Efficient Collusion with Private Monitoring

Grigory Kosenok[*,†]

November 8, 2005

Abstract

This paper considers an infinitely repeated Cournot duopoly with Imperfect Monitoring. Each firm does not observe the production level of the other firm, but instead observes only a noisy private signal (the price of the product). We show that if the support of the signal is not too large, there is an equilibrium in which both firms produce the cartel level of output. This equilibrium is a result of a slight modification of the grim trigger strategy, showing that the concept of a grim trigger strategy works in a more general context than has previously been envisioned. A Folk Theorem is also established for our game.

[*]This paper is based on my PhD dissertation at the University of Wisconsin-Madison. I am very grateful to my advisors Larry Samuelson, Bart Lipman and William Sandholm for their guidance and encouragement. Also I wish to thank participants of seminars at the University of Wisconsin-Madison, New Economic School of Moscow and Central European University of Budapest for helpful suggestions and comments. It is a pleasure to acknowledge useful conversations with Stanislav Anatolyev, Oleg Balashov, Yeon-Koo Che, Philip Haile, Peter Norman, Menesh Patel and Sergei Severinov.

[†]New Economic School, Nakhimovsky Prospekt 47, Suite 1721 (11), 117418 Moscow, Russian Federation, gkosenok@nes.ru.

1 Introduction

The question of how cooperative motives can arise in a competitive industry has been debated for a long time. Starting with Stigler's work (1964), the dynamic motivation for forming a cartel has been extensively considered. The idea stems from restricting industry production to the monopoly level in the long run. This level of production gives an extra profit to each member of the cartel compared to the short run Nash equilibrium - e.g. the Cournot outcome. The monopoly output is supported by the threat that a deviation of any member from the cooperative cartel output is punished by moving the whole industry to the static Nash Equilibrium outcome. This threat is sequentially rational, and generates a self-enforcing cartel in the industry.

The kind of threats that can be credibly made is sensitive to the information structure available to the firms. We can identify three cases.

First, the case of *Perfect Monitoring* is one in which there is perfect observability of every firm's output. For this case, it has been shown that for a sufficiently high discount rate, the firms can cooperate forever (e.g. Friedman (1971) and Abreu (1982)).

In the case of *Public Monitoring* firms know only their own output and a publicly observed signal related to the joint actions of all firms - typically, the price of the product in the market. For this case, the following industry behavior has been shown to sustain the cartel output: firms cooperate as long as the price is higher than some threshold price level, and the industry falls to the non-cooperative phase for several periods in response to a low price realization. After a low-price phase, firms curtail production and the industry returns to one characterized by cooperation. Hence in equilibrium, production fluctuates according to the dynamics of the signal. More details can be found in Green and Porter (1984) and Abreu, Pearce and Stacchetti (1986).[1]

This paper examines a model with the third type of information structure, *Private Monitoring.* Private monitoring occurs when each firm observes its own private signal of the actions of the other firms. In the industry modeled, there are two firms interacting for infinitely many periods. The product is homogeneous, and each firm observes its own pro-

[1] Empirical support for this type of price behavior can be found in Brander and Zhang (1993), Ellison (1994) and Levenstein (1997).

duction and a realized price, the latter of which is imperfectly related to the actions of the other firm. This price is, in fact, a signal imperfectly related to the actions of the other firm. We assume that the realization prices may differ between the firms although the product is homogenous. Specifically, each firm observes a separate realized price that is a function of the aggregate output and a firm specific random shock. The shocks are assumed to be independent and thus a firm's realized price is private information. This is a novel assumption in the oligopoly literature and not only captures the empirical evidence that prices do differ for the same good and but is also closely related to the specification used in models of differentiated products.[2] In the model, a firm uses its realized price and its knowledge of the structure of the demand curve to obtain an estimate of the other firm's production level. We refer to this estimate, which is used heavily in the paper, as a firm's *signal*.

Because of their more realistic nature, private monitoring games have recently become quite popular, with Prisoner's Dilemmas being the most commonly studied stage game.[3] In private monitoring games there is asymmetric information, as each player observes her opponents' actions imperfectly. So far it has been shown that for infinitely repeated games, similar to the game considered in the present paper, the Folk Theorem generally fails.[4] This happens because unlike in a public or perfect monitoring environment, firms do not have any joint knowledge of the behavior of each other; therefore, they cannot coordinate a switch to a punishment phase.[5] The complications stemming from the private monitoring information structure have made it hard to examine under what conditions, if any, firms would be able to coordinate their actions. The central goal of this paper is to address this question.

The game theory literature offers some insight into this problem. Matshushima (1991) and Compte (2002) examine a general class of games with a simple structure (a finite number of possible actions and signal realizations). They show that, in general, no degree of cooperation can be maintained. Malaith and Morris (2002) consider games of "almost public monitoring"

[2] Tedeschi (1994) uses similar approach in the public monitoring contest.

[3] See, for example, Bhaskar (1999), Ely and Valimaki (2002), Mailath and Morris (2002), Piccione (2002) and Sekiguchi (1998).

[4] Another name for this phenomenon is the "anti-folk theorem" result. See Compte (2002) and Matsushima (1991).

[5] Although, if communication is allowed, which is excluded in our model, then is possible to have positive result. See Aoyagi (1997), Compte (1998) and Kandori and Matsushima (1998).

in which each player observes the same signal with probability close to one. They find an almost efficient outcome in those games. An approach similar to theirs applied to oligopoly games *might* yield an outcome *close to* full cooperation. Therefore, the literature suggests that in oligopoly models, there might be at most partial cooperation when the Private Information structure is close to the Public one.

The information structure is a key factor preventing the lack of perfect cooperation under private monitoring. However, it is not the only obstacle, as evidenced by the fact that perfect cooperation is not attainable under public monitoring. Another factor that prevents full cooperation is the range of signal realizations. In fact, previous research has only considered a constant range for signal realizations - namely, the support of the signal realizations does not depend on firms' actions.

Consider an equilibrium strategy that forces all players to play the mono-poly outcome always. This means that each firm should produce the cooperative outcome for every possible realization of its signal. When a particular firm takes an action, each of its competitors observes a noisy signal of that firm's action. By assumption, the supports of those signals are independent of the firm's action. Since in the proposed equilibrium the opponents always produce at the collusive level, the best response for the firm is to produce an amount higher than the collusive level. In other words, it is not rational for the firm to sustain cooperation all the time, which means that there is no full cooperation in these games. Instead, these games display *almost* full cooperation. To attain it, the support of the signal received by a firm is usually divided into two regions: one where the firm sustains cooperation, and another where the firm starts punishment. When the firms cooperate, a firm is more likely than not to receive signals that induce it to sustain cooperation, but occasionally it gets signals that prompt it to punish, despite the fact that no cheating has occurred.

In the contrast to previous work the key property of the model considered in this paper is that the support of the signal, which is assumed to be bounded, changes with the actions of the firms. In this case, if a firm observes a signal outside the range that corresponds to cooperation, it is certain that a deviation has occurred, but this fact is not common knowledge. Since this information is the private knowledge of the detecting firm, a strategy that calls for an immediate coordinated punishment cannot be implemented. Although a shifting support for the signal does not remove the informational asymmetry, it creates conditions which allow full cooperation and for the Folk Theorem to hold.

Under a shifting support for the signal we show that for signals without too much noise - signals with sufficiently narrow support - all firms produce the monopoly output level all the time in equilibrium. The equilibrium shown is the result of firms following a modified "grim trigger" strategy, which among other things specifies that a firm detecting a deviation in the current period produces a high output level in the next period. This action transmits to the industry the information that a deviation has been detected, thereby triggering all industry to punish the deviator in all future periods. Other components of the strategy make this initiation of punishment optimal.

It will be shown that this strategy allows the firms to realize the full monopoly profit for the same range of discount factors as in the full information case, as long as the range of varying of realized prices is not too wide. This finding thus demonstrates that a key result from an environment with perfect monitoring will hold in an environment with private monitoring, and that simple variants of the Grim Trigger Strategy are applicable in more general contexts. Additionally, we obtain the following Folk Theorem result for my model: for any individually rational feasible outcome there exists an equilibrium strategy that supports this outcome, as long as the discount factor is sufficiently high and the range of prices is sufficiently concentrated.

The structure of the paper is as follows. In Section 2 we give a formal description of the model. Sections 3, 4 and 5 provide results for the perfect, public and private monitoring cases, respectively. Section 6 considers the sensitivity of our results with respect to the shape of the signal's density, and Section 7 explores the Folk Theorem for our game. Section 8 presents some concluding observations.

2 The Model

Consider the following Cournot duopoly model with infinitely many periods. There are two identical firms that produce the same product. In each time period t, firm 1 produces quantity q_t^1 and firm 2 quantity q_t^2. Since the firms are identical, the superscript i is used for the firm of interest and j for her opponent. In every period each firm has zero cost of production. In period t, firm i's realized price, p_t^i, is imperfectly related to the current total production of two firms: $q_t^i + q_t^j$. In the model, we assume the conventional linear dependence between price and total quantity, but unlike most previous models price is assumed as random. Namely, the price is affected by a shock which is additive and not related to the total quantity produced. The mathematical form of this relation is[6]

$$p_t^i = 1 - (q_t^i + q_t^j) - \varepsilon_t^j \tag{1}$$

where ε_t^j is the price shock. The reason for this notation will be clear shortly.

As mentioned before, the scope of the paper is to consider the two situations of imperfect monitoring: public and private. In both cases the player does not observe the rival's action, q_t^j, but rather observes her own realization price, p_t^i, a price which may differ from that of the other firm. When the realized prices do not differ, the public monitoring situation takes place. On the other hand, the realized prices might be different. Such the situation will arise, for example, if firms sell differentiated products where each firm faces the same price disturbed by an idiosyncratic "brand shock". This moves us to the private monitoring environment which is the primary interest of this paper.

Let us describe the information available to a firm. At the beginning of the period t firm i decides the amount to produce, q_t^i. The realized price of the product p_t^i becomes known at the end of period t after the quantities q_t^i and q_t^j are chosen. Firm i never observes the quantity q_t^j produced by the other firm. Instead, because the firm knows the market demand is given by (1), the "estimate" of the other firm's production $\tilde{q}_t^j = p_t^i + q_t^i - 1$ can be generated at the end of period t. We refer to the *signal* as either p_t^i or $\tilde{q}_t^j$, where the latter entity receives such a denotation since it can be derived directly from the price realization, p_t^i. The next formula reflects the relation between the actual production of the rival and the estimate:

$$\tilde{q}_t^j = q_t^j + \varepsilon_t^j. \tag{2}$$

From all of the above it follows that at the end of period t firm i gets actual profit $p_t^i q_t^i$. By formulae (1) and (2) the same profit expressed in terms of quantity produced and the

[6]This function without loss of generality presents any linear demand after proper rescaling.

estimate received in the end of the period is

$$P(q_t^i, \tilde{q}_t^j) = (1 - q_t^i - \tilde{q}_t^j)q_t^i.$$

Next, it is assumed that each firm has the intertemporal discount factor δ, so firm i's utility, u^i, will be the discounted sum of its profits:

$$u^i = (1-\delta)\sum_{t=0}^{+\infty} \delta^t P(q_t^i, \tilde{q}_t^j).$$

In the paper we assume that the shocks $(\varepsilon_t^1, \varepsilon_t^2)$ are independent and identically distributed (i.i.d.) over time with zero expected value.[7] Furthermore, each shock ε_t^j has a limited support. Notice that the support of player i's estimate $\tilde{q}_t^j$ of player j's action is always centered around q_t^j due to formula (2), so that this support moves together with q_t^j. This specification contrasts with the one usually made in most of the imperfect monitoring literature, where each player's action has no effect on the support of the signal observed by her opponent. For demonstration of the main results of the paper, we assume that ε_t^j is uniformly distributed on $[-r, r]$, an assumption which makes our problem more tractable mathematically. A more general form for the support of the noise is considered in Section 6.

As we mentioned above, each firm makes its decision before its profit is realized. The decision is based upon beliefs about the action of the other player. Hence for the further analysis it is necessary to introduce the notion of a firm's expected profit - the profit which is expected by a firm given that it knows the production of the other firm:

$$\pi_E(q_t^i, q_t^j) \equiv E\{P(q^i, \tilde{q}^j)|q^j\} = [1 - q^i - q^j]q^i. \tag{3}$$

where $E\{.|.\}$ stands for conditional expectation.

