Costly Collusion in Differentiated Industries

Ki-Eun Rhee
Raphael Thomadsen*
Columbia University
August 22, 2004

JEL Codes:
D40: Market Structure and Pricing
L120: Monopoly: Monopolization Strategies
L130: Oligopoly and Other Imperfect Markets

Abstract:
This paper demonstrates that, under a set of weak assumptions, increased product differentiation will make it more difficult to sustain collusion when it is costly either to coordinate or to maintain collusion. These results contrast with the previous theoretical literature, which shows that, in the absence of these costs, greater differentiation can help foster collusion under some common models of product differentiation, but is consistent with the empirical literature that suggests that collusion tends to occur most among homogeneous firms. Further, we show that costly monitoring can be compatible with fully cooperative outcomes.

Keywords: collusion, product differentiation, oligopoly, monitoring

* The authors thank Kyle Bagwell, Bjorn Jorgenson, Mike Riordan and Margaret Slade for their helpful comments. Send correspondence to Raphael Thomadsen, 3022 Broadway, Uris Hall Room 613, New York, NY 10027. Phone: (212) 854-4708. Fax: (212) 316-9219. Email: rct27@columbia.edu.
1. **Introduction**

This paper demonstrates that, under relatively unrestrictive assumptions, collusion becomes more difficult to sustain as the level of product differentiation increases when there are costs associated with maintaining or coordinating collusion that are high enough. This lack of ambiguity stands in contrast with the results of the literature on collusion and product differentiation, which finds that different models of product differentiation predict different relationships between the degree of product differentiation and cartel stability, and that collusion can be easier to support when there are greater levels of product differentiation under many of the most popular models.

At the heart of this issue is a simple tradeoff: while the homogenous firms have the most to gain from collusion, they also have the most to gain from cheating on the collusive agreement. The result is that the exact demand specification determines which effect dominates, a fact that is matched by the ambiguity of the theoretical literature. Chang (1991) uses a Hotelling-style differentiation model to show that collusion in a Bertrand game is easier to sustain as firms are further apart. Ross (1992) shows that greater homogeneity can reduce cartel stability under other similar demand models. Using different models, Rothschild (1992) and Häckner (1994) find the opposite result, with the latter study examining vertically differentiated industries. Deneckere (1983) obtains mixed results when he analyzes the ease of collusion under both Bertrand and Cournot competition using a special case of Dixit’s (1979) model of product differentiation; he finds that collusion under Cournot competition is easier to sustain when the goods are more heterogeneous, but that the relationship between ease of collusion and product differentiation is not monotonic under Bertrand competition, with
increased product differentiation making collusion harder up to a certain level of
differentiation, but then making collusion easier for any further increase in
differentiation. Finally, Tyagi (1999) shows that the shape of the demand curve
determines whether collusion can be more easily sustained in homogeneous or
heterogeneous markets under quantity competition.

Despite this theoretical ambiguity, there is empirical support to the idea that
collusion occurs more frequently among homogenous firms. Hay and Kelley (1974) state
that most of the industries that they study where there was successful prosecution of anti-
trust cases involved firms that sold products that are subjectively homogeneous.
Similarly, Levenstein and Suslow (2002) conclude that price fixing occurs more
frequently in homogeneous product markets in their article summarizing the empirical
literature on cartels. These observations match the general findings among empirical
papers examining whether firms are colluding: those looking at homogenous firms have
generally found that there is tacit collusion,\(^1\) while those examining heterogeneous firms
have generally not been able to reject that firms are competing using static Nash
outcomes.\(^2\) Homogeneity can also increase the stability of the cartel; Jacquemin, Nambu
and Dewez (1981) find that cartels among firms selling homogenous products tend to
last longer than cartels among firms selling heterogeneous products. Finally, Asch and
Seneca (1976) provide indirect evidence that collusion is most common among
homogeneous firms. They find that collusion tends to occur among firms that are not

---

\(^1\) See Gollop and Roberts (1979) study of coffee roasting, Spiller and Favaro (1984) study of banking,

very profitable, and that collusion is more likely to occur among firms selling producer goods, both of which are consistent with what would be observed if collusion were most common among homogeneous firms.

In this context, several authors have tried to establish mechanisms that imply that greater homogeneity fosters collusion. Raith (2001) finds that collusion is easiest among firms that are homogeneous when local demand conditions are uncertain and there is imperfect monitoring. Other sources, starting with Chamberlin (1929), suggest that collusion is easiest among homogeneous firms because it is easier to coordinate what the collusive outcome should be. For example, Scherer and Ross (1990) claim that “cooperation to hold prices above the competitive level is less likely to be successful … the more heterogeneous, complex and changing the products supplied are.” This intuition is repeated in the 1992 US Department of Justice and Federal Trade Commission Horizontal Merger Guidelines, which claims that “Reaching terms of coordination may be facilitated by product or firm homogeneity.”

In this paper we propose a more general reason why collusion may be easiest among homogeneous firms – that there is a cost to maintaining or coordinating collusion. Examples of costs of sustaining cooperation include monitoring costs, communication costs or the continual potential of being caught and punished for illegally colluding, while examples of costs of coordinating collusion include costs of building supportive infrastructure for collusion and the cost of negotiating the collusive outcome. We show

---

3 Asch and Seneca assert that the producer goods are more homogeneous than consumer goods. Of course, there are other differences between these markets.

4 See section 2.11.
that, under a set of weak assumptions, the presence of these costs increases the difficulty of sustaining collusion more for firms that are more differentiated, and that collusion becomes monotonically harder to sustain as products grow more heterogeneous when these costs are large enough.

While the model we use in this paper is relatively generic, we limit our analysis by considering only the difficulty of sustaining full collusion between firms, defined as selecting an outcome lying on the Pareto optimal frontier (from the firms’ points of view). Also, we focus most of our analysis on punishments using Nash Reversion Grim trigger strategies, ala Friedman (1971). The use of these punishments is common in the literature,⁵ and we argue that there are features of anti-trust laws that imply that this is often the correct benchmark punishment to use as a response to a breakdown of collusion. However, Abreu (1986, 1988) has shown that punishments more severe than Nash reversion are sometimes feasible. Thus, we also provide a set of assumptions under which our results hold using optimal punishments.

