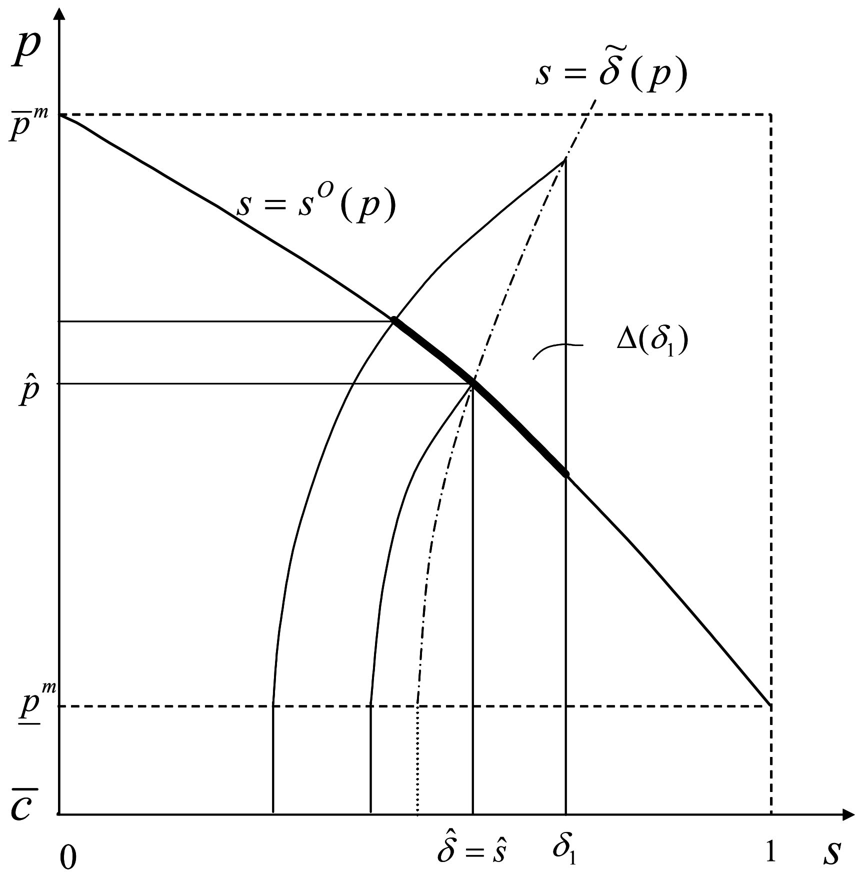

Figure 2: Efficient stationary collusion

To study fully efficient collusion, we use the concept of correlated equilibrium. That is, we will suppose that by throwing a dice or observing some common signal at the beginning of each period, the firms can assign the monopoly to the low-cost firm with probability α, and to the high-cost firm with probability $(1-\alpha)$. The probability α is part of the collusive arrangement, much as the market share earlier on. If the low-cost firm is the monopolist, it charges its monopoly price $\underline{p}^m$ and serves the whole market, while the high-cost firm charges (slightly more than) $\underline{p}^m$ and makes no sales at all. If the high-cost firm is the monopolist, it charges $\overline{p}^m$ and serves the whole market, while the high-cost firm charges (slightly more than) $\overline{p}^m$ and makes no sales at all. This way, the firms can achieve any point on the unconstrained linear Pareto profit frontier by letting α vary between zero and one. There is no loss of generality in restricting attention to lotteries with a stationary α.[13]

It is easy to generalize the optimal stick-and-carrot punishments to allow for lotteries along the collusive path. The stick price as well as the length of the stick period simply need to be such that the low-cost firm's *expected* continuation profits are zero. Consequently, the optimal punishments as maximal again.

Obviously, each firm's incentives to deviate are strongest when its rival holds the monopoly position. Therefore, the low-cost firm never wants to deviate from an efficient allocation if

$$\underline{\pi}(\underline{p}^m) \leq 0 + \frac{\delta}{1-\delta}\alpha\underline{\pi}(\underline{p}^m).$$

Similarly, the high-cost firm has no incentive to deviate if

$$\overline{\pi}(\underline{p}^m) \leq 0 + \frac{\delta}{1-\delta}(1-\alpha)\overline{\pi}(\overline{p}^m).$$