Note the following two facts:

- The expected profit has the same form as in the perfect monitoring case (there is no shock)[8];

- Our model incorporates the Perfect Information situation as a special case when $r = 0$.

Regarding the joint distribution of ε_t^1 and ε_t^2, we consider three cases. The first one is the case of perfect monitoring. The second is the case of public monitoring, in which shocks are

[7] The zero mean makes $\tilde{q}_t^j$ an unbiased estimator of q_t^j

[8] This is due to the linear dependence between player i's profit and the signal.

the same $\varepsilon_t^1 = \varepsilon_t^2$, i.e. the two firms face the same price, but this price is still random and does not give exact information about the quantity produced by the other firm. The third case is the private monitoring case, which arises when ε_t^1 and ε_t^2 are independent.[9]

3 Perfect monitoring and the "grim trigger" strategy

The model becomes a perfect monitoring model when $r = 0$. In this situation the so called "grim trigger" strategy profile helps to support the cooperative outcome.

Briefly recall the concept of this kind of strategy. First of all, the efficient symmetric outcome for this game takes place when each player chooses to produce $Q_P = 1/4$, the Pareto efficient quantity for both players. This gives each player the maximum possible symmetric profit $\pi_P = 1/8$. We refer to this strategy as the restricted, optimal, cooperative or cartel output. In a one period setting, this game has a unique Nash equilibrium, where each player plays the Cournot outcome quantity $Q_C = 1/3$, and obtains a profit of $\pi_C = 1/9$. We refer to this strategy as noncooperative or Cournot output. The well known grim strategy, which allows for certain values of δ the efficient outcome in an infinitely repeated game with perfect monitoring, is stated below.

The grim trigger strategy:

Produce the efficient quantity Q_P initially and as long as every player produced Q_P in the last period; otherwise produce the disagreement quantity Q_C.

The above strategy is sequentially optimal only for certain values of δ. The next well known result is quite simple and specifies this range. We provide it here with the proof because the range of δ turns out to be universal for all results of the paper.

Result 1: *The grim trigger strategy is a subgame perfect equilibrium strategy for any* $\delta \in [\underline{\delta}, 1)$, *where* $\underline{\delta} = 9/17$.

PROOF: To find the condition on the values of δ required to support equilibrium we need to check that any one-period deviation from cooperation does not offset the long run losses.

[9] Of course, there is an intermediate case of partial correlation between the ε's, but the analysis of such a case is beyond the scope of this paper.

The highest possible short run gain comes from playing 3/8, which is the static game best response to the opponent's play of $Q_P = 1/4$. This deviation yields a profit of 9/64, which is higher than the one period equilibrium payoff 1/8. In every subsequent period the deviator is punished and gets the Cournot payoff 1/9. Hence, the sequential optimality of the grim trigger strategy holds when the following inequality is true:

$$\frac{1}{8} \geq (1-\delta)\frac{9}{64} + \delta\frac{1}{9} \Longleftrightarrow \delta \geq \underline{\delta} = \frac{9}{17}. \tag{4}$$

Q.E.D.

Note that $\underline{\delta}$ will remain the lower bound for δ in the imperfect monitoring case, because the deviation described in the proof of Result 1 is also possible in any other case with expected short run payoffs of the same form as in the perfect monitoring case.

4 Public monitoring and public grim trigger strategy

Under public monitoring both firms sell their products for the same random price. In any time period t, the firm knows its own production levels in all previous periods and all past price realizations. A firm does not know the production of its opponent, but does know that its opponent's realization price is the same. For this reason, the price allows the firms to synchronize their actions. In other words, if the strategy of a firm only depends on only the price realization, firms can cooperate while observing "good" prices and initiate an immediate punishment after any "bad" price.

Now suppose that both firms produce the cooperative outcome $(Q_P, Q_P) = (1/4, 1/4)$ at some period $t-1$. According to formula (1), the realization of the price will be in $[1-2Q_P-r, 1-2Q_P+r] = [1/2-r, 1/2+r]$. Any realization of price below *the threshold level* $1-2Q_P-r$ means that the rival has produced more than Q_P and as a result she can be punished by moving to the Cournot outcome $(Q_C, Q_C) = (1/3, 1/3)$ forever.[10] By formula (1), the price is below $1-2Q_P-r$ if and only if the production q^i_{t-1} of firm i plus

[10] The firm may also punish for too high price realizations, ones which are higher than $1-2Q_P+r$. However, such action is not optimal since these signals are generated by low production, which decreases the opponent's profit.

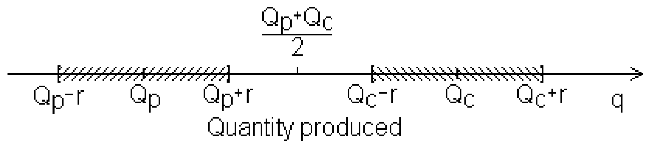

Figure 1: Signal realizations with non overlapping supports. We have drawn the possible signal realizations when a player's opponent plays either Q_P or Q_C. The corresponding supports are separated. Hence if the player knows for sure that one of Q_P and Q_C has been played she can correctly guess the right one.

the "estimate" of its rival's production, $\tilde{q}^j_{t-1}$, is less than $2Q_P + r$. We use this observation to construct the public grim trigger strategy as a minor adaptation of the Green and Porter (1984) strategy.

Public grim trigger strategy:

$$q^i_t = s(q^i_{t-1}, \tilde{q}^j_{t-1}) = \begin{cases} Q_P & \text{if } q^i_{t-1} + \tilde{q}^j_{t-1} \leq 2Q_P + r \\ Q_C & \text{otherwise} \end{cases} \tag{5}$$

The above strategy helps to support the cartel outcome for the whole duration of the game. Now the only question that remains is, for which values of δ and r does the above strategy constitute a Perfect Bayesian Equilibrium in Pure Strategies?

From now on we will impose a restriction on highest possible level of r. This restriction dramatically simplifies the equilibrium analysis and as it is shown below is not an "active" restriction for the main result of the paper. The model requires that the support of the signal under the opponent's play Q_P does not overlap with the support under Q_C. So for the rest of the paper we consider noise such that

$$r \leq \frac{Q_C - Q_P}{2} = \frac{1}{24}. \tag{6}$$

Figure 1 provides a graphical presentation at this assumption.

The result below describes the set of values for δ and r for which the strategy under interest supports an equilibrium.

Result 2: *Given restriction (6), the Public Grim Trigger Strategy is a Perfect Bayesian Equilibrium in Pure Strategies strategy for any*

$$\delta \geq \underline{\delta} = \frac{9}{17} \text{ and } r \leq \frac{\delta}{36(1-\delta)}.$$

The borderline for r is an increasing function of δ. The intuition is the following. Suppose we increase the noisiness of the signal to $r' > r$. This lowers the probability of punishment for the same level of deviation because the density of the noise becomes smaller. So to make r' a border point we need to increase the discount factor δ to some $\delta' > \delta$. In this case the level of long run punishment becomes higher and the short run benefits get smaller to compensate having a lower probability of punishment.

PROOF: Any output level produced as a deviation is denoted by q. There are two types of deviations. The first type is a *"large" deviation,* with $q \geq Q_P + 2r$. This deviation is detected by the opponent for sure, so the rest of the game continues with the certain outcome (Q_C, Q_C). Since the one period best response to $Q_P = 1/4$, which is 3/8, is itself a large deviation, the analysis is equivalent to that of the perfect monitoring case, which yields the restriction on δ given by condition (4).

The second type of deviation is a *"small" deviation.* This happens when $Q_P \leq q \leq Q_P + 2r$ or $q \in \left[\frac{1}{4}, \frac{1}{4} + 2r\right]$. In this situation, with probability $\frac{q-Q_P}{2r}$ the price for the current period will be lower than the threshold level, and the game will end up in the noncooperative outcome (Q_C, Q_C) which gives the one period payoff 1/9. Otherwise, the game will continue in the cooperative phase with payoff 1/8 as prescribed by the equilibrium strategy. For our strategy to be optimal the set of following inequalities must hold for all such q's.

$$\frac{1}{8} \geq (1-\delta)(1-1/4-q)q + \delta\left\{\left(1 - \frac{q-1/4}{2r}\right)\frac{1}{8} + \frac{q-1/4}{2r}\frac{1}{9}\right\}.$$

The left side of inequality represents the equilibrium payoff. On the right side, which is payoff from deviation, the first summand stands for the payoff during the deviation period and the second summand corresponds to the expected continuation payoff after deviation.

Now notice that the deviation payoffs are concave in q and equal to 1/8 when $q = 1/4$.

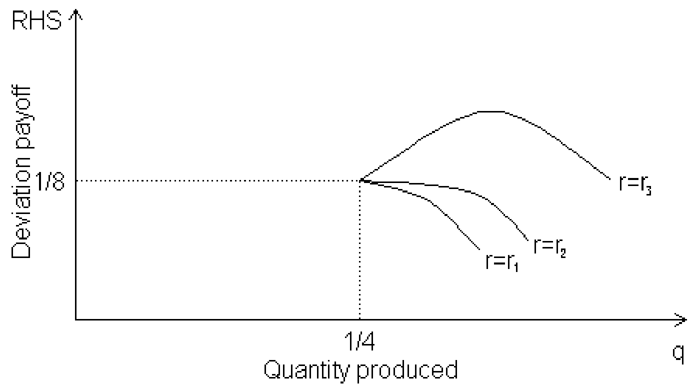

Figure 2: Deviation payoffs for public monitoring. We have drawn the deviation payoff for three different values of r ($r_1 < r_2 < r_3$) under fixed δ. The higher r the larger the payoff from playing q. The line with zero derivative w.r.t. q at $q = 1/4$ corresponds to $r = r_2$. For any $r > r_2$ the payoff is higher than $1/8$ for some $q > 1/4$ and for $r \leq r_2$ the payoff from deviation is always less than $1/8$.

Hence, the derivative of these payoffs with respect to q at $q = 1/4$ must be non-positive or

$$(1-\delta)\frac{1}{4} - \delta\frac{1}{144r} \leq 0 \Longleftrightarrow r \leq \frac{\delta}{36(1-\delta)}. \tag{7}$$

The specific form of the relationship between r and δ arises because given q the deviation payoff is an increasing function of r. This happens because the probability of punishment $(q - 1/4)/(2r)$ (see Figure 3) becomes smaller the larger is r. We draw the borderline for r in Figure 8 (r_2 on Figure 2 is a point of the borderline).

The only thing left to specify is the players' beliefs. Clearly, the trivial belief that the opponent will play Q_P given the game has been in cooperation ($q^i_{t-1} + \tilde{q}^j_{t-1} \leq 2Q_P + r$) and Q_C otherwise, makes our strategy optimal. This belief is consistent with Bayes' rule on the equilibrium path of game. On this path every player always plays Q_P and observes that $\tilde{q}^j_{t-1} \leq Q_P + r$. The proof is complete by specifying that off the equilibrium path a firm believes that its opponent will play Q_C.[11] Q.E.D.

Notice that the lowest level of the signal's noisiness (level of r) for which the cooperation outcome cannot be sustained takes place for the lowest level of the discount factor, and is

[11] We are free to choose any beliefs off the equilibrium path.

equal to 1/36. This means that if r is no more then one third of the distance between the cooperative and noncooperative production level then collusion is sustainable for the same range of the discount factor as in the perfect monitoring environment.

5 Private monitoring and private grim trigger strategy

Now consider the private monitoring case, in which the signals are independent across players. Recall that shocks that satisfy restriction (6) are only considered. These values of r are such that any estimate of the opponent's play when she actually plays $Q_P = 1/4$ differs from any estimate when her opponent actually plays $Q_C = 1/3$.

For this informational situation it is possible to construct an analog of the grim trigger strategy. Suppose that both firms agree to support the Pareto efficient outcome by producing the same output Q_P, which gives the one-period expected profit $\pi_P = 1/8$ to each firm. As with perfect monitoring, each firm has the temptation to increase production in order to increase its current profit. As in the public monitoring case, two types of deviation are identified.