We first discuss the cost of maintaining or coordinating collusion in Section 2. Then we show that product homogeneity helps foster collusion in the presence of maintenance costs in Section 3, and in the presence of coordinating costs in Section 4. In both cases, we investigate how costly collusion affects the relationship between product homogeneity and cartel sustainability and demonstrate that collusion becomes monotonically harder to sustain as the degree of product differentiation increases with sufficiently large costs of collusion. Section 5 discusses the broad implications of these

⁵ For example, see Friedman (1971), Deneckere (1982), Chang (1991), Ross (1992), Rothschild (1992), and Häckner (1994).
results and concludes.

2. Cartel Costs

The premise of this paper is that there are costs to coordinating and maintaining collusion. In this section, we discuss the sources from which these costs may arise along with empirical evidence of their existence.

We first examine the costs of coordinating collusion in Section 2.1, and find that due to a lack of focal outcomes, collusion has been documented to require many face-to-face meetings. Also, studies have found that firms need to invest in institutions specific to collusion in order for collusion to be successful. One of the most common institutional investments is that of forming a trade association that coordinates and monitors collusion.

We then examine the costs of maintaining collusion in Section 2.2. One cost of maintaining collusion is the cost of monitoring the other firm’s actions. The firms often pre-commit to spending the monitoring costs through the activities of the trade association, but in the appendix we present a model where firms will choose to monitor without pre-commitment even though firms always abide by the collusive agreement. Other costs of maintaining collusion are similar to those required for coordinating collusion – coordinating ongoing revisions to the collusive agreement and conducting ongoing discussions and renegotiations to maintain the cartel. Finally, the longer the cartel continues the greater the chances of prosecution for price-fixing.

2.1 Coordinating Costs

Two of the largest costs of coordinating collusion, according to numerous studies
of cartels, are costs of negotiating what the collusive outcome should be and the costs of creating the social and institutional structures needed to sustain collusion.

One reason that negotiations over collusive outcomes can be so costly is that there is no focal point for the firms to select as the equilibrium if there are asymmetries between the firms. Since the folk theorem suggests that there are an infinite number of potential collusive outcomes, different authors have chosen different approaches to selecting which collusive outcome they believe is most reasonable. Many authors, including d’Aspremont et al (1983), Donsimoni (1985), and Athey and Bagwell (2001) use joint-profit maximization. Jehiel (1992) takes another approach and assumes that firms collude by using the Nash bargaining solution. On yet another note, Friedman and Thisse (1993) consider only equilibria where the ratio of the firms’ profits under collusion is proportional to the ratio of those profits that would be realized under the non-cooperative equilibrium.\(^6\)

Knittel and Santiago (2003) show that the lack of focal collusive outcomes can inhibit tacit collusion. They find that collusion was more prevalent in credit card markets where there were laws dictating price ceilings than in markets without such regulations because these regulations acted as focal points for what the collusive outcome should be. Other empirical and case studies of collusion have found that selecting which collusive outcome to implement requires face-to-face meetings, which cost money, time, effort, and because these meetings are illegal, a probability of prosecution. In the Lysine and Citric acid market case, Connor (2001) discusses how difficult it was to reach an

\(^6\) Harrington (1991) shows that even these latter two approaches have complicating factors, compounding this issue.
agreement over what the collusive outcome should be. For example, at first the lysine producers agreed to set prices but not market shares. Only later were they able to set up a more efficient outcome by allocating market shares. Many other sources, including Hay and Kelly (1974), Stocking (1954), and Scherer and Ross (1990), also discuss how difficult negotiations over the collusive agreements can be. Indeed, Levenstein and Suslow (2002) find that bargaining problems are a major cause of cartel breakdowns. All of these factors can be seen by examining how difficult it is for OPEC, a cartel that can work in public without fear of reprisals, to reach agreements on how much oil each member nation should be allowed to produce.

Studies of collusion have also shown that collusion flourishes best where there is a social context to the interactions between the firms, which can set a more collusive tone for both initial and recurring negotiations. Levenstein and Suslow (2002) note that industries with cooperative cultures, such as the diamond industry or the rayon industry, had an easier time colluding than those without such a culture. Creating this social atmosphere requires an investment on the part of the firms. Connor (2001) notes that when a couple of ADM officers wanted to start a citric acid cartel they began doing this by meeting managers of other companies with what were described as “simply introductory, get acquainted sessions.” This practice is not new; Judge Elbert H. Gary, chairman of U.S. Steel’s board of directors, organized the famous Gary dinners from 1907 to 1911, where the leaders of the steel industry were invited to the Waldorf Astoria in New York to discuss steel industry practices.

In addition to building social infrastructure, firms often have to make internal and external institutional investments to support collusion. An example of internal adaptation
is that the firms participating in the lysine cartel had to figure out how to adhere to their quota without alerting either the middle managers or their sales force that such restrictions existed in order to try to avoid getting caught. Externally, collusion often has involved creating trade associations that coordinate the collusive schemes and publish and verify the output of cartel members in order to facilitate punishing firms that violate the agreements. These trade associations also provide a forum for discussion and renegotiation among the firms. While many of the costs of the trade associations are actually associated with maintaining collusion, there are also costs associated with initially establishing a strong trade association as well.

Levenstein and Suslow (2002) note that it takes cartels several attempts before they learn to develop the institutions that support a stable cartel. This is often because the collusive agreements require a high level of detail in order to keep incentives aligned. For example, gypsum producers needed to include a clause requiring their producers to use only rail delivery in order to support their price-fixing scheme.8 Another form of building institutions can involve building official regulatory agencies, as the railroads did with the ICC.9

These institutions are important in coordinating collusion even when parties can set up agreements publicly. Bagwell and Staiger (2002) discuss how the creation GATT fostered collusion in an international trade environment. Before instituting GATT, governments were hesitant to undertake cooperative behaviors due to mutual mistrust

---

7 This led to even further meetings among the cartel participants as they met to get good ideas about what to tell their sales force. See Connor (2001) for a detailed description.

8 See Scherer and Ross, page. 236.

9 See Scherer and Ross and Levenstein and Suslow.
after the tariff war in the 1920’s and 1930’s, even though governments were fully aware of potential gains from cooperative trade policy. The problem began to resolve following the establishment of GATT, which specified the detailed rules for multilateral negotiations and the method for enforcing the rules.