[13]To see this, suppose firms collude on a path of probabilities $\{\alpha_t\}_{t=0}^{\infty}$ from today $(t=0)$ to infinity. The probabilities α_t can also be interpreted as expected probabilities when firms collude only on the probability distribution(s) from which the α_ts are drawn. Define

$$\alpha \equiv \frac{(1-\delta)}{\delta}\sum_{T=\widehat{t}+1}^{\infty}\delta^{T-\widehat{t}}\alpha_T$$

as the constant probability which would yield the same continuation payoff from time $\widehat{t}$ onwards as the sequence $\{\alpha_t\}_{t=\widehat{t}}^{\infty}$. Then, the non-deviation constraints of the non-stationary lottery at time $\widehat{t}$ coincide with the non-deviation constraints of a stationary lottery with probability α. Since the non-deviation constraints must be satisfied in every time period, this implies that there is no loss of generality from considering only stationary lotteries.

Adding up these constraints and rearranging terms shows that efficient collusion can be sustained if the discount factor satisfies

$$\delta \geq \frac{\overline{\pi}(\overline{p}^m) + \overline{\pi}(\underline{p}^m)}{2\overline{\pi}(\overline{p}^m) + \overline{\pi}(\underline{p}^m)} > \frac{1}{2}. \tag{5}$$

Hence, fully efficient collusion requires discount factors above the threshold for efficient collusion under cost symmetry, in which case no randomisation is needed to achieve full efficiency.

4.3 The Pareto Frontier of Sustainable Allocations

We now analyze the subset of Pareto undominated allocations within the set of sustainable outcomes for each discount factor. Unlike in the previous section, we do not ask when a given efficient allocation is sustainable, but rather which of the sustainable allocations for a given discount factor are undominated. This approach takes account of the methodological point, underlined by Harrington (1991), that an allocation only provides a sensible collusive outcome if it is indeed implementable by a self-enforcing agreement. By restricting attention to the set of sustainable collusive equilibria a priori, the firms automatically solve this implementation problem.

Proposition 2 *The set of Pareto-undominated sustainable allocations is*

$$\Omega(\delta) = \Delta(\delta) \cap \left[\left((p,s) \mid (p \in \left[\underline{p}^m, p^O(\delta)\right), s = \delta) \right) \cup \left((p,s) \mid p = p^O(s), s \in [0,1] \right) \right].$$

Proof. First of all, unconstrained Pareto optimal allocations are obviously undominated if sustainable. For $\delta \geq \widehat{\delta}$, the set of Pareto undominated sustainable allocations therefore always includes part of the Pareto frontier $p^O(s)$.

Note also that any allocation (p,s) in $\Delta(\delta)$ with $p < \underline{p}^m$ is Pareto dominated by the allocation $(\underline{p}^m, s)$, which is also included in $\Delta(\delta)$.

Therefore, we can restrict attention to sustainable allocations with prices at least equal to the low-cost firm's monopoly price. For such prices, the low-cost firm's isoprofit lines are strictly increasing and concave in the (s,p) space. In fact, the isoprofit line for profit level $\Pi = (1-\delta)\underline{\pi}(\underline{p}^m)$ coincides with $\underline{C}(\delta)$. Profit levels are increasing in the southeast direction, as the low-cost firm prefers a higher market share s and prices closer to its own monopoly price.

The high-cost firm's isoprofit lines are increasing and convex in the (s,p) space. For allocations below $p^O(s)$ they are flatter than, for allocations on $p^O(s)$ tangent to, and for allocation above $p^O(s)$ steeper than the isoprofit lines of the low-cost firm. Moreover, profit levels are increasing in the northwest direction, as the high-cost firm prefers a higher market share $(1 - s)$ and higher prices.

Given this, it is straightforward that any sustainable allocation strictly above $p^O(s)$ is Pareto-dominated: moving along the low-cost firm's isoprofit curve towards $p^O(s)$ always increases the high-cost firm's profits without hindering collusive sustainability. Now consider any allocation strictly below $p^O(s)$. If firms are able to move northeast along the low-cost firm's isoprofit line without violating sustainability, a Pareto improvement within $\Delta(\delta)$ can be achieved: the high-cost firm is strictly better off thanks to the price increase although its market share $(1-s)$ is lower. The only sustainable allocations strictly below $p^O(s)$ that are not dominated are then those for which the high-cost firm's non-deviation constraint is binding, i.e. $s = \delta$, so that no further northeast moves are feasible. ■

Undominated sustainable allocations thus either lie on the high-cost firm's non-deviation constraint or/and are unconstrained Pareto optima. In the former case, prices lie between the low-cost firm's monopoly price and $p^O(\delta)$. Figure 3 illustrates the sets of Pareto-undominated sustainable allocations for two different discount factors, δ_1 and δ_2, one below and one above $\widehat{\delta}$.