The first type of deviation is the "large" one. This happens when the player produces an output q higher than $Q_P + 2r$. In this case the opponent's estimate of the player's action will lie between $q - r$ and $q + r$, while under cooperative play Q_P the signal would be in the range $[Q_P - r, Q_P + r]$. Hence the opponent for sure will detect the deviation and be able to punish the rival by producing the noncooperative outcome Q_C. Because the level of the noise r is restricted (restriction (6)) to make the Cournot outcome be a "large" deviation, the punishment will definitely be recognized by the deviator. The rest of the game will continue in the noncooperative state (Q_C, Q_C), which will give the one-period expected profit $\pi_C = 1/9$ to each firm. In short the large deviations can be punished exactly in the same way as in the perfect monitoring case, so we can apply the analysis that already exists in the literature.

The situation is different when the player tries to make a "small" ($Q_P \leq q \leq Q_P + 2r$) deviation. Suppose a player (deviator) decided to produce more than Q_P, say, q. This action induces a signal $\tilde{q}$ not known by the player but received by the opponent which might be lower or larger than $Q_P + r$. If the signal's realization is less than this value then the opponent will not detect the deviation and no punishment will take place. When the signal $\tilde{q}$ is higher than $Q_P + r$, the opponent will know for sure that a deviation occurred, and she can implement the long run punishment by producing Q_C forever. The period after the deviation, the deviator will find out for sure that the deviation has been detected. Therefore, she will have to play Q_C from then on.

The above analysis implies that it might be possible to look for a strategy where any detected deviation ($\tilde{q} > Q_P + r$) is immediately punished by the opponent by playing Q_C for the rest of the game. Let us now find out what is rational for the deviator to play in the next period after deviation under these circumstances.

To simplify the exposition for the remainder of the section we refer to the subsequent period after the deviation as the current period t.

While previously playing $q > Q_P$ the deviator (player i) "alarmed" the opponent with probability

$$\alpha(q) = \text{Prob}\{\tilde{q}^i > Q_P + r | q^i = q\},$$

which has the simple functional form given by

$$\alpha(q) = \frac{q - Q_P}{2r}. \tag{8}$$

Figure 3 provides a description of $\alpha(q)$.

From now on we will refer to α simply as "the detection probability". Notice that α uniquely specifies the last period deviation q by equation (8), so we can safely deal with the probability instead of the action and use q for the current period action. Now let us analyze what the player might do after producing the high output level that induced detection probability α. Before doing so, let us introduce notation for the expected one period payoff during the period after the deviation. The function $\pi(q, \alpha)$ is the short notation for the expected payoff from playing q in the current period when the opponent mixes between Q_P and Q_C with probabilities $(1 - \alpha)$ and α respectively. This function has the following form

$$\pi(q, \alpha) = (1 - \alpha)(1 - Q_P - q)\, q + \alpha (1 - Q_C - q)\, q. \tag{9}$$

Figure 4 shows the shape of function $\pi(q, \alpha)$.

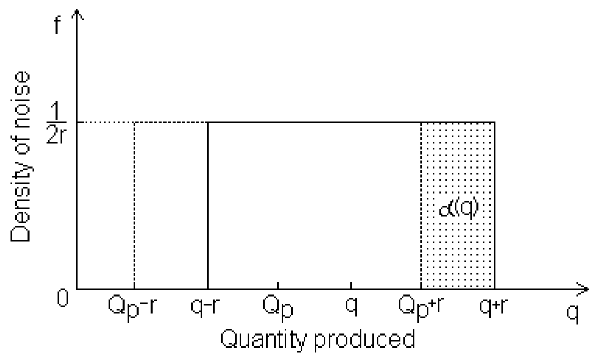

Figure 3: The probability of detection. The shaded area denotes the probability of being detected after a deviation to q. This probability equals probability of getting a signal in the range $[Q_P + r, q + r]$, which is $\alpha(q) = (q - Q_P)/(2r)$.

One possibility is to play Q_P, the "penitent" action. Then in the next period the deviator finds out with probability α that she is punished, and both players will play the Cournot outcome (Q_C, Q_C) for the rest of the game. With probability $(1 - \alpha)$ the game continues in the cooperative phase (Q_P, Q_P). The decision to play Q_P gives the expected payoff

$$\Pi_P(\alpha, \delta) = (1 - \delta)\pi(Q_P, \alpha) + \delta\left[(1 - \alpha)\pi_P + \alpha\pi_C\right], \tag{10}$$

where the subscript P denotes the penitent action payoff.

Another possibility is to produce outcome $\hat{Q}(\alpha)$ which maximizes her expected current period profit given by formula (9). By taking the derivative of $\pi(q, \alpha)$ with respect to q, and setting this derivative equal to zero, we get the expression for $\hat{Q}(\alpha)$:

$$\hat{Q}(\alpha) = (1 - \alpha)\frac{1 - Q_P}{2} + \alpha\frac{1 - Q_C}{2}. \tag{11}$$

The level of production $\hat{Q}(\alpha)$ is always no less than the Cournot level Q_C for any value of the probability α. So this kind of action is a large deviation, and the game will definitely continue at the Cournot level afterwards. This deviation will be detected for sure, and the expected continuation payoff is

$$\Pi_F(\alpha, \delta) = (1 - \delta)\pi(\hat{Q}(\alpha), \alpha) + \delta\pi_C, \tag{12}$$

where the subscript F refers to "the fatal action" $\hat{Q}(\alpha)$.

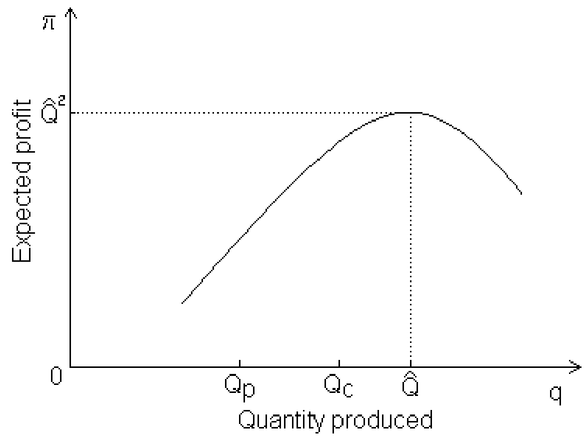

Figure 4: Expected one period profit. The one period expected profit, $\pi(q,\alpha)$, is a parabola with negative curvature. The maximum is at point $\hat{Q}(\alpha) = (1-\alpha)\frac{1-Q_P}{2} + \alpha\frac{1-Q_C}{2}$ with value $\hat{Q}(\alpha)^2$.

If the player is restricted in choosing between the penitent action Q_P or the fatal action $\hat{Q}(\alpha)$, she will pick the one which will give her the highest continuation payoff. Intuitively, when the detection probability α is low enough one should expect that the action Q_P will give a higher payoff, while the action $\hat{Q}(\alpha)$ is preferable for a high level of α. This means that there is a critical value of the probability under which the penitent and fatal actions give the same continuation payoff. We notate this probability as $\alpha^*(\delta)$. It comes as the solution to the system $\Pi_P(\alpha^*,\delta) = \Pi_F(\alpha^*,\delta)$. By solving the this system we get

$$\alpha^*(\delta) = \frac{4\sqrt{2\delta^2-\delta}-7\delta+3}{1-\delta}. \tag{13}$$

The graph of $\alpha^*(\delta)$ is shown on Figure 5.

If the deviator plays Q_P when $\alpha \leq \alpha^*(\delta)$ and $\hat{Q}(\alpha)$ otherwise, then her opponent's next period reaction Q_C to the signal observed in the range $(Q_P + r, Q_P + 3r]$ can be justified by the belief that the deviator plays action $Q_P + 2r$ in the previous period. This action generates the detection probability $\alpha = 1$ and so in the current period the deviator's best response is $\hat{Q}(1) = Q_C$. Therefore, in the next period the punisher is consistent by playing Q_C while observing the signal in the range $(Q_P + r, Q_P + 3r]$: she plays a best response to her rival's action Q_C. This finishes the construction of the joint behavior of the players with self-sustained beliefs.

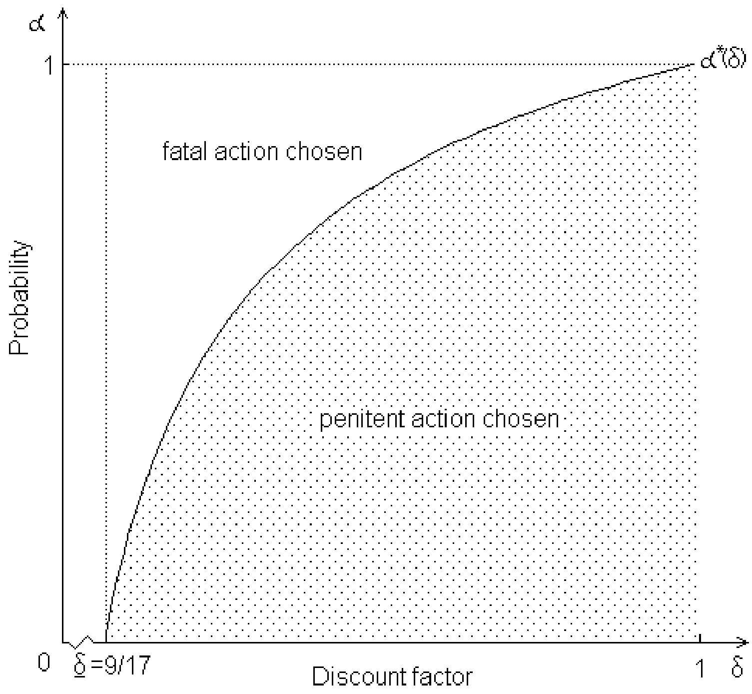

Figure 5: Probability Cutoff Level $\alpha^*(\delta)$. The threshold level of the probability $\alpha^*(\delta)$ is monotonically increasing in δ. This happens because the higher the δ, the more significant the long run punishment compared to the short run benefits from deviation. For the same reason $\alpha^*(\underline{\delta}) = 0$ and $\lim_{\delta \to 1} \alpha^*(\delta) = 1$. The shaded area shows the value of α for which the deviator will want to produce Q_P, while for α above the shaded area the deviator produces $\hat{Q}(\alpha)$.

Having described the key features of the equilibrium strategy, we can present them formally. The action q_t^i of player i at time period t depends only on the information received at period $t-1$ $(q_{t-1}^i, \tilde{q}_{t-1}^j)$ and has the following form

Private grim trigger strategy:

$$s(q_{t-1}^i, \tilde{q}_{t-1}^j) = \begin{cases} Q_P & \text{if } \alpha(q_{t-1}^i) \leq \alpha^*(\delta) \text{ and } \tilde{q}_{t-1}^j - Q_P \leq r \\ \hat{Q}\left(\alpha(q_{t-1}^i)\right) & \text{if } \alpha(q_{t-1}^i) > \alpha^*(\delta) \text{ and } \tilde{q}_{t-1}^j - Q_P \leq r \\ Q_C & \text{otherwise} \end{cases} \tag{14}$$

where functions $\alpha(\cdot)$, $\alpha^*(\cdot)$ and $\hat{Q}(\cdot)$ are given by formulae (8), (13) and (11), respectively.

This strategy is presented on Figure 6.