2.2 Costs of Maintaining Collusion

Costs of maintaining collusion include costs of monitoring the other players’ actions, costs of coordinating any collusive outcomes that vary over time, costs of renegotiation, and costs of the continual risk of prosecution. We discuss each of these in turn.

Monitoring can be a significant cost of maintaining collusion. Some readers may be concerned about whether monitoring can be a credible component of an equilibrium; in the appendix we discuss issues about the credibility of costly monitoring and present a simple game theoretic framework which has an equilibrium where both players always abide by the agreement and yet pay to monitor on the equilibrium path. However, firms that are members of cartels often pre-commit to incurring these monitoring costs by allocating these tasks to a third party. This can involve paying a trade association to collect and distribute very detailed information about each firm at a level of detail that allows the other firms to verify whether each firm adhered to the agreement. Firms also often consent to having this data verified either by the trade association or by an auditing firm hired by the trade association. Connor (2001) notes that this monitoring has often included opening up facilities and inventories to onsite inspections.

Maintaining collusion can often also require coordinating each period’s collusive
outcome and renegotiating over potential collusive outcomes. Athey and Bagwell (2001) discuss one such case, where firms have an incentive to use communication to truthfully reveal their true costs each period in order to determine how the firms should split the market in any particular period. If a firm insists that they should get more sales in one period then they reward their partner in collusion by allowing them to get more sales in a later period.\textsuperscript{10} The citric acid cartel implemented a similar scheme where firms that went over quota compensated those under quota by allowing the other firms to make up their sales in later periods.\textsuperscript{11} Another example of repeated coordination costs is the periodic updating of the pricing guides published by trade associations. This is often a non-trivial task given the potential complexity of the products.

Whether coordinated through a trade association or not, it is clear from case studies examining collusion that maintaining collusion requires constant communication and frequent meetings, as documented in Levenstein and Suslow (2002), Hay and Kelly (1974), Connor (2001), Stocking (1954) and Scherer and Ross (1990).\textsuperscript{12}

Finally, along with the direct costs described above, setting up cartels also involve indirect costs through the fact that setting up collusion is generally illegal, so whenever the members communicate, hold meetings or distribute documentation there is a risk that the members of the cartel will get prosecuted for these actions. The longer the collusion lasts, and the more phone calls and communication that goes on, the greater the chance of prosecution. This is both because there is a greater chance that a member of the cartel

\textsuperscript{10} While Athey and Bagwell (2001) do not include a cost of this communication, their result would hold if there were communication costs which were not too high.

\textsuperscript{11} See Connor (2001) and Levenstein and Suslow (2002) for more details on the citric acid cartel.

\textsuperscript{12} Athey and Bagwell (2001) provides a theoretical reason why this communication could be necessary.
will be a mole or a strong piece of evidence that allows for conviction gets into the wrong hands, and because the longer the cartel goes on the greater the probability that the government will take notice and begin investigating the activities of the member firms.

3. Collusion with Costs of Maintaining Collusion

In this section we demonstrate that collusion becomes more difficult to maintain as product differentiation increases when costs of maintaining collusion are high enough.

The model that we use is very basic. For notational simplicity, we assume that firms are symmetrical, although the model can be trivially modified to handle cases where this does not hold. Let \( d \) be an index that captures the level of differentiation. Denote the (assumed stationary expected value of) collusive profits of each firm as \( \pi^C(d) \), the static Nash equilibrium profits as \( \pi^N(d) \), and the profit that a firm derives from unilaterally deviating from a collusive agreement as \( \pi^D(d) \). Let \( \delta \) be the discount factor, which is the rate at which firms value a payment in the next period compared with an equal payoff in the current period. We assume that \( \pi^C(d) \), \( \pi^D(d) \), and \( \pi^N(d) \) are continuous and once-differentiable at all but a finite number of levels of differentiation.

Note that \( \pi^D(d) \geq \pi^C(d) \geq \pi^N(d) \).

We first consider collusion supported with grim trigger strategy punishments, where firms play Nash outcomes for eternity once collusion has been broken. Collusion can then be sustained whenever

\[
\frac{1}{1 - \delta} \pi^C(d) \geq \pi^D(d) + \frac{\delta}{1 - \delta} \pi^N(d),
\]

or equivalently,

\[
\frac{1}{1 - \delta} \pi^C(d) \geq \pi^D(d) + \frac{\delta}{1 - \delta} \pi^N(d) + 1.
\]

13 In the Lysine price-fixing scheme, a mole wore a wire to many meetings where the price-fixing was being held. See Connor (2001) for more information. Hay and Kelley (1974) note that some large illegal cartels had details of their meetings printed in large newspapers.
whenever
\[
\delta \geq \hat{\delta}(d) = \frac{\pi^D(d) - \pi^C(d)}{\pi^D(d) - \pi^N(d)}.
\] (1)

Note that \( \hat{\delta}(d) \) will be continuous and differentiable at all but a finite number of levels of differentiation given the assumptions on \( \pi^D(d) \), \( \pi^C(d) \), and \( \pi^N(d) \). We say that collusion is easier to support in market A than in market B when the discount factor required to sustain collusion is smaller in market A than in market B. That is, when
\[
\hat{\delta}(d_A) = \frac{\pi^D(d_A) - \pi^C(d_A)}{\pi^D(d_A) - \pi^N(d_A)} < \hat{\delta}(d_B) = \frac{\pi^D(d_B) - \pi^C(d_B)}{\pi^D(d_B) - \pi^N(d_B)}.
\]

Given this framework, we make a few assumptions:\(^{14}\)
\[
\frac{\partial (\pi^C - \pi^N)}{\partial d} < 0 \quad \text{(A1)}
\]
\[
\frac{\partial (\pi^D - \pi^C)}{\partial d} < 0 \quad \text{(A2)}
\]

When there is no limit to the level of product differentiation that can be obtained in the model\(^ {15}\) then it may be convenient to assume that
\[
\lim_{d \to \infty} (\pi^C - \pi^N) = 0. \quad \text{(A3)}
\]

Assumption (A1) states that the gain from colluding is highest among homogeneous products, and decreases as the level of product differentiation increases.