5 Collusion with Side Payments

Side payments are often ruled out in the theory of collusion,[14] since antitrust law forbids overt monetary transfers between firms in most jurisdictions. Nonetheless, our stance is that side payments can occur in reality in various disguised forms.

Firstly, some indirect forms of monetary tranfers are conceivable. Competitors that are also business partners, for example, can quite easily manipulate transfer prices. Joint ventures may also serve as a vehicle for such transfers. Moreover, many industries exhibit a high degree of cross-ownership between firms. Firms are thus often direct claimants to a share of the profits of one of their rivals.[15] Finally, payments in kind may sometimes be another possibility to avoid antitrust suspicions.

In the following analysis, there are no restrictions at all on side payments. Thus, firms can collude on outcomes in which only the efficient firm produces, but transfers part of its profits to the inefficient firm. This is clearly an extreme case that does not reflect reality, yet it allows

[14] One exception is Jehiel (1992).

[15] The effect of cross-ownership on collusion has recently been analyzed by Gilo, Spiegel & Moshe (2004) who find that cross-ownership, even if passive, usually facilitates collusion.

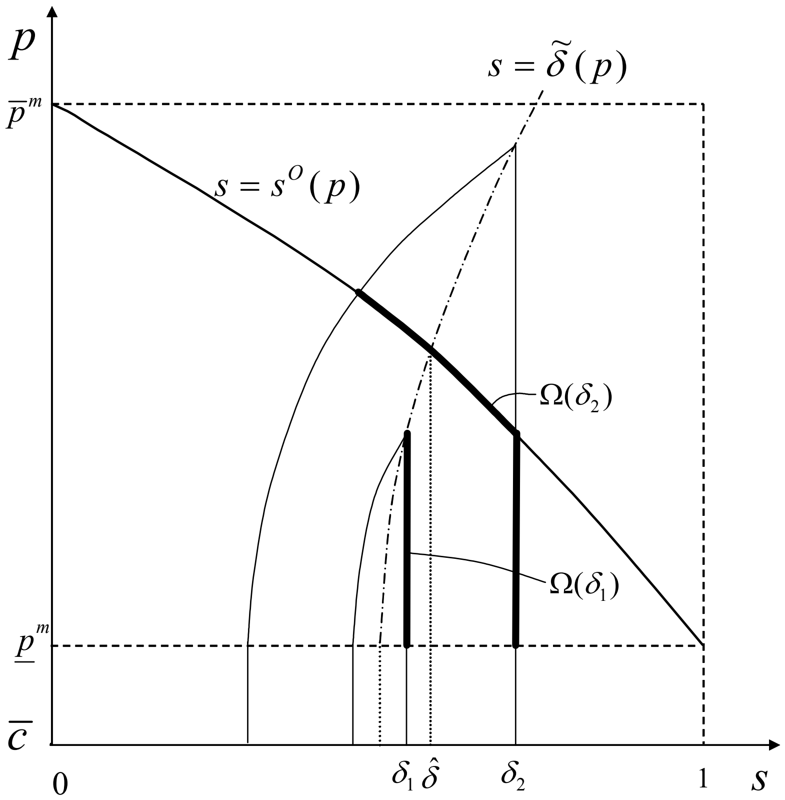

Figure 3: The Pareto frontier of sustainable allocations

us to identify the mechanisms by which cost asymmetry affects collusive sustainability when side payments are feasible. The main qualitative insight of our analysis, namely that cost asymmetry decreases the critical discount factor for collusion, still holds if the extent of side payments is limited.

5.1 Sustainable Collusive Outcomes and Efficiency

Without loss of generality, we restrict attention to collusive allocations, in which the low-cost firm carries out all the production in every period. Obviously, letting the high-cost firm produce a positive quantity in some or all periods would imply lower collusive profits. Deviation profits, however, could not be lowered this way. A collusive outcome is then defined by a couple (p, S), where p is the price, and S the profit sharing rule. In each period, the low-cost firm charges a price p, and the high-cost firm a price (slightly above) p.[16] The low-cost firm thus serves the whole demand, and then makes a side payment $(1-S)\underline{\pi}(p)$ to the high-cost firm.