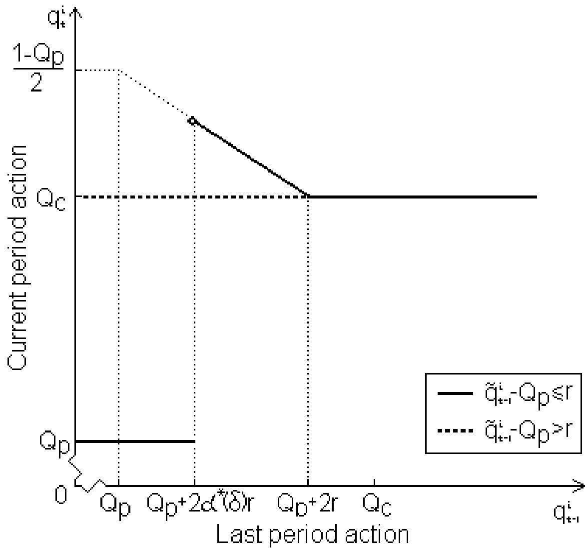

Figure 6: Private Grim Trigger Strategy $q_t^i = s(q_{t-1}^i, \tilde{q}_{t-1}^j)$. The dotted line shows the strategy played by the player when the signal is "bad", i.e. if a deviation of the opponent has been detected ($\tilde{q}_{t-1}^j - Q_P > r$). In this case, the player always plays the Cournot outcome Q_C. The solid line shows the strategy under a "good" signal ($\tilde{q}_{t-1}^j - Q_P \leq r$). The player follows the collusive agreement Q_P when its last period production q_{t-1}^i generates a probability of being punished by its opponent no higher than the threshold level $\alpha^*(\delta)$ (left part of the graph). In the case of a certain detection, the player plays Q_C (right part of the graph). The middle part is a linear combination of one period best responses Q_C and $\frac{1-Q_P}{2}$ on Q_C and Q_P, respectively. The weights are equal to the probabilities of these actions played by the opponent.

We have already described two possible actions that the deviator may choose after her own deviation. By previous construction the equilibrium payoff function is the maximum of the penitent action payoff function (10) and the fatal action payoff function (12)

$$\Pi^*(\alpha, \delta) = \max\{\Pi_P(\alpha, \delta), \Pi_F(\alpha, \delta)\}. \tag{15}$$

Figure 7 shows the equilibrium payoff function.

Another possible action after the deviation is to play some "small" deviation.[12] This action induces a detection probability β which can calculated in exactly the same fashion as

[12]By the construction $\hat{Q}(\alpha)$ gives highest continuation payoff among all large deviations.

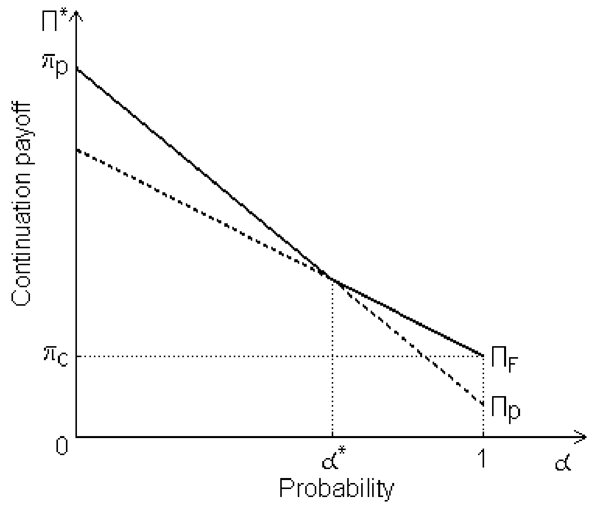

Figure 7: Equilibrium Payoff Function $\Pi^*(\alpha, \delta)$. The solid line marks the equilibrium payoff function, $\Pi^*(\alpha, \delta)$, which is the maximum of the penitent action payoff, $\Pi_P(\alpha, \delta)$, and the fatal action payoff, $\Pi_F(\alpha, \delta)$. $\Pi^*(\alpha, \delta)$ is decreasing function of α. When δ goes up function Π_P pivots up around point the $(0, \pi_P)$ and function Π_F is "pressed" down to the level π_C.

α. Given that the players will follow the equilibrium strategy (14) afterwards, the calculation of the expected payoff after two small deviations with probabilities α and β is analogous to the calculation of the expected payoffs $\Pi_F(\alpha, \delta)$ and $\Pi_P(\alpha, \delta)$ and is equal to

$$\Pi'(\alpha, \beta, \delta, r) = (1 - \delta)\pi(Q_P + 2\beta r, \alpha) + \delta\left[(1 - \alpha)\Pi^*(\beta, \delta) + \alpha\pi_C\right]. \tag{16}$$

It turns out that the payoff functions given by formulae (15) and (16) are enough to characterize the set of δ and r for which the above strategy constitutes the Nash Equilibrium.

We would also like to show that this equilibrium is one in which players' out of equilibrium beliefs are reasonable. Unfortunately, Kreps and Wilson's definition of Sequential Equilibrium (1982) only pertains to the finite games - games with a finite number of information sets and a finite number of actions available at each information set. Our game is not finite in either of these aspects. We therefore introduce a consistency principle in the spirit of Kreps and Wilson to get a refinement of a Perfect Bayesian Equilibrium. The following definition uses the same concepts as the definition of Sequential Equilibrium applied and it will be applied to our game.[13] We call it:

[13] Similar concepts were used, for example, in Chatterjee and Samuelson (1990) and Simon and Stinchcombe

Definition 1: *Nash Equilibrium with Consistent Beliefs in Pure Strategies:*

Let H denote the collection of information sets of the game and let $s{:}H \rightarrow A(H)$ specify the pure strategy behavioral strategies of players. Then these strategies yield an equilibrium if there exists an assessment (μ, s)[14] consisting of beliefs μ and behavioral strategies s with the following two properties:

i) Sequential rationality: for the given beliefs, the behavioral strategies yield the largest continuation payoffs at each information set.

ii) Consistency: Let s_n be a set of completely mixed behavioral strategies. By the notion of complete mixing we mean that for any information set h, $\text{supp}(s_n) \equiv A(h)$. For each s_n, Bayes' rule uniquely specifies beliefs μ_n. Consistency means that there is a weakly converging sequence of completely mixed strategy profiles s_n to s such that the induced beliefs μ_n weakly converge to μ. In short, $(\mu, s) = \lim_{n \rightarrow \infty}(\mu_n, s_n)$.

The application of this concept yields the central result of the paper.[15]

Result 3: *Fix a pair (δ, r). The private grim trigger strategy described by (14) gives a Nash Equilibrium with Consistent Beliefs in Pure Strategies if and only if following set of inequalities holds*

$$\Pi^*(\alpha, \delta) \geq \Pi'(\alpha, \beta, \delta, r) \; \forall \alpha, \beta \in [0, 1], \tag{17}$$

where, the payoff functions Π^* and Π' are given by formulae (15) and (16).

The above result demonstrates the interesting features of the constructed strategy. First, any private history, regardless of its length and the sequence of observations, influences the player's actions only through α, the probability of punishment. Second, to check the optimality of the strategy we need to check only one period deviations which are characterized by the probability of inducing punishment β from them.

The next result gives a closed form description of the set of parameter pairs (δ, r) for which the private grim trigger strategy is an equilibrium one and shows that this set is

(1995).

[14] Kreps and Wilson use the symbol π instead of s.

[15] The proof is quite tedious and it is presented in Appendix $(A1)$.

nonempty.[16]

Result 4: *Let Σ be the set of pairs (δ, r) such that condition (17) holds. Then:*

i) $\exists\ \bar{r}(\delta) \in C[\underline{\delta}, 1)$ such that $\bar{r}(\delta) > \frac{1}{48}$ for $\forall \delta \in [\underline{\delta}, 1)$ and $\Sigma = \{(\delta, r) : \delta \in [\underline{\delta}, 1), 0 \leq r \leq \bar{r}(\delta)\}$.

ii) Moreover $\bar{r}(\delta) = \min\{r_1(\delta), r_2(\delta)\}$, where $r_1(\delta)$ and $r_2(\delta)$ are given by:

$$
\begin{aligned}
r_1(\delta) &= \frac{1}{48} \frac{\delta(3-\delta)\left(3\delta - 1 - 2\sqrt{2\delta^2 - \delta}\right)}{(1-\delta)(\delta - \sqrt{2\delta^2 - \delta})} \text{ and} \\
r_2(\delta) &= \frac{1}{204} \frac{9\delta - 9\sqrt{2\delta^2 - \delta} + 2\sqrt{2}\sqrt{\delta(17\delta - 9)\left(3\delta - 1 - 2\sqrt{2\delta^2 - \delta}\right)}}{1-\delta}.
\end{aligned}
$$

iii) $\bar{r}(\underline{\delta}) = \frac{9}{272}$, $\lim\limits_{\delta \to 1} \bar{r}(\delta) = \frac{1}{48}$.

The intuition behind the functions $r_1(\delta)$ and $r_2(\delta)$ is the following. Increasing the probability of detection α decreases the long run equilibrium payoff at a constant rate, until $\alpha = \alpha^*(\delta)$. From then on, the rate of decrease in the long run payoff drops since the player switches from the penitent action to the fatal action. We show in Appendix (A2) that the immediate benefit from deviating first offsets the long run costs when $\alpha = \alpha^*(\delta)$. Then, by imposing the condition that the penitent action ($\beta \leq \alpha^*(\delta)$) can not be profitable we get the $r_1(\delta)$ part of the curve $\bar{r}(\delta)$. The proof in the appendix shows that values of $\beta \simeq 0$ are critical, because the long run loss is linearly increasing with a increase in β, while the short run profits increase at a decreasing rate. Taking care of the fatal action ($\beta > \alpha^*(\delta)$) gives the $r_2(\delta)$ part of the curve $\bar{r}(\delta)$. In this case, the critical values of β are difficult to determine explicitly. Graphical analysis shows a negative relationship between δ and the critical value of β.

The function $\bar{r}(\delta)$ is depicted on Figure 8 along with the borderline for the public monitoring case.

From the picture we can see that the values of r for which cooperation can be sustained are comparable to the ones in the public monitoring case. Also for $r \leq 1/48$, the cooperative

[16] The proof of this result is technical and can be found in Appendix ($A2$).

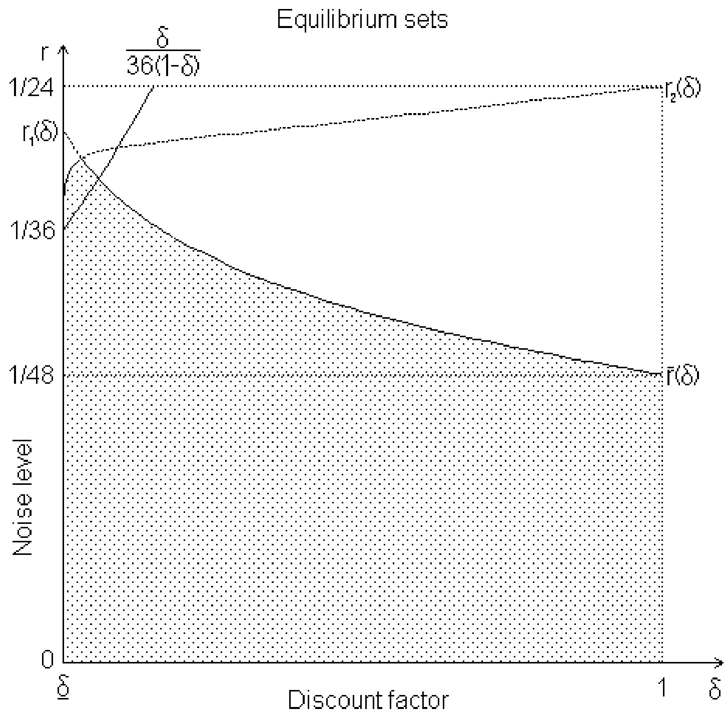

Figure 8: The sets for which the Public and Private Strategies are equilibrium ones. The shaded area marks possible pairs (δ, r) for which the private strategy gives an equilibrium. The line $\frac{\delta}{36(1-\delta)}$ restricts the area for the public monitoring case. The dotted lines show the curves $r_1(\delta)$ and $r_2(\delta)$.

equilibrium is sustainable for the same range of δ as in the perfect monitoring case. Notice that unlike under public monitoring, the borderline $\bar{r}(\delta)$ is not monotonic in δ. This is a feature of our particular modification of the grim trigger strategy. However, it might possible to find other modifications for this strategy which would make the borderline monotonic.

As mentioned above, restriction (6) on r has no effect on the set Σ because $\bar{r}(\delta) < 1/24$ for $\delta \in [\underline{\delta}, 1)$. Possible realizations of the signal under $r \leq \bar{r}(\delta)$ when the opponent plays either Q_P or Q_C are shown in Figure 9. We can see clearly that for any δ the supports do not intersect, so our restriction on possible values of r is not critical.