---

\(^{14}\) These assumptions only need to hold for the levels of differentiation where these functions are differentiable. For notational simplicity, we suppress the dependence of the profits on the level of differentiation hereafter.

\(^{15}\) For example, Hotelling models over a compact convex interval have a limit on the amount of product differentiation that is possible. Infinite Hotelling lines and Dixit (1979) do not have such a limit.
This may be expected because competition tempers prices less as products become more differentiated. Assumption (A2) states that the gain from deviating from a collusive agreement is also greatest among homogeneous products. The intuition for (A2) is that a firm will gain fewer consumers from undercutting a competitor’s price under greater levels of differentiation, so the one-time gain from cheating on a collusive agreement is smaller. Finally, assumption (A3) states that there is essentially no gain from colluding when the firms produce products that are differentiated enough, as would be expected if the product differentiation can become large enough that the firms effectively are competing in separate markets. This assumption is only used in Proposition 1.

Assumptions (A1)-(A3) do not constrain $\frac{\partial \delta}{\partial d}$ to be either positive or negative, and each of these assumptions hold for the most commonly used models of product differentiation, including all of those cited in this article.

Now suppose that there is a cost, $C$, of maintaining collusion. Examples of these costs are discussed in Section 2.2. The profit accrued in each period under collusion becomes $\pi^C = \pi^C - C$. The following propositions show that in the presence of this cost, it is easier to sustain collusion among firms selling homogeneous goods than among those selling more differentiated ones.

**Proposition 1:** Assume that there is no limit to the amount of product differentiation that is feasible and that (A3) holds. Then for any $C>0$ there exists some level of differentiation where collusion cannot be sustained for any greater level of differentiation under any level of discounting.

Proof: Collusion is unsustainable under any level of discounting whenever $C > \pi^C(d) - \pi^C$. 

\[ \pi^N(d) \]. By (A3), for every \( \varepsilon > 0 \) there exists a level of differentiation, \( \hat{d} \), such that \((\pi^C - \pi^N) < \varepsilon \) for all \( d > \hat{d} \). The proof is completed by substituting \( C \) for \( \varepsilon \). Q.E.D.

**Proposition 2:** Assume that (A1) holds. Then the range of product differentiation where collusion can be sustained at some level of discounting will collapse towards the most homogeneous markets as \( C \), the cost of maintaining collusion, increases.

Proof: Collusion is sustainable for some level of discounting only if \( C < \pi^C(d) - \pi^N(d) \). Assumption (A1) dictates that the right-hand side decreases as a function of differentiation, indicating that the range of differentiation supporting collusion at some level of discounting consists of an interval \( d \in [0, \hat{d}) \), where \( \hat{d} \) is lesser of the \( d \) that sets \( C = \pi^C(\hat{d}) - \pi^N(\hat{d}) \) and the maximum level of differentiation allowed in the model. As \( C \) increases, the right-hand side must also increase, which implies that \( \hat{d} \) decreases. Since there exists a \( \hat{C} \) such that \( \hat{C} = \pi^C(0) - \pi^N(0) \), and the function mapping \( C \rightarrow \hat{d} \) is (weakly) monotonic and continuous, this collapse is continuous and, eventually, complete. Q.E.D.

While propositions 1 and 2 show that the presence of costs to maintaining collusion limits the range over which collusion is feasible, they do not address whether differentiation always makes collusion more difficult within this range. To see that this is also the case, note that the critical discount factor becomes \( \hat{\delta}^* = \frac{\pi^D - \pi^C}{\pi^D - \pi^N} + \frac{C}{\pi^D - \pi^N} \) with these costs. To get the intuition of why this increases with differentiation when the costs of maintaining collusion are high enough, note that when these costs are small then
the first term will tend to determine the effect of differentiation on collusion. This effect is ambiguous and dependent on the functional forms in the model. However, when these costs are sufficiently large then the effect of the second term will dominate, and the fact that the denominator shrinks with product differentiation will cause the critical discount factor to increase. In fact, if costs are large enough then collusion is monotonically more difficult to sustain as the level of differentiation increases.

**Proposition 3**: Assume that (A1) and (A2) hold. Then there exists a level of costs for sustaining collusion, \( \hat{C} \), such that collusion is (weakly)\(^{16}\) monotonically more difficult to sustain as the level of differentiation increases for any \( C > \hat{C} \).

Proof: \( \frac{\partial \hat{\delta}^*}{\partial d} > 0 \) iff \( \frac{\partial (\ln \hat{\delta}^*)}{\partial d} > 0 \).

\[
\frac{\partial \ln(\hat{\delta}^*)}{\partial d} = \frac{\partial \ln(\pi^D - \pi^C + C)}{\partial d} - \frac{\partial \ln(\pi^D - \pi^N)}{\partial d}.
\]

\[
= \frac{\partial (\pi^D - \pi^C)}{\partial d} \frac{\partial (\pi^D - \pi^N)}{\partial d}.
\]

Assumptions (A1) and (A2) jointly imply that

\[
\frac{\partial (\pi^D - \pi^N)}{\partial d} = \frac{\partial (\pi^D - \pi^C)}{\partial d} + \frac{\partial (\pi^C - \pi^N)}{\partial d} < \frac{\partial (\pi^D - \pi^C)}{\partial d} < 0.
\]

Thus, equation (2) is positive whenever \( C \) is large enough; a sufficient condition for this to be true is that \( \pi^D - \pi^C + C \geq \pi^D - \pi^N \). This implies that \( \hat{C} < \pi^C_0 - \pi^N_0 \), where \( \pi_o \)

---

\(^{16}\) This relationship will be strictly monotonic over the range of differentiation where collusion is feasible (i.e., \( \hat{\delta}^*(d) < 1 \)), but for high levels of differentiation collusion is impossible for all \( \delta \).
denotes the profit levels that arise when \(d=0\), since \(\pi^D - \pi^C + (\pi^C_0 - \pi^N_0) \geq \pi^D - \pi^N\) for any level of differentiation by (A1), with equality only at \(d=0\). \textbf{Q.E.D.}

Figures 1-3 illustrate the implications of these three Propositions. Figures 1 and 2 demonstrate this relationship for Deneckere’s (1983) model under Bertrand and Cournot competition, respectively. Demand for each product is characterized as \(P_i = 1 - q_i + E q_i\), where \(E \in [-1,0]\). Note that the level of product differentiation increases as \(E\) is closer to zero. Figure 3 plots the critical discount factor against different levels of product differentiation for a Hotelling-style model, as detailed in Chang (1991). In this model, consumers are located uniformly on the interval \([0,1]\) and firms are located symmetrically at \(x\) and \((1-x)\). In all three cases, firms produce with zero marginal costs.