We again restrict price to $p \in (\overline{c}, \overline{p}^m]$, and use the same punishments as before. Since these are maximal, there is no point in introducing side payments during punishment phases.

The low-cost firm could optimally deviate from the collusive outcome by refusing to make the side payment, and charging its monopoly price if $p \geq \underline{p}^m$ or remaining at p otherwise. The low-cost firm's no-deviation constraint is

$$\underline{\pi}\left(\min[p, \underline{p}^m]\right) \leq \frac{1}{1-\delta} S\underline{\pi}(p),$$

which is equivalent to

$$(1-\delta)\frac{\underline{\pi}\left(\min[p, \underline{p}^m]\right)}{\underline{\pi}(p)} \leq S. \qquad (\underline{D})$$

The high-cost firm could deviate from the collusive outcome by producing itself and undercutting the price. Hence, the high-cost firm' non-deviation condition is

$$\overline{\pi}(p) \leq \frac{1}{1-\delta}(1-S)\underline{\pi}(p).$$

[16]This supposes that the high-cost firm remains ready to produce in every period without ever making any sales along the collusive path.

In an alternative scenario, the high-cost firm is shut down and can only start production with some time lag. This would affect the deviation possibilities of the low-cost firm. The low-cost firm's optimal deviation price would always be $\underline{p}^m$, even if $p < \underline{p}^m$. If the inefficient firm could restart production fast enough however, the discount factor threshold for collusion would still lie below $\frac{1}{2}$.

This condition is equivalent to

$$S \leq 1 - (1-\delta)\frac{\overline{\pi}(p)}{\underline{\pi}(p)}. \qquad (\overline{D})$$

Adding up the two no-deviation constraints show that collusion on price p is sustainable if and only if :

$$\delta \geq \widetilde{\delta}^S(p) \equiv 1 - \frac{\underline{\pi}(p)}{\overline{\pi}(p) + \underline{\pi}\left(\min[p, \underline{p}^m]\right)}. \qquad (6)$$

Collusion with side payments, even efficient collusion, is sustainable for discount factors below $\frac{1}{2}$ now. In fact, the threshold function $\widetilde{\delta}^S(p)$ is strictly increasing in p: over the whole range of prices, collusion becomes more difficult as the price increases.

For $p \geq \underline{p}^m$, a price rise clearly increases the deviation incentives of both firms. Since the profit ratio $\frac{\overline{\pi}(p)}{\underline{\pi}(p)}$ is increasing in p, deviations become more profitable for the high-cost firm. For the low-cost firm, collusive profits are decreasing in p, while deviation profits are constant.

For $p < \underline{p}^m$, a price reduction alleviates the high-cost firm's non-deviation constraint without affecting the low-cost firm's deviation incentives. In fact, $\widetilde{\delta}^S(p) \to 0$ as $p \to \overline{c}$, so that some collusion is sustainable for any $\delta > 0$.[17]

An outcome is Pareto efficient if $p = \underline{p}^m$, so that the firms cannot jointly gain by either reallocating production or changing the price. Efficient collusion is sustainable for any discount factor above $\widetilde{\delta}^S(\underline{p}^m) = \frac{\overline{\pi}(\underline{p}^m)}{\overline{\pi}(\underline{p}^m)+\underline{\pi}(\underline{p}^m)} < \frac{1}{2}$. Figure 4 illustrates the sets of sustainable outcomes with side payments for $\widetilde{\delta}^S(\underline{p}^m)$ and for $\delta_1 > \widetilde{\delta}^S(\underline{p}^m)$.

Comparison with the Previous Analysis Under cost asymmetry, side payments facilitate collusion. Suppose the allocation (p, s) is sustainable without side payments. Then, the allocation $(p, S = s)$ is always sustainable with side payments. The high-cost firm has higher collusive profits than without side payments (since $s\underline{\pi}(p) > s\overline{\pi}(p)$), whereas its deviation profits are unchanged. The low-cost firm's deviation concern, on the other hand, remains unchanged. Intuitively, side payments permit greater "flexibility": production can be allocated efficiently without inhibiting collusion

[17] In the alternative scenario described in footnote 16, the discount factor threshold does not tend towards zero. However, collusion is still sustainable for discount factors below $\frac{1}{2}$ if the time lag is not too long. In particular, if the time lag is only one period, the discount factor threshold for efficient collusion, $\widetilde{\delta}^S(\underline{p}^m)$, remains unchanged.