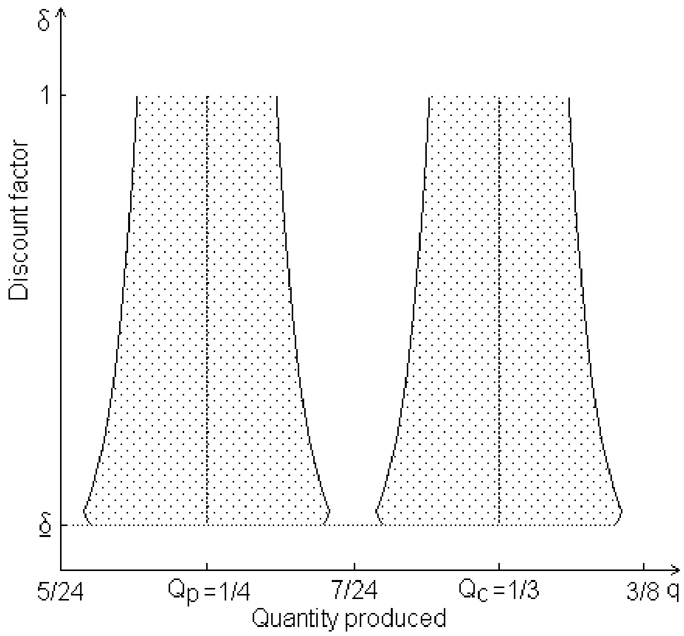

Figure 9: Possible signal realizations under private monitoring. The possible signal realizations are marked by the shaded area when the player's opponent plays either Q_P or Q_C, for r's, which make our private grim trigger strategy an equilibrium one. Note that the realizations are separated for $q = Q_P$ and $q = Q_C$.

6 Arbitrary distribution of noise

Our analysis so far has relied on the assumption that the signal noises are uniformly distributed. Surprisingly, the private monitoring results of this paper do not substantially change if this assumption is relaxed. We demonstrate these results without providing proofs.

Suppose that the noise, ε, is distributed on a limited support whose width equals $2r$, and has zero expected value. Let the value of the right end of the support be E_R. The noise can be fully described by the decumulative distribution function measured from the right end of the support:

$$F(x) \equiv \text{Prob}(\varepsilon > E_R - 2rx).$$

Note that $F(0) = 0$ and $F(1) = 1$. Assume, further, that this distribution function has a continuous density $f(x)$.

Now the analog of formula (8) which gives the detection probability from playing q is

$$\alpha(q; F(\cdot)) = F\left(\frac{q - Q_P}{2r}\right).$$

The equilibrium strategy is

$$s(q^i_{t-1}, \tilde{q}^j_{t-1}; F(\cdot)) = \begin{cases} Q_P & \text{if } \alpha(q^i_{t-1}; F(\cdot)) \le \alpha^*(\delta) \text{ and } \tilde{q}^j_{t-1} - Q_P \le E_R \\ \hat{Q}\left(\alpha(q^i_{t-1}; F(\cdot))\right) & \text{if } \alpha(q^i_{t-1}; F(\cdot)) > \alpha^*(\delta) \text{ and } \tilde{q}^j_{t-1} - Q_P \le E_R \\ Q_C & \text{otherwise} \end{cases}$$

where the functions $\alpha^*(\cdot)$ and $\hat{Q}(\cdot)$ are given by formulae (13) and (11).

Result 3 which gives the necessary and sufficient condition for having an equilibrium stays the same. In it, the function $\Pi'(\alpha, \beta, \delta, r)$ should be understood as

$$\Pi'(\alpha, \beta, \delta, r; F(\cdot)) = (1 - \delta)\pi(Q_P + 2F^{-1}(\beta)r, \alpha) + \delta\left[(1 - \alpha)\Pi^*(\beta, \delta) + \alpha\pi_C\right],$$

where $F^{-1}(\cdot)$ is inverse function of $F(\cdot)$.

As it will be shown shortly, we have the analog of Result 4 with nonempty set of pairs (δ, r) when we have *positive density of the right end of the support* ($F'(0) > 0$). Under this property of the distribution of the noise we have the borderline $\bar{r}(\delta; F(\cdot))$ as the minimum of two continuous functions, $r_1(\delta; F(\cdot))$ and $r_2(\delta; F(\cdot))$, which can assume any nonnegative values. In general, it is impossible to find the shape of the functions $r_1(\delta; F(\cdot))$ and $r_2(\delta; F(\cdot))$ explicitly. Still the curve $r_1(\delta; F(\cdot))$ is related to curve $r_1(\delta)$ as follows:

$$r_1(\delta; F(\cdot)) \le f(0)r_1(\delta).$$

If the density on the right end of support is equal to zero, then our modified grim trigger strategy no longer supports cooperation. This is because the short run profit gain from an infinitesimal increase of production dq is proportional to dq, while the expected long run costs are proportional to $F(dq)$. Therefore there always exists a small dq such that it will be optimal for a player to deviate.[17]

In the case of nonzero density on the right end of support it is possible to show that, analogous to the uniform case, $\bar{r}(\delta; F(\cdot)) > 0$ for any $\delta \in [\underline{\delta}, 1)$.

[17] The same problem takes place in the public monitoring situation.

As a concluding remark to this section, notice that the condition of nonzero density narrows the applicability of our approach. Still for most economic situations possible values of price and estimate of production are not continuous. The production and estimate usually take discrete values of the form $\{k\Delta : \Delta > 0,\ k = 0, 1, 2, ...\}$.[18] Given that any noise with bounded support always has nonzero density on the right end of the support, the direct translation of our approach for the discrete case yields cooperation with any distribution of signal with sufficiently small noisiness.[19]

7 Nash Reversion Folk Theorem

In this section we show that a private grim trigger strategy may also be used to support any feasible individually rational outcome with uniform density of noise.[20] By an individually rational outcome for our game we mean a pair of payoffs $(\bar{u}^1, \bar{u}^2)$ which gives players no less than the Cournot outcome $\pi_P = 1/9$ each.

Now it is necessary to choose a pair of actions which yield the pair of payoffs $(\bar{u}^1, \bar{u}^2)$. The candidates $(\bar{Q}^1, \bar{Q}^2)$ are solutions to the following system of equations:

$$\begin{aligned} \pi(\bar{Q}^1, \bar{Q}^2) &= \bar{u}^1 \\ \pi(\bar{Q}^2, \bar{Q}^1) &= \bar{u}^2 \end{aligned} \tag{18}$$

where function $\pi(\cdot, \cdot)$ is given by formula (3). The system (18) might exhibit multiple solutions, but if at least one solution exists we can always choose the pair where $(\bar{Q}^1, \bar{Q}^2) \leq (5/18, 5/18) < (Q_C, Q_C) = (1/3, 1/3)$.[21] In fact, one can establish the following claim, which we provide without proof:

Claim: *For any pair of individually rational feasible payoffs $(\bar{u}^1, \bar{u}^2) > (\pi_C, \pi_C)$ there exists a pair of production levels $(\bar{Q}^1, \bar{Q}^2)$ such that $\bar{Q}^i < 5/18 < Q_C = 1/3$ and $(\bar{Q}^1, \bar{Q}^2)$*

[18] For example, car and airline industries measure production in units.

[19] The situation with zero density on the right end of the support can also been remedied by allowing grim trigger strategy to be ε-equilibrium. The same type of equilibria are used in Fudenberg and Levine (1991) and Lehrer (1992).

[20] The Folk Theorem can naturally be extended for the noise with nonzero density on the right end of the support. For notational simplicity we demonstrate the result with the uniform noise.

[21] The corresponding closed form solution is $\bar{Q}^i = \frac{1-\sqrt{1-4(\bar{u}^i+\bar{u}^j)}}{4} + \frac{\bar{u}^i+\bar{u}^j}{1+\sqrt{1-4(\bar{u}^i+\bar{u}^j)}}$.

yield payoff $(\bar{u}^1, \bar{u}^2)$.

The strategy in (14) can be used as a starting point for constructing the strategy which generates our outcome as an equilibrium. For the chosen pair $(\bar{Q}^1, \bar{Q}^2)$ the following private grim trigger strategy for player i works.

$$s(q^i_{t-1}, \tilde{q}^j_{t-1}, \bar{Q}^i, \bar{Q}^j) = \begin{cases} \bar{Q}^i & \text{if } \alpha(q^i_{t-1}, \bar{Q}^i) \leq \alpha^*(\delta, \bar{Q}^i, \bar{Q}^j) \text{ and } \tilde{q}^j_{t-1} - \bar{Q}^j \leq r \\ \hat{Q}\left(\alpha(q^i_{t-1}, \bar{Q}^i), \bar{Q}^j\right) & \text{if } \alpha(q^i_{t-1}, \bar{Q}^i) > \alpha^*(\delta, \bar{Q}^i, \bar{Q}^j) \text{ and } \tilde{q}^j_{t-1} - \bar{Q}^j \leq r \\ Q_C & \text{otherwise} \end{cases} \tag{19}$$

Here, the function $\alpha(q, \bar{Q}^i)$ is the probability that player i is detected by the opponent when playing q instead of $\bar{Q}^i$. $\hat{Q}\,(\alpha, q)$ is the short run best response while the opponent mixes q and Q_C with probabilities $(1-\alpha)$ and α respectively.

The function $\alpha^*(\delta, \bar{Q}^i, \bar{Q}^j)$ is the cutoff probability level which equalizes the payoffs from playing penitent action $\bar{Q}^i$ and fatal action $\hat{Q}\left(\alpha, \bar{Q}^j\right)$ after a the deviation with a detection probability α. We do not present the formula for $\alpha^*(\delta, \bar{Q}^i, \bar{Q}^j)$ because it is quite bulky and there is no need for the closed form of this function. The payoffs from the penitent and the fatal actions are, respectively

$$\Pi_P(\alpha, \delta, \bar{Q}^i, \bar{Q}^j) = (1-\delta)\pi(\bar{Q}^i, \alpha, \bar{Q}^j) + \delta\left[(1-\alpha)\bar{u}^i + \alpha\pi_C\right] \tag{20}$$

and

$$\Pi_F(\alpha, \delta, \bar{Q}^i, \bar{Q}^j) = (1-\delta)\pi(\hat{Q}(\alpha, \bar{Q}^j), \alpha, \bar{Q}^j) + \delta\pi_C, \tag{21}$$

where $\pi(q, \alpha, \bar{Q}^j)$ is one period expected payoff from playing q when the opponent mixes $\bar{Q}^j$ and Q_C with probabilities $(1-\alpha)$ and α respectively. Again the equilibrium payoff function $\Pi^*(\beta, \delta, \bar{Q}^i, \bar{Q}^j)$ is maximum of $\Pi_P(\alpha, \delta, \bar{Q}^i, \bar{Q}^j)$ and $\Pi_F(\alpha, \delta, \bar{Q}^i, \bar{Q}^j)$.

As in Section 5 we need to care about payoffs which arise from initiating "small" deviations with probability β. Analogously we get

$$\Pi'(\alpha, \beta, \delta, r, \bar{Q}^i, \bar{Q}^j) = (1-\delta)\pi(\bar{Q}^i + 2\beta r, \alpha, \bar{Q}^j) + \delta\left[(1-\alpha)\Pi^*(\beta, \delta, \bar{Q}^i, \bar{Q}^j) + \alpha\pi_C\right]. \tag{22}$$

We can now describe our Folk Theorem.

Result 5 (Nash Reversion Folk Theorem):

i) Fix a pair of parameters (δ, r) *and a pair of Nash dominating payoffs* $(\bar{u}^1, \bar{u}^2) > (\pi_C, \pi_C)$. *The strategy given by (19) gives a Nash Equilibrium with Consistent Beliefs in Pure Strategies if and only if the following system of inequalities holds*

$$\Pi^*(\alpha, \delta, \bar{Q}^i, \bar{Q}^j) \geq \Pi'(\alpha, \beta, \delta, r, \bar{Q}^i, \bar{Q}^j), \ \forall \alpha, \beta \in [0, 1]. \tag{23}$$

Here the payoff functions are given by formulae (20-22) and pair $(\bar{Q}^i, \bar{Q}^j)$ *are those in the claim above.*

ii) For any pair of individually rational feasible playoffs $(\bar{u}^1, \bar{u}^2) > (\pi_C, \pi_C)$ *there exists an* $\epsilon = \epsilon(\bar{u}^1, \bar{u}^2) > 0$ *that for any* $\delta \geq 1 - \epsilon$ *and* $r \leq \epsilon$ *the system of inequalities (23) holds.*

Result 5 shows that for any individually rational feasible outcome there exists an equilibrium strategy that support this outcome if the discount factor is sufficiently high and the range of price realizations are sufficiently concentrated. The first statement of this result can be proved exactly in the same way as Result 3. Given the first part of the result, the proof of the second part is provided in Appendix $(A3)$.