Proposition 1 states that when there is no limit on the degree of product differentiation that is feasible then for any \(C > 0\) there is some level of differentiation above which collusion cannot be supported under any level of discounting. While there is a limit on the degree of product differentiation that is possible in the Hotelling model, the Deneckere-style models do not have such a limit since the products compete in independent markets when \(E = 0\). Thus, in Figures 1 and 2 there is always a point where the critical discount factor rises to one, even with small maintenance costs, and at greater levels of differentiation there is no feasible discount factor that can support collusion. Proposition 2 states that the range of differentiation levels where collusion can be sustained collapses towards the most homogeneous markets as the maintenance costs increase. We see this in the three figures by noting that the levels of differentiation where the curves hit the \(\delta = 1\) threshold diminishes as the maintenance costs increase.
Finally, Proposition 3 states that collusion becomes monotonically more difficult to sustain for higher levels of differentiation once the costs to maintaining collusion are high enough. In Figures 1-3 we see that the relationship between the level of differentiation and the critical discount factor moves closer to being monotonic as costs to maintaining collusion increase, and at high enough costs it becomes completely monotonic.

While the focus of this paper is mostly on supporting collusion using the grim trigger strategies, it is worth noting how things would change if collusion were instead supported by optimal punishments. Propositions 1 and 2 will generally still be true to the extent that firms will not choose to collude if Nash competition is more profitable, although technically the folk theorem means that firms will be able to support collusive payoffs that are below the Nash equilibrium levels as long as these payoffs are higher than the minimax.

It turns out that Proposition 3 still holds under optimal punishments with a modification of assumption (A1). To examine this, suppose that any deviation from collusion yields the following most severe implementable stream of payoffs to the deviator: $\lambda_1(\delta,d), \lambda_2(\delta,d), \lambda_3(\delta,d), \ldots$\textsuperscript{17} We assume that it is possible to construct such a series such that $\lambda_t(\delta,d) \leq \pi^C(d) - C$ for every period $t$\textsuperscript{18} We believe that implementing the optimal punishment requires monitoring, although this assumption is not necessary as long as the other assumptions hold. In this case, the payoffs, $\lambda_t$, include a monitoring cost.

\textsuperscript{17} Abreu (1988) shows that such a sequence exists.

\textsuperscript{18} If $\lambda_t(\delta,d) > \pi^C(d) - C$, it would imply that the one meting out the punishment was getting below the collusive payoff in that period. While we cannot prove that all optimal punishments in any game can be crafted to conform this assumption, this property holds in the optimal punishments of Abreu (1986) and Häckner (1996).
Denote the average payoff from such a punishment as $\theta(\delta, d) \equiv \left( \sum_{t=1}^{\infty} \lambda_t(\delta, d) \delta^t \right) \frac{(1-\delta)}{\delta}$.

In this context the critical discount factor becomes \( \hat{\delta}^* = \frac{\pi^D - \pi^C + C}{\pi^D - \theta} \). Also, denote the minimax payoff as \( \pi^M(d) \). Note that that $\pi^D(d) \geq \pi^C(d) \geq \pi^N(d) \geq \theta(\delta, d) \geq \pi^M(d)$, and that \( \lim_{\delta \to 1} \theta(\delta, d) = \pi^M(d) \) by the folk theorem. We assume that $\theta(\delta, d)$ and $\pi^M(d)$ are continuous and once differentiable in $d$, and that

\[
\frac{\partial (\pi^C(d) - \theta(\delta, d))}{\partial d} < 0. \tag{A4}
\]

**Proposition 4:** Assume that (A2) and (A4) hold. Then there exists a level of costs for sustaining collusion, $\hat{C}$, such that collusion is (weakly) monotonically more difficult to sustain as the level of differentiation increases for any $C > \hat{C}$.

**Proof:** Following the first steps to the proofs of Proposition 3, equation (2) becomes

\[
\frac{d \ln(\hat{\delta}^*)}{dd} \left[ \frac{(\pi^D - \theta) - \hat{\delta} \frac{\partial \theta}{\partial \hat{\delta}}}{(\pi^D - \theta)} \right] = \left[ \frac{\partial (\pi^D - \pi^C)}{\partial d} \right] - \left[ \frac{\partial (\pi^D - \theta)}{\partial d} \right]. \tag{4}
\]

We first prove that $(\pi^D - \theta) - \hat{\delta} \frac{\partial \theta}{\partial \hat{\delta}}$, and therefore \( \frac{(\pi^D - \theta) - \hat{\delta} \frac{\partial \theta}{\partial \hat{\delta}}}{(\pi^D - \theta)} \), will always be positive. To see this, first note that

\[
\frac{\partial \theta}{\partial \hat{\delta}} = \sum_{t=1}^{\infty} \frac{\partial \lambda_t}{\partial \hat{\delta}}(1-\delta) + \sum_{t=1}^{\infty} \left[ (t-1)\delta^{t-1} - t\delta^{t-1} \right] \lambda_t \leq \sum_{t=1}^{\infty} \left[ (t-1)\delta^{t-2} - t\delta^{t-1} \right] \lambda_t,
\]

because any punishment scheme that is implementable at a low discount factor must be...
implementable at a higher discount factor, implying that the first term is negative.\footnote{While we use notation that suggests that $\theta$ is differentiable, the sustainable punishments $\lambda$, and therefore $\theta$, need not be continuous. However, the change in $\theta$ from any sized change in $\delta$ must be less than the change that comes directly from the deltas, as represented above. Thus, we ignore this lack of differentiability for notational simplicity. We also suppress the notation $\hat{\delta}$ for notational simplicity.}