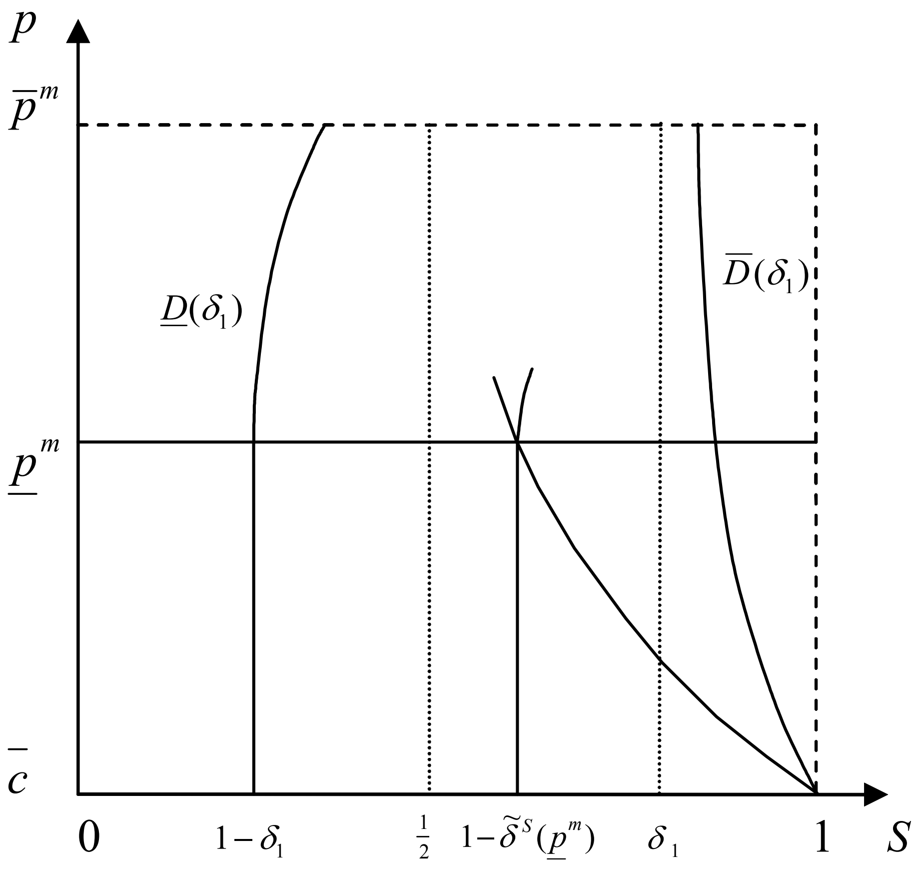

Figure 4: Sustainable Collusive Allocations with Side Payments

Collusion with side payments between asymmetric firms is also easier to sustain than collusion between symmetric firms.[18] This result occurs because cost asymmetry lowers the high-cost firm's incentive to deviate. With side payments, collusion with a more efficient competitor is more "valuable" than collusion with an identical firm. The low-cost firm's deviation incentives, on the other hand, are unaffected by the cost asymmetry.

5.2 The Pareto Frontier of Sustainable Allocations

With side payments, any allocation in which only the low-cost firm is an active producer and maximizes profits by setting its monopoly price is Pareto optimal for the firms. Interestingly, even if some unconstrained efficient outcomes are sustainable, the firms will not always want to select one of them. To see this, consider any given discount factor $\delta > \widetilde{\delta}^S(\underline{p}^m)$. Then, the preferred outcome of the high-cost firm is $(\underline{p}^m, 1-\delta)$. The low-cost firm's preferred allocation however is *not* such that the price is $\underline{p}^m$ and S as large as possible given the price. In fact, the low-cost firm prefers to move to a price strictly below $\underline{p}^m$. Such a move has a negative second-order effect on $\underline{\pi}(\underline{p}^m)$, but this effect is more than offset by a positive first-order effect on the low-cost firm's share S, since it alleviates the high-cost firm's no-deviation constraint.

It can be shown that the Pareto frontier of sustainable allocations has a similar shape as without side payments. For very small discount factors, only allocations on the high-cost firm's non-deviation constraint $\overline{D}(\delta)$ are undominated. For large enough discount factors, the Pareto frontier has two sections: the sustainable part of the unconstrained Pareto frontier ($p = \underline{p}^m$), and part of $\overline{D}(\delta)$ (namely, between $\underline{p}^m$ and some price strictly below $\underline{p}^m$).