8 Conclusion

This paper demonstrated that full cooperation is possible in a repeated Cournot Duopoly with private monitoring. The key property of the model sustaining this result is that the support of the price signals depends on the players' actions. We showed that for a moderate level of noise, the monopoly outcome can be supported for the same range of the discount factor as in the perfect monitoring case, and we established a Folk Theorem for our model.

We conclude the paper by discussing some possible directions of future research and complications which may arise along the way.

In Section 6 it was shown that the strategy we constructed fails when the density of the noise on the right end of the support is zero. Further analysis shows that in general the situation cannot be remedied by applying a grim trigger strategy which initiates a punishment phase after realizations of the signal close to the right end of the support. The difficulty is that there are always infinitesimally small profitable deviations involving either increased or decreased production. Hence, the situation in which the noise has zero density noise on the right end of its support needs a more elaborate investigation.

One might develop another modification of the grim trigger strategy which allows sustaining the monopoly outcome for noisiness of the signal which is larger than the ones established in this paper. One possibility is to impose a more severe punishment for deviation than moving to the Cournot outcome.

Additionally, the model can also be adapted to study cases with more than two firms. Here every firm observes the estimate of the total production of the rest of the industry. Similar to the duopoly case, the non-deviator plays the monopoly quantity while the signal is "good" and punishes using the Cournot quantity when a deviation has been detected.[22] The deviator plays the penitent action when the detection probability is lower than the threshold level, and plays a one period best response when the detection probability is higher. The primary difficulty of moving to the n firm case is not conceptual but computational: for example, the threshold probability is the solution of a polynomial of order of the number of firms minus one.

Another direction of research is to allow for correlation between signals. It would be interesting to see how the equilibrium set Σ from Section 5 changes while the correlation between signals increases from 0 to 1. Notice that this is a gradual transformation of the model from private monitoring to public monitoring situation.

The presence of nonlinear demand does not affect the applicability of the our model. In this case the expected price becomes dependent on the signal noisiness, which makes the monopoly and Cournot outcome sensitive to the level of noise of signal. The other components of the model and strategy construction stay intact.

Throughout the paper we imposed a restriction on the highest possible level of noisiness. It was assumed that the monopoly and Cournot levels of production are observationally different to allow for the construction of beliefs with one period memory. However, when consider overlapping supports, including infinite support, the same signal realizations for cooperative and non-cooperative actions of the other firm exist. In this case, beliefs will be functions of the firm's whole set of private information. This situation, which is more

[22]Note that the identity of the deviator is not known to the firm.

complex than the one presented in this paper, requires further research.

9 Appendix

A1. Proof of Result 3

Proof that the strategy described by formula (14) gives a Nash Equilibrium with Consistent Beliefs in Pure Strategies as given by definition 1. The proof is provided in a few steps.

Step 1: *The structure of a game history and an information set.*

Suppose the players have played t periods. According to the game structure, any t period history h_t can be represented as a set of the following t quadruples

$$h_t = \{(q_0^1, \tilde{q}_0^2, q_0^2, \tilde{q}_0^1), ..., (q_{t-1}^1, \tilde{q}_{t-1}^2, q_{t-1}^2, \tilde{q}_{t-1}^1)\} \in R^{4t}.$$

Fix a player i, and call her opponent j. After any time period $\tau < t$, player i observes only her action q_τ^i and the signal $\tilde{q}_\tau^j$ about the action q_τ^j of player j. So the information set[23] of player i at time t, h_t^i, can be represented by the following set

$$h_t^i = \Pi_{\tau=0}^{t-1}\{q_\tau^i\} \times \{\tilde{q}_\tau^j\} \times [\tilde{q}_\tau^j - r, \tilde{q}_\tau^j + r] \times [q_\tau^i - r, q_\tau^i + r] \in R^{4t},$$

where multiplication means the Cartesian product and the sequence of pairs $\{(q_\tau^i, \tilde{q}_\tau^j),\ \tau = 0, ..., t-1\}$ represents all of player i's observations at period t.

Definition 1 requires beliefs $\mu(h_t^i)$ for player i for every information set h_t^i. These beliefs should be sequentially consistent. This is the scope of the next two steps.

Step 2: *, Construction of a sequence of completely mixed behavioral strategies as required by Definition 1.*

Let $f(x; \omega, \sigma^2)$ denote the density function of the normally distributed random variable with expectation ω and variance σ^2. Then for each information set h_t^i we consider the totally

[23] The information set reflects a partitioning of all possible game histories according to player i's private information.

mixed strategy with the following density functions:

$$s_n(q_t^i; h_t^i) = \frac{n^2 - 2n - 1}{n(n-1)} f\left(q_t^i; s(q_{t-1}^i, \tilde{q}_{t-1}^j), \frac{1}{n}\right) + \sum_{k=-\infty}^{+\infty} \left(\frac{1}{n}\right)^{|k|+1} f\left(q_t^i; Q_P + 2kr, \frac{1}{n}\right), \quad (24)$$

where $s(.,.)$ is the equilibrium strategy.

So a mixture of equilibrium strategies puts $\sum_{k=-\infty}^{+\infty} \left(\frac{1}{n}\right)^{|k|+1} = \frac{n+1}{n(n-1)} \left(\underset{n\to+\infty}{\to} 0\right)$ amount of probability around points of the following uniform grid $\{Q_k = Q_P + 2kr, k = -\infty ... + \infty\}$ and the rest of the probability around the quantity prescribed by the equilibrium strategy.

By the construction $\text{supp}(s_n(q_t^i; h_t^i)) \equiv (-\infty, +\infty)$ and $\lim_{n\to\infty} s_n(q_t^i; h_t^i) = \delta(q_t^i - s(q_{t-1}^i, \tilde{q}_{t-1}^j))$,[24] what is required by definition 1.

Step 3: *Calculation of the limiting beliefs of a player for each information set.*

The sequence of totally mixed behavioral strategies is given by formula (24). For a given information set h_t^i , by Bayes' rule player i's beliefs $\mu_n(h_t^i)$ can be uniquely determined. Moreover, the limiting belief $\mu(h_t^i)$ may be found. $\mu(h_t^i)$ is in fact a density function over h_t^i, which shows how likely it is for the game to be at a certain node of h_t^i. This distribution, together with a particular form of the equilibrium strategy, gives a distribution of possible actions for player j at period t. Let us denote this distribution function of actions of j as $\tilde{\mu}(q_t^j; h_t^i)$. Any of functions $\tilde{\mu}$ and μ are sufficient for checking the optimality of player i's behavioral strategy. Unlike μ the description of $\tilde{\mu}$ is simple. From now on without loss of generality we deal with the distribution of possible actions of player j: $\tilde{\mu}(q_t^j; h_t^i)$. Lemma 1 summarizes the results for the current step.

The first statement of Lemma 1 shows that for any h_t^i the density function $\tilde{\mu}(q_t^j; h_t^i)$ describes a random action which has only two possible realizations $\{Q_P, Q_C\}$. In other words, to characterize the density function $\tilde{\mu}(q_t^j; h_t^i)$ we need a scalar function, say, $\alpha(h_t^i)$ which specifies the probability that player j's plays Q_C.[25]

The second result of Lemma 1 gives a closed form solution for the function $\alpha(\cdot)$. More importantly, it states that the probability of player j's playing Q_C at time t only depends on player i's private observation in period $t-1$.

[24] $\delta(x - a)$ is called delta function and it denotes the density of a deterministic random variable which takes the value a.

[25] $\alpha(q)$ as presented in paper is this probability.

Lemma 1 *i)* $\forall h_t^i$ *supp*$\{\tilde{\mu}(q_t^j; h_t^i)\} \subset \{Q_P, Q_C\}$.

$$ii)\ Prob(q_t^j = Q_C | h_t^i) = \begin{cases} 0 & \text{if } q_{t-1}^i \leq Q_P \text{ and } \tilde{q}_{t-1}^j - Q_P \leq r \\ \frac{q_{t-1}^i - Q_P}{2r} & \text{if } 0 < q_{t-1}^i - Q_P \leq 2r \text{ and } \tilde{q}_{t-1}^j - Q_P \leq r \\ 1 & \text{otherwise} \end{cases}$$

PROOF:

i) This statement we prove by using the principle of mathematical induction on t. For $t = 0$ it can be seen that $\sup p(\tilde{\mu}(q_0^j; h_0^i)) = \{Q_P\}$ because the mixture of behavioral strategies at time zero should converge to the equilibrium strategy. Therefore, it will be expected that player j plays Q_P for sure.

Now suppose that the statement (i) of the lemma is true for any $t \leq T-1$. Consider the information set h_T^i. By the inductive hypothesis, the beliefs $\mu_n(h_{T-1}^i)$ converge to the beliefs $\tilde{\mu}(q_{T-1}^j; h_{T-1}^i)$ which leave player j with two possible actions, Q_P and Q_C , in period $T-1$. So the behavioral mixtures will be "concentrated" around those two actions. In the limit as $n \to \infty$ the Bayesian rule together with the signal q_{T-1}^j gives player i what we call the "intermediate" belief that (at $T-1$) player j has played for sure either $Q_P + 2kr$ $(k \in Z)$ (see the formula (24)) or Q_C at $T-1$. This intermediate belief arises due to the additional information that arrives as $\tilde{q}_{T-1}^j$ get realized. Now consider player j. By the equilibrium strategy, the action of player j at $T-1$, which we described above, and the signal $\tilde{q}_{T-1}^i$ determine the action of player j at period T. She should follow the equilibrium strategy (14). So it can be checked that this action can be either Q_P or Q_C.

ii) From the proof of the statement above it follows that player i's subjective beliefs, say, that if the following two events have occurred: a) player j played $Q_P + 2kr$ $(k \leq 0)$ in the previous period, and b) the signal player j received was less than $Q_P + r$, then player j plays Q_P. When player i gets the signal $\tilde{q}_{t-1}^j \leq Q_P + r$, she believes that (a) happened for sure. The probability of the signal being lower than $Q_P + r$ or event (b) taking place is $1 - \frac{q_{t-1}^i - Q_P}{2r}$. All of this gives the expression for the probability of firm j's playing the noncooperative level of production. Q.E.D.

Step 4: *Proof of the one-stage deviation principle.*

The one-stage deviation principle is heavily used in the next step. We present it in the following

Lemma 2 *(One-stage deviation principle) Consider a multistage game with incomplete information and restricted one-stage payoff set. Then for a behavioral strategy to be the best response it is necessary and sufficient to check that any one-period deviation at any information set does not increase the continuation payoff of the player who makes the decision at the information set considered.*

PROOF:[26]

Necessity. Follows immediately.

Sufficiency. We start the proof from the contrary statement and then we find a contradiction. Suppose that for a player starting on period t there exists a set of deviations from the behavioral equilibrium strategy such that the player's payoff is increased by $\varepsilon > 0$. Let $\tilde{t}$ stands for the last period when there is an information set where a deviation takes place. There are two cases: in the first $\tilde{t}$ is finite, while in the second $\tilde{t}$ is infinite.

Now we show that without loss of generality we can restrict attention to the first case. This is done by reducing the second case to the first one. Suppose $\tilde{t}$ is infinite. Let U_H be the highest possible one-stage payoff of the player. No deviation generates a one-stage payoff higher than U_H, which implies that if deviations take place after period t, then the highest possible payoff is no larger than $\delta^{t+1}U_H$. Given that the equilibrium expected payoff is bounded from below by U_L it is possible to find $\hat{t}$ that $\delta^{\hat{t}+1}U_H - \delta^{\hat{t}+1}U_L < \varepsilon/2$. As a result we can remove all deviations after period $\hat{t}$, consider deviations until period $\hat{t}$, and get the first case with $\tilde{t} = \hat{t}$. The improvement of the payoff is no less than $\varepsilon/2$ and so it is sufficient to examine only the first case.