Also, since $\pi^D(d) - \pi^C(d) + C = \frac{\delta}{1 - \delta}(\pi^C(d) - C - \theta(\delta, d))$ at the critical discount factor, $\pi^D - \theta = \frac{\pi^C(d) - C - \theta(\delta, d)}{1 - \delta}$. Given this:

$$
(p^D - \theta) - \frac{\delta}{1 - \delta} \frac{\partial \theta}{\partial \delta} \geq \frac{\pi^C(d) - C - \theta(\delta, d)}{1 - \delta} - \frac{\sum_{t=1}^{\infty} (t-1)\delta^{t-2} - t\delta^{t-1}}{1 - \delta} \lambda_i
$$

$$
= \frac{\pi^C(d) - C}{1 - \delta} - \sum_{t=1}^{\infty} \delta^{t-1} \lambda_i - \sum_{t=1}^{\infty} [(t-1)\delta^{t-1} - t\delta^t] \lambda_i
$$

$$
= \frac{\pi^C(d) - C}{1 - \delta} - \sum_{t=1}^{\infty} [t\delta^{t-1}(1 - \delta)] \lambda_i.
$$

(5)

To see that this is positive, note that we can define a series $\lambda_i^* \equiv \lambda_i - \pi^C(d) + C \leq 0$. Then equation (5) becomes

$$
(p^D - \theta) - \frac{\delta}{1 - \delta} \frac{\partial \theta}{\partial \delta} \geq \frac{\pi^C(d) - C}{1 - \delta} - \sum_{t=1}^{\infty} [t\delta^{t-1}(1 - \delta)] \lambda_i^* - \sum_{t=1}^{\infty} [t\delta^{t-1}(1 - \delta)] (\pi^C(d) - C)
$$

$$
= \frac{\pi^C(d) - C}{1 - \delta} - \sum_{t=1}^{\infty} [t\delta^{t-1}(1 - \delta)] \lambda_i^*
$$

$$
- \sum_{t=1}^{\infty} [(t-1)\delta^{t-1} - t\delta^t] (\pi^C(d) - C)
$$

$$
= \frac{\pi^C(d) - C}{1 - \delta} - \sum_{t=1}^{\infty} [t\delta^{t-1}(1 - \delta)] \lambda_i^* - \sum_{t=1}^{\infty} \delta^{t-1} (\pi^C(d) - C)
$$

$\phantom{1\frac{\delta}{1 - \delta}}$
\[ \sum_{t=1}^{\infty} t(\delta^{t-1})(1-\delta) = -2, \]  

(6)

where the second-to-last equality holds because each \( t\delta^t \) term can be matched with the \( (t-1)\delta^{t-1} \) term for the next \( t \), and \( (t-1) = 0 \) for \( t=1 \). The last term in (6) is clearly positive.

The right-hand side of equation (4) is also positive if \( C \) is large enough. To see this, remember that

\[
\frac{\partial (\pi^D - \theta)}{\partial d} = \frac{\partial (\pi^D - \pi^C)}{\partial d} + \frac{\partial (\pi^C - \theta)}{\partial d} < \frac{\partial (\pi^D - \pi^C)}{\partial d} < 0.
\]  

(7)

Then \( \frac{\partial (\pi^D - \pi^C)}{\partial d} \) is certainly positive for all \( C \) such that \( \pi^D - \pi^C + C \geq \pi^D - \theta \). The Folk Theorem (Fudenberg and Maskin (1986)) tells us that \( \theta(\delta, d) \to \pi^M(d) \) as \( \delta \to 1 \). Thus, \( \hat{C} < \pi^C_0 - \pi^M_0 \), where \( \pi_0 \) denotes the profit levels that arise when \( d=0 \), since \( \pi^D - \pi^C + (\pi^C_0 - \pi^M_0) \geq \pi^D - \theta \) for any level of differentiation, with equality possible only at \( d = 0 \) because \( \frac{\partial (\pi^D - \pi^M)}{\partial d} < 0 \) by (A4) and the folk theorem. Q.E.D.

Assumption (A4) holds for many models, although it does not hold as generally as assumptions (A1) – (A3). To see this, note that the minimax of some games is zero at all levels of differentiation, while the collusive payoffs will increase for at least some range of product differentiation. However, it is worth noting that this is not true for all games, and that the lack of (A4) holding may be thought of as an undesirable property of these models rather than as a weakness in the general result. Games where (A4) holds include
the following games where prices are constrained to be non-negative: \(^{20}\) Bertrand competition on a \([0,1]\) Hotelling line with linear travel costs, Cournot competition in the Deneckere version of the Dixit (1979) model, and many common empirical models of product differentiation, including those of geographic competition used by Davis (2003) and Thomadsen (2003).

It is also worth noting that there are reasons to believe that Nash reversion punishments should be the benchmark punishments even when they are not the most severe punishments. First, implementing optimal punishments may require many of the same costs as collusion requires. On the other hand, Nash-reversion punishments are self-enforcing, so the firms do not have to worry about monitoring the actions of the other firms and seeing whether they complied with the punishment. Also, in many games it may not be obvious to all of the participating firms what the optimal punishments are, which may increase the coordination costs of collusion, which in turn will make collusion less appealing.\(^{21}\) Finally, the optimal punishments are not renegotiation proof. Given that the punishment gives the deviator, and often the one giving the punishment, a payoff that is below the Nash level, the deviator may be inclined to renegotiate. If the firm that gives the punishment also gets below Nash payoffs then this firm will be especially tempted to accept playing the Nash equilibrium forever, which they know to be credible, rather than taking a larger loss to punish what is at that point an irreversible deviation. Further, the firm that is being punished may have the unilateral ability to get out of the

\(^{20}\) Pricing goods at negative prices, or pricing goods that obviously have positive marginal costs at zero, is almost surely going to draw an anti-trust investigation.

\(^{21}\) In Section 4 we show that this increases the degree to which collusion occurs among homogeneous firms.
severe punishment in a way that is not present in many games: it can notify the government about the price-fixing scheme and demand that the government investigate whether the actions of the other firm are currently consistent with competitive behavior. This is made credible by laws that many countries have stating that the first firm that confesses to a price-fixing scheme is immune from prosecution. This, in turn, gives the firm meting out the punishment further incentive to accept a request to substitute Nash reversion for the optimal punishment once the digression has occurred since with Nash reversion, neither firm has a strict incentive to report their past violations, but this is not the case with more severe punishments.