6 Concluding Remarks

By using optimal punishments and allowing for side payments, this paper addresses two largely unexplored aspects in the existing literature on collusion between cost asymmetric firms. We have derived three main results: (i) Without side payments, some collusion is sustainable under cost asymmetry whenever collusion is sustainable under cost symmetry. (ii) Without side payments, efficient collusion is more difficult when costs are asymmetric. (iii) With side payments, cost asymmetries facilitate collusion. Hence, the general conclusion that cost asymmetry

[18] The feasibility of side payments is irrelevant under cost symmetry. In the absence of fixed costs, no advantage can be derived from allocating production to only one of the firms.

hinders the sustainability of collusion needs to be nuanced. In this sense, our paper provides a benchmark of "optimal collusion under cost asymmetry". Another key implication is that the feasibility of side payments between cartel members plays a particularly important role when firms have asymmetric cost structures.

This paper as well as some more recent research suggest that the use of grim trigger punishments is crucial for a number of the comparative statics results in the collusion literature. One example is the impact of horizontal differentiation on collusion in a simple Hotelling model when firms charge delivered prices. Relying on grim trigger strategies, Gupta & Venkatu (2002) find that in this set-up collusion is generally easiest under minimal differentiation. Our preliminary research shows that, in most situations, collusion sustained by optimal punishments is easiest when differentiation is maximal. The results obtained by relying on different punishment strategies are thus in stark contrast to each other.

Another interesting venue for future research may be to allow for more general cost structures so as to check the robustness of our results in that respect.

7 Appendix

Non-Stationary Collusion Here we show that allowing for non-constant prices and market sharing rules over time cannot facilitate collusion, i.e. permit collusion for discount factors below the threshold derived for stationary collusion. This result is derived in a slightly more general model than the one of the main text, allowing for n firms instead of just two. Without loss of generality, we suppose that these n firms are ordered by efficiency: $c_1 \leq c_2 \leq ... \leq c_n$. By revealed preferences, we then have that $p_1^m \leq p_2^m \leq ... \leq p_n^m$.

As shown in the main text, optimal punishments leave a deviator with continuation profits of zero in such a model. Deviation incentives are then minimized when firms share the market symmetrically, and charge a collusive price somewhere between the most inefficient firm's marginal cost c_n, and the most efficient firm's monopoly price p_1^m. The discount factor threshold for such a stationary collusive scheme is $\frac{n-1}{n}$.

A (possibly) non-stationary collusive sequence consists of a (common) price and a market sharing rule for each period from today to infinity: $\{p_t, (s_t^1, ..., s_t^n)\}_{t=0}^{\infty}$, with $s_t^i \in [0,1]$, and $\sum_i s_t^i = 1$.[19] With maximal punishments, firm i's non-deviation constraints are:

$$\pi_i(\min[p_t, p_i^m]) \leq \sum_{T=t}^{\infty} \delta^{T-t} s_T^i \pi_i(p_T) \text{ for } t \in [0, \infty), \tag{7}$$

By backward induction, if one of the non-deviation constraints is violated in any future period, collusion will already break down today.

Lemma 2 *For $\delta < \frac{n-1}{n}$, there does not exist any sustainable collusive sequence of prices and market sharing rules.*

Proof. First suppose that $p \equiv sup\{p_t\}_{t \geq 0}$ exists. Then for any $\varepsilon > 0$, there exists a period t_ε such that $\frac{\pi_i(p_{t_\varepsilon})}{\pi_i(p)} > (1-\varepsilon)$ for all i.

[19] Since we are interested in the collusive arrangement easiest to sustain, restricting attention to a common price for all firms is perfectly legitimate. To see this, consider a collusive arrangement such that firms charge different prices in some period, and denote the lowest price charged in that period by p_M. As their is no reason to ration consumers, the market price will be p_M, and only those firms that charge this price can have strictly positive market shares. Then a collusive arrangement with the same market shares but a common price equal to p_M is at least as easy to sustain than the one where firms charge different prices. The firms' collusive profits are exactly the same in the two arrangements, and their deviation profits are either unchanged or even reduced as profitably upward deviations can be ruled out.

Also, collusion could not be facilitated by allowing for $\sum_i s_t^i < 1$, since this would clearly reduce joint collusive profits.