[26] The proof of the lemma's result is provided in almost exactly the same way as the one for the "one-stage deviation principle" in infinitely played games with perfect information (see Fuderberg and Tirole (1991)). This is due to two reasons. First, we are dealing with information sets, which "evolve" in the same tree like fashion. Second, by dealing with the behavioral strategy at all information sets we avoid the issue of reaching some information sets with probability zero.

Consider case 1. There are deviations in information sets from time period t until $\tilde{t}$. Those deviations improve the payoff by ε. Now starting with information sets whose time duration is $\tilde{t}$, let us replace deviations by equilibrium behavioral strategies. By the one-stage deviation principle these replacements cannot reduce the continuation payoff at these information sets. Hence these replacements yield another example of case 1, except that the terminal period is now reduced in 1, and the payoff improvement is no less than ε. So in a finite number of steps we get $\tilde{t} = t$, which means that we found the one period deviations that improved the continuation payoff in at least $\varepsilon > 0$. We reached a contradiction, which proves the sufficiency statement of this lemma. Q.E.D.

Step 5: *Derivation of necessary and sufficient conditions.*

From Definition 1 and the result of the previous step it follows that we need to make sure that no deviation at any information set yields a larger continuation payoff. From step 3 we know that any information set may be "labeled" with the probability that the opponent plays Q_C instead of Q_P. We denote this probability as α. Any one period deviation at time t is characterized by the quantity produced, which is denoted by q.

There are three possible situations:

i) $q \leq Q_P$. In this situation, due to equilibrium strategy all these levels of q provide the same continuation play of the opponent. But $q = Q_P$, which is prescribed by equilibrium strategy, give the uniformly highest payoff for any level of α. So any deviation $q < Q_P$ cannot be a profitable deviation.

ii) $q \geq Q_P + 2r$. The opponent will detect the deviation with probability 1 and continue to behave by producing Q_C. Again analogous to the previous situation all q's but $\hat{Q}(\alpha) \geq Q_P + 2r$ cannot be optimal and $\hat{Q}(\alpha)$ is part of the equilibrium strategy.

iii) $Q_P < q < Q_P + 2r$. This action can be uniquely characterized by the probability of inducing a signal higher than $Q_P + r$. This probability is $\beta = \alpha(q)$, where the function $\alpha(q)$ is given by (8). The continuation payoff is $\Pi_{\alpha\beta}(\alpha, \beta, \delta, r)$. Hence, if the continuum set of inequalities (17) hold any deviation of situation (iii) in not profitable.

A2. Proof of Result 4

We develop the proof in several steps.

Step 1: *Proof that the system holds when $r = 0$.*

Substituting formulae (15) and (16) into (17) when $r = 0$ we get

$$\max\{\Pi_P(\alpha,\delta),\Pi_F(\alpha,\delta)\} \geq (1-\delta)\pi(Q_P,\alpha) + \delta\left[(1-\alpha)\Pi^*(\beta,\delta) + \alpha\pi_C\right] \tag{25}$$

Now the RHS[27] becomes $\Pi_P(\alpha,\delta)$ when $\beta = 0$. The proof of this step follows from the fact that $\Pi^*(\beta,\delta)$ is a nonincreasing function of β.

Step 2: *Modification of the system of inequalities while r is changing.*

When r is nonzero the initial system of inequalities (17) can be modified in the following form

$$\Delta\Pi(\alpha,\beta,\delta) \geq (1-\delta)\Delta\pi(\alpha,\beta,r)\ \forall\alpha,\beta \in [0,1], \tag{26}$$

where the function $\Delta\pi(\alpha,\beta,r) = \pi(Q_P + 2\beta r,\alpha) - \pi(Q_e,\alpha)$ or

$$\Delta\pi(\alpha,\beta,r) = \frac{\beta(3-\alpha)}{6}r - 4\beta^2 r^2 \tag{27}$$

and $\Delta\Pi(\alpha,\beta,\delta) = \Pi^0_E(\alpha,\delta) - \Pi^0_D(\alpha,\beta,\delta)$. The functions $\Pi^0_E(\alpha,\delta)$ and $\Pi^0_D(\alpha,\beta,\delta)$ are the LHS and RHS of inequality (25) correspondingly.

The properties of the new function $\Delta\Pi(\alpha,\beta,\delta)$ are studied in the next steps. As for the function $\Delta\pi(\alpha,\beta,r)$, it is continuous and increasing in r on the interval $[0,1/24]$ with $\Delta\pi(\alpha,\beta,0) = 0$. Given that $\Delta\Pi(\alpha,\beta,\delta)$ is nonnegative (see the previous step) and the properties of $\Delta\pi(\alpha,\beta,r)$, it follows immediately that for a given δ the set of r's for which the set of inequalities holds has the form $[0,\bar{r}(\delta)]$.

In order to find $\bar{r}(\delta)$ we gradually increase r until for some pairs (α,β) the inequality (26) binds for some $r < 1/24$, where $1/24$ comes from restriction (6). Such pairs we call "critical" pairs. In the next step we find the restriction for critical values of α.

Step 3: *For any critical pair $\alpha = \alpha^*(\delta)$.*

The functions $\Delta\Pi(\alpha,\beta,\delta)$ and $\Delta\pi(\alpha,\beta,r)$ are continuous in their arguments. Also it can be seen that the function $\Delta\pi(\alpha,\beta,r)$ is linear in α. In turn the function $\Delta\Pi(\alpha,\beta,\delta)$ is piecewise linear in α with the kink at $\alpha = \alpha^*(\delta)$. Hence any critical value of α is critical

[27] Throughout we will refer to the right hand side and the left hand side of any relationship as RHS and LHS, respectively.

together with one of three possible values $\{0, \alpha^*(\delta), 1\}$. The rest of this step shows that 0 and 1 cannot be critical values without $\alpha^*(\delta)$ being a critical value as well.

Case $\alpha = 1$: From the system of inequalities (25) and formula (27) it follows that

$$\begin{aligned}\Delta\Pi(\alpha &= 1, \beta, \delta) = (1-\delta)\frac{1}{144} \text{ and} \\ \Delta\pi(\alpha &= 1, \beta, r) = \frac{\beta}{3}r - 4\beta^2 r^2.\end{aligned}$$

The highest possible level of $\Delta\pi(1, \beta, r)$ is when β and r reach their highest possible values in which case $\Delta\pi(\alpha = 1, \beta = 1, r = 1/24) = 1/144$. Hence the inequality (26) holds for all values of β and r, so it is not possible to have critical pairs (α, β) with $\alpha = 1$.

Case $\alpha = 0$: In this case it can be seen that for $\alpha \leq \alpha^*(\delta)$ the function $\Delta\Pi(\alpha, \beta, \delta)$ has the form $\Delta\Pi(\alpha, \beta, \delta) = (1-\alpha)\varphi(\beta, \delta)$, where the function $\varphi(\beta, \delta)$ is nonnegative. Now suppose that at $\alpha = 0$ and for some $\hat{\beta}$ and $\hat{r}$ the inequality (26) holds with equality or $\varphi(\hat{\beta}, \delta) = (1-\delta)\Delta\pi(0, \hat{\beta}, \hat{r})$. Then, because $\Delta\pi(1, \hat{\beta}, \hat{r}) \geq 0$ and $\Delta\pi(\alpha, \hat{\beta}, \hat{r})$ is linear in α it follows that $\Delta\Pi(\alpha^*(\delta), \hat{\beta}, \delta) \equiv (1-\alpha^*(\delta))\varphi(\hat{\beta}, \delta) \leq (1-\delta)\Delta\pi(\alpha^*(\delta), \hat{\beta}, \hat{r})$. Hence if the system of inequalities (26) binds at $\alpha = 0$ and some r, then it binds at $\alpha = \alpha^*(\delta)$ and the same or lower level of r. So $\alpha = 0$ cannot be a critical value without $\alpha^*(\delta)$ being a critical value as well.

The original system of inequalities is parameterized by two parameters α and β. This step fixes the first parameter α, what helps to decrease the dimensionality of the system and moves us to the construction of the boundary curve $\bar{r}(\delta)$.

Step 4: *Derivation of two boundary curves $r_1(\delta)$ and $r_1(\delta)$, their values are higher than* 1/48.

From the previous step we found that in order to find the highest possible value of r at which the system of inequalities (26) holds, we can restrict our attention to values of α equal to $\alpha^*(\delta)$. Given that the function $\Delta\Pi(\alpha^*(\delta), \beta, \delta)$ can be represented as

$$\Delta\Pi(\alpha^*(\delta), \beta, \delta) = (1-\alpha^*(\delta))\delta \min\{\Delta\Pi_1(\beta, \delta), \Delta\Pi_2(\beta, \delta)\},$$

where

$$\Delta\Pi_1(\beta,\delta) = \frac{3-\delta}{144}\beta \text{ and}$$
$$\Delta\Pi_2(\beta,\delta) = \frac{1}{8} - \left[(1-\delta)\left(\frac{9-\beta}{24}\right)^2 - \delta\frac{1}{9}\right].$$

The function $\Delta\Pi(\alpha^*(\delta),\beta,\delta)$ is the minimum of two well behaved functions. One way to find the largest r, for which the function $\Delta\Pi(\alpha^*(\delta),\beta,\delta)$ still dominates $\Delta\pi(\alpha^*(\delta),\beta,r)$, is the following. First, we find the largest r for the function $\alpha^*(\delta)\delta\Delta\Pi_1(\beta,\delta)$, which we refer to as $r_1(\delta)$. Then we do the same for $\alpha^*(\delta)\Delta\Pi_2(\beta,\delta)$ and call the associated value of r as $r_2(\delta)$. Clearly the function of interest $\bar{r}(\delta)$ is the minimum of the functions $r_1(\delta)$ and $r_2(\delta)$.

Before studying each case separately let us prove the following lemma which is very handy throughout this step.

Lemma 3 $(1-\alpha^*(\delta))\delta\frac{1}{72} > (1-\delta)\frac{2-\alpha^*(\delta)}{144}$, *where the function* $\alpha^*(\delta)$ *is given by (13).*

PROOF: Below we present a series of simultaneous transformations applied to both sides of the inequality. They preserve the above inequality and facilitate the task of checking whether it holds. The sequence of inequality preserving operations on both sides is: multiply by 144, subtract $(1-\delta)(1-\alpha^*(\delta))$, multiply by $(1-\delta)$, plug in formula (13), add $\left\{4(3\delta-1)\sqrt{2\delta^2-\delta}-(1-\delta)^2\right\}$, take the square and finally subtract $-16(3\delta-1)^2(2\delta^2-\delta)$. These transformations brings us to the inequality $(1-\delta)^4 > 0$, which is obviously holds. Q.E.D.

Now let us derive the two curves in a sequence.

Curve $r_1(\delta)$: The function $\Delta\Pi_1(\beta,\delta)$ is linearly increasing in β and $\Delta\Pi_1(0,\delta) = 0$. In turn the function $\Delta\pi(\alpha^*(\delta),\beta,r)$ is quadratic in β, has a nonpositive second derivative and $\Delta\pi(\alpha^*(\delta),0,r) = 0$. Moreover, the function $\Delta\pi(\alpha^*(\delta),\beta,r)$ is increasing in r. From all of the above it follows that we can increase r until the slopes (derivatives in β) of the two functions $\alpha^*(\delta)\delta\Delta\Pi_1(\beta,\delta)$ and $(1-\delta)\Delta\pi(\alpha^*(\delta),\beta,r)$ at $\beta = 0$ are get equal. This gives us the necessary condition for the function $r_1(\delta)$

$$(1-\delta)\frac{3-\alpha^*(\delta)}{6}r_1(\delta) = (1-\alpha^*(\delta))\delta\frac{3-\delta}{144}, \tag{28}$$

where the LHS is the derivative of $(1-\delta)\Delta\pi$ and the RHS is the derivative of $(1-\alpha^*(\delta))\delta\Delta\Pi_1$ w.r.t. β at $\beta = 0$.