4. Collusion with Costs of Coordinating Collusion

In the last section, we demonstrated that collusion is easiest among homogeneous firms when there are costs associated with maintaining collusion. This section shows that costs of coordinating collusion have a similar effect.

The main role that costs of coordinating collusion have on discouraging collusion is that if collusion is sufficiently expensive to coordinate then the firms will choose to compete through Nash competition rather than to collude.\(^{22}\) Given this, it is natural to ask why there are costs to coordinating collusion that do not exist for Nash competition. This is answered in detail in Section 2.1, but it pays to briefly summarize the main points.

One likely cost to coordinating collusion is the cost of negotiating a collusive agreement. While non-cooperative game theory requires that any collusive arrangement

---

\(^{22}\) Other papers, including Athey and Bagwell (2001) use Nash equilibrium as the benchmark for a collusive outcome to beat.
be self-enforcing, it is silent on how the players know the strategies of the other players. (See Jacquemin and Slade (1989), Binmore (1992), Mas-Collel, Whinston and Green (1995).) While many oligopoly models have a unique Nash equilibrium, these same models often have no focal point for what the optimal collusive arrangement looks like. As a result, arranging collusion requires communication and discussion that Nash competition does not. Similarly, collusion requires some sort of verification of the competitors’ actions in order to keep the firms honest, while Nash competition is self-enforcing. Thus, the institutional and social investments that are necessary to make collusion credible are not necessary for Nash competition since Nash competitors will not be tempted to deviate from their strategy.

Suppose that firms initially compete in a way that delivers static Nash equilibrium profits in each period, but that the firms may move to collusive behavior after they both incur a cost, $C$, of coordinating collusion. Firms will pay this cost if (i) coordination is sustainable, i.e. $\delta \geq \frac{\pi^D - \pi^C}{\pi^D - \pi^N}$ as derived in the previous section, and (ii)

$$-C + \frac{\pi^C}{1-\delta} \geq \frac{\pi^N}{1-\delta} \iff \delta \geq 1 - \frac{\pi^C - \pi^N}{C}$$

Put together, the critical discount factor becomes $\hat{\delta}^* = \max(1 - \frac{\pi^C - \pi^N}{C}, \frac{\pi^D - \pi^C}{\pi^D - \pi^N})$.

**Proposition 5:** Assume that (A1) and (A2) hold. Then there exists a level of costs of coordinating collusion, $\hat{C}$, such that collusion becomes monotonically more difficult to sustain as the level of product differentiation increases for any $C > \hat{C}$.

**Proof:** A sufficient condition for this relationship to be monotonic is that
1 - \frac{\pi^C - \pi^N}{C} \geq \frac{\pi^D - \pi^C}{\pi^D - \pi^N} \text{ for all levels of product differentiation, because then}

\frac{\partial \hat{d}^*}{\partial d} = -\frac{1}{C} \frac{\partial (\pi^C - \pi^N)}{\partial d} > 0 \text{ since } \frac{\partial (\pi^C - \pi^N)}{\partial d} < 0 \text{ by (A1). } 1 - \frac{\pi^C - \pi^N}{C} \geq \frac{\pi^D - \pi^C}{\pi^D - \pi^N} \text{ iff } C \geq \pi^D - \pi^N, \text{ which will hold for all levels of differentiation if } C \geq \pi_o^D - \pi_o^N, \text{ since }

\frac{\partial (\pi^D - \pi^N)}{\partial d} < 0 \text{ (see equation (3)). Thus, } \hat{C} \leq \pi_o^D - \pi_o^N. \text{ Q.E.D.}

It is worth noting that the results in Proposition 5 are not dependent upon whether deviations from collusion are punished by Nash reversion or by optimal punishments because 

1 - \frac{\pi^C - \pi^N}{C} > \frac{\pi^D - \pi^C}{\pi^D - \pi^N} \geq \frac{\pi^D - \pi^C}{\pi^D - \pi} \text{ for high enough } C, \text{ so the sufficient condition prescribed in the proposition is also sufficient under optimal punishments.}

5. Conclusions

Despite the fact that the literature has shown that the functional forms of demand dictate whether collusion is easiest to sustain for homogeneous or heterogeneous products, anti-trust policy, supported by empirical studies, reflects the belief that collusion generally occurs among firms selling homogeneous goods. This paper bridges these two streams of literature by noting that the theory predicts that collusion will be easiest to support among homogeneous firms, regardless of the shape of demand, when there are costs associated with coordinating or maintaining collusion that are large enough. Not coincidently, these costs of collusion have been the stated justification by the U.S. DOJ and FTC of why collusion is most likely to occur among homogeneous
We assumed in this paper that the costs of maintaining or coordinating collusion are constant across the levels of differentiation. However, there may be cases where this does not hold. For example, if we think of maintenance costs as monitoring costs, then it may be reasonable to assume that the costs increase in the level of differentiation because it is more difficult to observe prices or detect secret defections at outlets that are far away. Thus, one could place the model of Raith (2001) in the context of our model, where the cost of maintaining collusion increases with the level of product differentiation instead of remaining constant. Coordination costs may also increase in the level of product differentiation if we think of coordination costs as bargaining costs because greater asymmetry between firms may increase the costs of negotiations. For example, when products are homogeneous, the firms can agree on a single price or level of output for the product, but when the products are more heterogeneous it is possible that the firms would have to set a schedule of prices and output levels (See, for example, Scherer and Ross (1990), Chamberlin (1929), Jacquemin and Slade (1989), U.S. DOJ and FTC (1992).)

While endogenizing these costs is beyond the scope of this paper, it is worth noting that when these costs increase in the level of differentiation our result that sustaining collusion is more difficult for more heterogeneous products is only reinforced.

The results of this paper imply that the FTC is justified to focus on collusion among relatively homogeneous firms because it is likely that such costs do exist. Also, they suggest that the effects of an amnesty program for price-fixing cases is likely to have a larger effect on stopping collusive behavior among heterogeneous firms than among homogeneous ones. This is because an amnesty program increases the probability that
the participants in a price-fixing scheme are caught and punished, increasing the cost of coordinating collusion,\(^{23}\) which we have shown obstructs collusion most among heterogeneous firms.