We know that $p_1^m \leq p_2^m \leq ... \leq p_n^m$. There is hence some $I \in [0, n]$ such that $p > p_i^m$ for all $i \leq I$, but $p \leq p_i^m$ for all $i > I$. For sufficiently small ε, this implies that $p_{t_\varepsilon} > p_i^m$ for all $i \leq I$. The non-deviation constraints in period t_ε for the I most efficient firms ($i \leq I$) are then

$$\pi_i(p_i^m) \leq \sum_{T=t_\varepsilon}^{\infty} \delta^{T-t_\varepsilon} s_T^i \pi_i(p_T). \tag{8}$$

By the definition of the monopoly price, $\pi_i(p_T) \leq \pi_i(p_i^m)$ for all T. Therefore, (8) implies

$$\pi_i(p_i^m) \leq \sum_{T=t_\varepsilon}^{\infty} \delta^{T-t_\varepsilon} s_T^i \pi_i(p_i^m) \text{ for } i \leq I,$$

or equivalently

$$1 \leq \sum_{T=t_\varepsilon}^{\infty} \delta^{T-t_\varepsilon} s_T^i \text{ for } i \leq I. \tag{9}$$

For the less efficient firms $i \in [I+1, n]$, the non-deviation constraint in period t_ε are

$$\pi_i(p_{t_\varepsilon}) \leq \sum_{T=t_\varepsilon}^{\infty} \delta^{T-t_\varepsilon} s_T^i \pi_i(p_T).$$

Since $p_t \leq p$ for all t by definition, $p \leq p_i^m$ for $i > I$, and $\pi_i(\cdot)$ is increasing below p_i^m, this condition implies that

$$\pi_i(p_{t_\varepsilon}) \leq \sum_{T=t_\varepsilon}^{\infty} \delta^{T-t_\varepsilon} s_T^i \pi_i(p) \text{ for } i > I,$$

which in turn implies

$$(1-\varepsilon) < \sum_{T=t_\varepsilon}^{\infty} \delta^{T-t_\varepsilon} s_T^i \text{ for } i > I. \tag{10}$$

Adding up the necessary conditions in (9) and (10) for all n firms, using the fact that market shares add up to 1 in each period, yields the following necessary condition for collusion:

$$I + (n-I)(1-\varepsilon) \leq \frac{1}{1-\delta},$$

which rewrites as

$$\frac{n-(n-I)\varepsilon-1}{n-(n-I)\varepsilon} \leq \delta.^{20} \tag{11}$$

[20]This inequality is in fact strict unless $I = n$. In the latter case, we are done at this point, as we have derived the necessary condition $\frac{n-1}{n} \leq \delta$.

Now suppose, by contradiction, that the discount factor is $\widehat{\delta} < \frac{n-1}{n}$. Then there always exists some sufficiently small $\widehat{\varepsilon}$ such that $\frac{n-(n-I)\widehat{\varepsilon}-1}{n-(n-I)\widehat{\varepsilon}} > \widehat{\delta}$, so that the necessary condition (11) is violated. This completes the proof when the supremum of the price sequence exists.

We still need to consider the case when prices "explode" at some point approaching infinity. However, this analysis is trivial given the previous explanations. In fact, for large enough prices, charging its monopoly price is the optimal deviation for each of the n firms. Thus, everything is as if $I = n$ in the previous analysis, and the necessary condition (11) is simply $\frac{n-1}{n} \leq \delta$. ■

Proof of Proposition 1. As shown in the text preceding the proposition, the minimum discount factor for which some efficient collusion is sustainable must be such that both firms are indifferent between complying and deviating. The function $\widetilde{\delta}(p)$ was previously defined as the discount factor at which both non-deviation constraints are binding given the price, and for this discount factor only the market sharing rule $s = \widetilde{\delta}(p)$ can be sustained. Define $\widehat{p}$ then as the price at which the two non-deviation constraints are simultaneously binding at a Pareto-efficient market sharing rule $s = s^O(\widehat{p})$:

$$s^O(\widehat{p}) = \widetilde{\delta}(\widehat{p}).$$

Such a $\widehat{p}$ exists and is unique because $s^O(p)$ is strictly decreasing with $s^O(\underline{p}^m) = 1$ and $s^O(\overline{p}^m) = 0$, while $\widetilde{\delta}(p)$ is strictly increasing over the relevant range and $\widetilde{\delta}(\underline{p}^m) = \frac{1}{2}$. It also follows that $\widehat{p} \in \left(\underline{p}^m, \overline{p}^m\right)$, which implies that

$$\widehat{\delta} \equiv \widetilde{\delta}(\widehat{p}) > \frac{1}{2}.$$

The Pareto-efficient outcome $(\widehat{p}, \widehat{s})$, where $\widehat{s} \equiv s^O(\widehat{p}) = \widehat{\delta}$, is thus sustainable as long as

$$\delta \geq \widehat{\delta},$$

and $\widehat{\delta}$ is the lowest discount factor for which some efficient collusion is sustainable.