Now by substituting $r_1(\delta) = 1/48$ into the LHS we have $(1-\delta)\frac{1+(1-\alpha^*(\delta))/2}{144}$, which is strictly less than $(1-\alpha^*(\delta))\delta\frac{1}{72}$ by lemma 3, while the RHS is strictly larger than $(1-\alpha^*(\delta))\delta\frac{1}{72}$. This gives

Lemma 4 $r_1(\delta) > \frac{1}{48}$ *for all* $\delta \in [\underline{\delta}, 1)$.

By solving equation (28) for $r_1(\delta)$ and by substituting formula (13) for $\alpha^*(\delta)$ we get

$$r_1(\delta) = \frac{1}{48}\frac{\delta(3-\delta)\left(3\delta - 1 - 2\sqrt{2\delta^2-\delta}\right)}{(1-\delta)(\delta - \sqrt{2\delta^2-\delta})}. \tag{29}$$

The above function is continuous on $[\underline{\delta}, 1)$ except at the points where the denominator is zero and the expression under the square root is negative. It can be checked that those points are outside the region $[\underline{\delta}, 1)$.

The result of the next lemma helps to get the limiting value of the function $r_1(\delta)$ at $\delta = 1$.

Lemma 5 $\lim\limits_{\delta\to 1}\frac{\delta-\sqrt{2\delta^2-\delta}}{1-\delta} = \frac{1}{2}$ *and* $\lim\limits_{\delta\to 1}\frac{3\delta-1-2\sqrt{2\delta^2-\delta}}{(1-\delta)^2} = \frac{1}{4}$.

PROOF: The above limits result from a Taylor expansion of the numerator around the point $\delta = 1$. Q.E.D.

From the above lemma it follows that $\lim\limits_{\delta\to 1} r_1(\delta) = 1/48$ and it can be checked that $r_1(\underline{\delta}) = 21/544$.

The monotonicity properties are hard to investigate but from the graphical representation of the function on Figure 9 we see that it is a monotonically decreasing function of δ.

Curve $r_2(\delta)$: In this case both functions $\Delta\Pi_2(\beta, \delta)$ and $\Delta\pi(\alpha^*(\delta), \beta, r)$ are increasing and quadratic in β with a negative curvature. At $\beta = 0$ the function $\Delta\Pi_2(\beta, \delta)$ is nonnegative and at $\beta = 1$ it is equal to $1/72$. The highest possible level of the function $\Delta\pi(\alpha^*(\delta), \beta, r)$ is achieved at $\beta = 1$ and $r = 1/24$, in which case the function equals $\frac{2-\alpha^*(\delta)}{144}$. From lemma 3 it follows that the two functions $(1-\alpha^*(\delta))\delta\Delta\Pi_1(\beta,\delta)$ and $(1-\delta)\Delta\pi(\alpha^*(\delta),\beta,r)$ never meet at the point $\beta = 1$.

Before finding the closed form of $r_2(\delta)$, let us show that all values of $r_2(\delta)$ are strictly higher than 1/48.

Lemma 6 *$r_2(\delta) > 1/48$ for all $\delta \in [\underline{\delta}, 1)$.*

PROOF: By plugging in 1/48 for r and $\alpha^*(\delta)$ for α into formula (27) we get

$$\Delta\pi(\alpha^*(\delta),\beta,\frac{1}{48}) = \frac{(3-\alpha^*(\delta))}{288}\beta - \frac{1}{576}\beta^2.$$

The above function is concave in β. Let us consider a linear function of β, $L(\beta,\delta)$ which at $\beta = 0$ has the same value and derivative as the above function $(1-\delta)\Delta\pi(\alpha^*(\delta),\beta,\frac{1}{48})$. This function has the form

$$L(\beta,\delta) = (1-\delta)\left(\frac{3-\alpha^*(\delta)}{288}\right)\beta.$$

Next, the following two inequalities hold

$$\begin{aligned}
(1-\alpha^*(\delta))\delta\Delta\Pi_2(0,\delta) &= (1-\alpha^*(\delta))\delta\frac{17\delta-9}{576} \geq 0 = L(0,\delta) \text{ and} \\
(1-\alpha^*(\delta))\delta\Delta\Pi_2(1,\delta) &= \frac{(1-\alpha^*(\delta))\delta}{72} > (1-\delta)\left(\frac{3-\alpha^*(\delta)}{288}\right) = L(1,\delta),
\end{aligned}$$

where the first inequality follows from the fact that $\delta \geq 9/17$, and the second inequality follows from lemma 3.

Because the function $\Delta\Pi_2$ is concave in β, it follows that the function $(1-\alpha^*(\delta))\delta\Delta\Pi_2$ strictly dominates the linear function $L(\beta,\delta)$ except at the point $\beta = 0$ when $\delta = 9/17$. As a result when $r = 1/48$, the function $(1-\alpha^*(\delta))\delta\Delta\Pi_2$ strictly dominates the function $(1-\delta)\Delta\pi$, except at the point $(\beta,\delta) = (0, 9/17)$. Comparing the derivatives w.r.t. β of these two functions at the point $(\beta,\delta) = (0, 9/17)$, we get the strict inequality $9/544 > 1/204$, so the result of the lemma holds with strict inequality. Q.E.D.

We now return to determining the form of $r_2(\delta)$. The only possibility left is that the two functions considered above are tangent in β at the highest possible level of $r = r_2(\delta)$. The tangency condition gives us the necessary system of equations for β_2 (the β where two curves are tangent) and $r_2(\delta)$

$$\begin{aligned}(1-\alpha^*(\delta))\delta\left\{\frac{1}{8}-\left[(1-\delta)\left(\frac{9-\beta}{24}\right)^2-\delta\frac{1}{9}\right]\right\} &= (1-\delta)\left\{\frac{\beta_2(3-\alpha^*(\delta))}{6}r_2(\delta)-4{\beta_2}^2r_2(\delta)^2\right\}\\ \text{and } \frac{1}{288}(1-\alpha^*(\delta))\delta(9-\beta_2) &= (1-\delta)\left\{\frac{3-\alpha^*(\delta)}{6}r_2(\delta)-8\beta_2r_2(\delta)^2\right\},\end{aligned}$$

where the first equation states that the functions are equal and the second equation states that their derivatives coincide.

By solving for β_2 from the second equation which is linear in β_2 and by substituting it into the first one we get a polynomial of fourth order in $r_2(\delta)$, and then we can write down four possible solutions for $r_2(\delta)$. The first two solutions are $\pm\frac{\sqrt{(1-\alpha^*(\delta))\delta}}{48}$ which do not generate values greater than 1/48. For the other two solutions by applying lemma 5 we get that one of the roots has the property $\lim_{\delta\to 1} r_2(\delta)=1/408<1/48$, while the other one has the property $\lim_{\delta\to 1} r_2(\delta)=1/24>1/48$. So the last one is the solution we are looking for and it is given by formula[28]

$$r_2(\delta)=\frac{1}{204}\frac{9\delta-9\sqrt{2\delta^2-\delta}+2\sqrt{2}\sqrt{\delta(17\delta-9)\left(3\delta-1-2\sqrt{2\delta^2-\delta}\right)}}{1-\delta}.$$

It can be checked that this function is continuous in δ on $[\underline{\delta},1)$ and $\lim_{\delta\to 1} r_2(\delta)=1/24$ and $r_2(\underline{\delta})=9/272$.

Monotonicity properties is hard to investigate but from a graphical representation of the function (see Figure 9) we see that the function $r_2(\delta)$ is a monotonically decreasing function of δ.

A3. Proof of the Folk Theorem

Proof that there exists sufficiently small $\epsilon(\bar{u}^1,\bar{u}^2)>0$ such that for any $\delta\geq 1-\epsilon(\bar{u}^1,\bar{u}^2)$ and $r\leq\epsilon(\bar{u}^1,\bar{u}^2)$ the system of inequalities in (23) holds.

The system of inequalities (23) may be rewritten as

$$\Pi^*(\alpha,\delta,\bar{Q}^i,\bar{Q}^j)\geq(1-\delta)\pi(\bar{Q}^i+2\beta r,\alpha,\bar{Q}^j)+\delta\left[(1-\alpha)\Pi^*(\beta,\delta,\bar{Q}^i,\bar{Q}^j)+\alpha\pi_C\right].$$

[28] To get the formula for the other root in which $\lim_{\delta\to 1} r_2(\delta)=1/408$ ths plus sign before $2\sqrt{2}$ should be replaced by the minus sign.

Note that the last term of the RHS is increasing in β and the RHS equal to $\Pi_P(\alpha, \delta, \bar{Q}^i, \bar{Q}^j)$ when $r = 0$ and $\beta = 0$. The first term of RHS is increasing in β and the second term is decreasing in β. The rate of increase of the first term declines till 0 with a decrease of r and so we can always find sufficiently small r which makes the above inequality true. The proof is complete if we take the bound for δ from the Perfect Information case.

10 References

Abreu D., "Repeated games with discounting: A general theory and an application to oligopoly", mimeo, Princeton University, 1982

Abreu D., Pearce D. and Stacchetti E., "Optimal cartel equilibria with imperfect monitoring", Journal of Economic Theory, 39: 251-269, 1986

Aoyagi M., "A collusion mechanism in dynamic Bertrand oligopoly under private observation of correlated demand shocks", Mimeo, University of Pittsburgh, 1997

Bhaskar V., "Sequential equilibria in the repeated prisoner's dilemma with private monitoring", Mimeo, University of Essex, 1999

Brander J.A. and Zhang A., "Dynamic oligopoly behavior in the airline industry", International Journal of Industrial Organization, 11: 407-435, 1993

Chatterjee K. and Samuelson L., "Perfect equilibria in simultaneous-offers bargaining", Journal of Game theory, 19: 237-267, 1990

Compte O., "Communication in repeated games with imperfect private monitoring", Econometrica, 66: 597-626, 1998

Compte O., "On failing to cooperate when monitoring is private", Journal of Economic Theory, 102(1): 151-188, 2002

Ellison G., "Theories of cartel stability and the Joint Executive Committee", RAND Journal of Economics, 25: 37-57, 1994

Ely J.C. and Valimaki J., "A robust folk theorem for the prisoner's dilemma", Journal of Economic Theory, 102: 84-105, 2002

Friedman J., "A non-cooperative equilibrium for supergames", Review of Economic Studies 38: 1-12, 1971

Fudenberg D. and Levine D., "An approximate folk theorem with imperfect private information", Journal of Economic Theory, 54: 26-47, 1991

Fudenberg, D. and Tirole J., "Game theory", MIT Press, 1996

Green E.J. and Porter R.H., "Noncooperative collusion under imperfect price information", Econometrica, 52: 87-100, 1984

Kandori M. and Matsushima H., "Private observation, communication and collusion", Econometrica, 66: 627-652, 1998

Kreps D. and Wilson R., "Sequential equilibrium", Econometrica, 50: 863-894, 1982

Levenstein M.C., "Price wars and the stability of collusion: A study of the pre-World War I bromine industry", The Journal of Industrial Economics, 45: 117-137, 1997

Lehrer E., "On the equilibrium payoffs set of two player repeated games with imperfect monitoring", International Journal of Game Theory, 20: 211-226, 1992

Malaith G. and Morris S., "Repeated games with almost-public monitoring", Journal of Economic Theory, 102(1):189-228, 2002

Matsushima H., "On the theory of repeated games with private information. Part I: Anti-folk theorem without communication", Economic Letters, 35: 253-256, 1991

Piccione M., "The repeated prisoner's dilemma with imperfect private monitoring", Journal of Economic Theory, 102(1):70-83, 2002

Sekiguchi T., "Efficiency in repeated prisoner's dilemma with private monitoring", Journal of Economic Theory 76: 345-361, 1997

Simon L. K. and Stinchcombe M.B.,"Equilibrium refinement for infinite normal-form games", Econometrica, 63: 1421-1443, 1995

Stigler G.J., "A theory of oligopoly", Journal of Political Economy, 72: 44-61, 1964

Tedeschi P., "Cartels with homogeneous and differentiated products and imperfect monitoring", International Economic Review, 35: 635-656, 1994