Finally, this paper also has implications for geographic competition. While economists generally believe that prices will generally be lowest in markets where firms are located in close proximity, this may not be the case if only firms in close proximity collude. If this is the case then there likely exists an optimal distance of differentiation – that where firms stop colluding and start competing – where prices will be lowest. Finding this level of differentiation, and verifying that firms in closer proximities are more likely to collude, are topics for empirical work. Intuitively, this would involve comparing prices for outlets that are located at different distances from their nearest competing outlets. A finding that prices are higher among firms with nearby competitors than those further away from their nearest competitors would be evidence that homogeneous firms are, on average, colluding more than heterogeneous firms. Also, one can establish the level of differentiation that is likely to yield the lowest prices to the extent that one can plot the evolution of average prices against different levels of differentiation.

\(^{23}\text{An amnesty program may perhaps increase the costs of maintaining collusion, too, to the extent that the longer the parties participate in a collusive scheme the more likely they are to be caught and punished, and perhaps the more severe the punishment will be.}\)
Appendix

Credible Monitoring with Complete Cooperation

While we note in Section 2.2 that firms often pre-commit to monitor the actions of other firms, some firms may wish to collude without building a trade association. It is natural to ask whether it is credible for the firms to incur the costs of monitoring in this context. The intuition behind doubt about the credibility of monitoring is that it would seem that the firms would have an incentive to shirk on any monitoring they were instructed to carry out if firms always abide by the collusive agreement in equilibrium. If this were the case then the only equilibrium is a mixed-strategy equilibrium where firms follow the collusive agreement with some probability and also monitor with some probability. Several papers in the principle-agent literature on moral hazard have found this result.\textsuperscript{24}

Various approaches have been used to try to solve this conflict. Kandori and Matsushima (1998) examines the problem in a differentiated product market where each firm receives a different private signal, and suggests that the problem of secret price cutting can be resolved by communication, showing that under certain circumstances players will communicate truthful information that can lead to optimal outcomes.

Bernheim (2004) takes a different approach and seeks to explain why we observe price undercutting in collusive agreements which are not followed by price wars. In his differentiated product model, firms agree never to undercut competitors below a certain

\textsuperscript{24} See Baiman and Demski (1980), Dye (1986), Holmström (1979), and Kanodia (1985), for example. The probability of monitoring can either be due to a mixed monitoring strategy, or from a pure strategy based on whether observable results fell below a certain threshold.
minimum price. Selling a product to a consumer requires a sales call, but firms cannot observe whether their competitor visited a customer unless they pay some cost. The consumer will purchase from the lowest priced firm that visited them. The equilibrium outcome is that firms never cheat on the absolute price floor, but choose prices and monitoring following a mixed-strategy. Thus, a partial collusion is obtained, although the agreement itself is never violated.

However, it is possible to construct a simple game framework that obtains pure strategy equilibria where firms play pure strategies of cooperation and choose to pay monitoring costs. Consider a repeated game where each stage of the supergame has the following substages: (1) Firms choose their action, which is not observed by the other firm. (2) A random payoff is obtained. (3) Firms each decide whether to pay to monitor the action of the other firm. Each firm can observe whether the other chooses to pay this cost. We say that monitoring is successful if both firms monitor and find that the other firm played the cooperative action.

One can verify that both firms playing the following strategy forms a collusive equilibrium as long as the discount factor is high enough: at substage (1), play the cooperative action as long as cooperation and monitoring has been successful in every previous period, play the Nash outcome otherwise. At substage (3), the firm pays to monitor if and only if it played the cooperative action in substage (1) of the same stage of the supergame and cooperation and monitoring has been successful in every previous period.

25 We have in mind that either this payoff is not observable to the other firm, or if the other firm can see the payoffs then this is not sufficient information to verify whether the collusive agreement was followed or not. While these restrictions on observability seem natural for markets where firms monitor, they are not actually necessary for the equilibrium laid out here.
The monitoring need not occur with probability 1. For example, if there is a public signal then the equilibrium above can be modified to say that in substage (3) the firms only monitor when certain signals are observed. The players would continue to cooperate until either monitoring reveals that the other player did not cooperate or the other player refuses to pay for the monitoring when it is prescribed. This modified set of strategies will be an equilibrium as long as the probability of monitoring is high enough.

If some of the information about the firms’ profits are known then there exist equilibria in the spirit of the lower-tailed monitoring papers such as Dye (1986) that have full cooperation but involve monitoring only in states where either of the firms’ profits are low enough. For example, if it is possible for both firms to observe prices but not quantities in a quantity-setting game with random per period demand, then the strategy for substage (3) could be modified to be to pay to monitor only if the firm played the cooperative action in substage (1) and either player’s prices are below some cutoff.

There is also an analogy between this equilibrium and Green and Porter (1984). In equilibrium, there is no cheating from the collusive agreements in both models, but in both cases firms implement a response to low payoffs that need to be endured by the firms to keep cooperation credible. In this case, the monitoring substitutes for the price war that the firms need to endure in Green and Porter to keep monitoring credible.
Bibliography:


Bernheim, B.D., 2004, Cheating within Imperfect Cartels, Mimeo, Stanford University.


Davis, P., 2003, The Effect of Local Competition on Retail Prices: The US Motion Picture Exhibition Market, Mimeo, London School of Economics.


Dixit, A., 1979, A Model of Duopoly Suggesting a Theory of Entry Barriers, Bell Journal
of Economics 10, 20-32.


Willig, Elsevier, 415-473.


Perloff, J., V. Suslow and P. Seguin, 1996, Higher Prices from Entry: Pricing of Brand-Name Drugs, Mimeo, University of California, Berkeley.

Raith, M., 2001, Competition, Risk and Managerial Incentives, Mimeo, University of Chicago.


Spiller, P. T. and E. Favaro, 1984, The Effects of Entry Regulation on Oligopolistic Interaction: The Uruguayan Banking Sector, RAND Journal of Economics 15, 244-54.


Figure 1. Deneckere's Model with Bertrand Competition

Figure 2. Deneckere's Model with Cournot Competition
Figure 3. Hotelling Model a la Chang