Now suppose $\delta > \widehat{\delta}$. Within the set of Pareto-efficient allocations, the high-cost firm's deviation incentives are increasing and the low-cost firm's deviation incentives decreasing with the collusive market sharing rule. The maximum s must hence be such that the high-cost firm is just indifferent between complying and deviating from the efficient allocation, whereas the minimum s is such that the low-cost firm is indifferent between complying and deviating.

Given any price and the discount factor, the low-cost firm's non-deviation constraint is binding at $\widetilde{s}(\delta, p)$. Let $\widetilde{p}^O(\delta)$ be the price as a function of the discount factor such that the low-cost firm's non-deviation constraint is binding at a Pareto-efficient market sharing rule:

$$s^O(\widetilde{p}^O(\delta)) = \widetilde{s}(\widetilde{p}^O(\delta), \delta).$$

The function $\widetilde{p}^O(\delta)$ is uniquely defined by this condition, and is strictly increasing in δ. The high-cost firm's non-deviation constraint is binding at a Pareto-efficient allocation market sharing rule for:

$$s^O(p) = \delta.$$

Thus, the range of market sharing rules for which Pareto-efficient collusion is sustainable becomes

$$\left[s^O(\widetilde{p}^O(\delta)), \delta\right].$$

This range is non-empty if and only if $\delta \geq \widehat{\delta}$. ■

References

[1] Abreu, D. (1986), "Extremal Equilibria of Oligopolistic Supergames", *Journal of Economic Theory*, Vol.39: 191-225.

[2] Abreu, D. (1988), "Towards a Theory of Discounted Repeated Games", *Econometrica*, Vol.56: 383-396.

[3] Bae, H. (1987), "A Price-Setting Supergame between Two Heterogeneous Firms", *European Economic Review*, Vol.31: 1159-1171.

[4] Bain, J. S. (1948), "Output quotas in imperfect cartels", *Quarterly Journal of Economics*, Vol.62: 617-622.

[5] Blume, A. (2003), "Bertrand without fudge", *Economics Letters*, Vol.78: 167-168.

[6] Compte, O., Jenny, F., and Rey, P. (2000), "Capacity Constraints, Mergers and Collusion", *European Economic Review*, Vol.46: 1-29.

[7] Davidson, C. and Denerecke, R. (1990), "Excess Capacity and Collusion", *International Economic Review*, Vol.31(3): 521-541.

[8] Gilo, D, Spiegel, Y. and Moshe, Y. (2004), "Partial Cross Ownership and Tacit Collusion", mimeo, Tel Aviv University.

[9] Gupta, B., and Venkatu, G. (2002), "Tacit Collusion in a Spatial Model with Delivered Pricing", *Journal of Economics*, Vol.76(1): 49-64.

[10] Harrington, J. E. Jr. (1991), "The Determination of Price and Output Quotas in a Heterogeneous Cartel", *International Economic Review*, Vol.32(4): 767-792.

[11] Harrington, J. E. Jr. and Skrzypacz, Andrzej (2005), "Collusion under Monitoring of Sales", mimeo.

[12] Jehiel, P. (1992), "Product differentiation and price collusion", *International Journal of Industrial Organization*, Vol.10: 633-641.

[13] Lambson, V. E. (1987), "Optimal Penal Codes in Price-setting Supergames with Capacity Constraints", *Review of Economic Studies*, Vol.54: 385-297.

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

[15] Rothschild, R. (1999), "Cartel Stability when Costs are Heterogeneous", *International Journal of Industrial Organization*, Vol.17: 717-734.

[16] Scherer, F. M. (1980), *Industrial Market Structure and Economic Performance*, Houghton Mifflin Company, Boston.

[17] Schmalensee, R. (1987), "Competitive Advantage and Collusive Equilibria", *International Journal of Industrial Organization*, 5: 351-368.

[18] Tirole, J. (1988), *The Theory of Industrial Organization*, MIT Press, Cambridge.

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