Collusion Enforcement with Private Information and Private Monitoring*

Jimmy H. Chan† Wenzhang Zhang‡

December 2014

Abstract

This paper shows that a cartel that observes neither costs, prices, nor sales may still enforce a collusive agreement by tying each firm's continuation profit to the truncated current profits of the other firms. The mechanism applies to both price and quantity competition, and the main features are broadly consistent with common cartel practice identified by Harrington and Skrzypacz (2011).

JEL classification: C73; D43; D82; L13; L40

Keywords: Collusion; Oligopoly; Repeated game; Private information; Private monitoring

*An earlier version of this paper has been circulated under the title "Price Wars, Side-Payments, and Collusion Enforcement."

†School of Economics, Fudan University, Shanghai 200433, China. E-mail: jimmy.hing.chan@gmail.com.

‡Corresponding author. School of Economics, Shanghai University of Finance and Economics, Shanghai 200433, China. Tel: (86 21) 6590-2948. E-mail: wenzhang.zhang@gmail.com.

1 Introduction

Despite an extensive literature on repeated games, there has been little work that seeks to provide a theoretical justification of observed cartel practice. In an important exception, Harrington and Skrzypacz (2011) note that many cartels used transfer schemes to enforce collusion. Under these schemes, colluding firms agreed to a set of sales quotas, and firms that reported sales above quota compensated firms that reported sales below through inter-firm purchases. In one case, Haarmann & Reimer purchased 7000 tons of citric acid from Archer Daniel Midlands. Inspired by Harrington and Skrzypacz (2011), we propose a new collusion enforcement scheme in a repeated game with both private monitoring and private information. We show that despite observing neither costs, prices, nor sales, a cartel can still enforce a collusive agreement that maximizes joint profit. When the firms are patient and the demand shocks small, the equilibrium cartel profit could be close to the monopoly level.

We call the static version of our scheme an output-target mechanism. Under this mechanism firms first agree to a set of output targets. At the end of each period the cartel can calculate the "reported" profit of each firm based on the reported output and cost of each firm. The cost reports and profit shortfalls—the difference between the profit targets (i.e., the firm's profit if it makes the output target) and the reported profits—jointly determine the side-payments between firms and the probability of switching to a non-collusive continuation path (i.e., having a price war) in the next period. A firm tends to pay less when it reports lower sales, but the gain is offset by a higher probability of a price war. Taking into account its payments to other firms and the loss in case of a price war, each firm's continuation profit in equilibrium is linear in the reported output shortfalls of the other firms.

One can view the mechanism as a *truncated* Clarke-Groves mechanism (Vickrey, 1961, Clarke, 1971 and Groves, 1973). If the output targets were set higher than the maximum outputs, then the continuation profit of each firm (including side-payments) would be equal to the reported profits of the other firms (minus a constant), and our scheme would resemble a classic Clarke-Groves mechanism. But since our cartel problem involves hidden actions, we cannot balance the budget by transferring the penalty of one firm to another firm. Instead, all penalties are destroyed through a price war. If the output targets were set above the maximum outputs, the amount to be destroyed could be very large. The first contribution of this paper is to show that we can reduce the efficiency loss by lowering the output targets. When the demand shocks are small, the efficiency loss would be small when the output targets are set near the expected output levels. However, when output targets are set below the maximum outputs, the transfer scheme would no longer fully capture the external effects of the firms' actions. A firm that has secretly cut price may escape punishment when a positive demand shock masks the effect of the price cut. A second contribution of the paper is to show that despite this "truncation" problem, the mechanism can still enforce an efficient collusive agreement under a fairly weak condition.

The main feature of our mechanism—that a firm is punished when the reported outputs of the other firms fall below certain targets—is consistent with the transfer schemes described by Harrington and Skrzypacz (2011). Since these schemes are common, it is important to understand how they work. Harrington and Skrzypacz (2011) show that the main features of these schemes are consistent with equilibrium behavior in a model of repeated Bertrand competition with private monitoring. Our model is more general in that it allows private information and applies to both Bertrand and Cournot competition. In Harrington and Skrzypacz (2011), inter-firm transfers serve as a linear output tax. By choosing the right tax rate, a cartel can discourage the firms from over-production. However, since a firm is required to pay more when it reports higher sales, it has an incentive to under-report. Harrington and Skrzypacz (2011) show that it is possible to induce truth-telling when the industry demand is completely inelastic. Under our mechanism, firms have no incentive to mis-report as the continuation profit of a firm does not depend on its own report.

The standard approach to enforce an efficient outcome in repeated games with private monitoring is to punish a player whenever he fails a statistical test (Kandori and Matsushima, 1998, Compte, 1998, Fudenberg and Levine, 2007, Obara, 2009, and Zheng, 2008). Our approach is to make each firm internalize the external effects of its action. The main advantage of this approach is that it is very robust. The existing literature on collusion enforcement can be divided into two strands. One strand assumes that prices are publicly observable and focuses on the issue of private cost information (Athey and Bagwell, 2001, Athey and Bagwell, 2008).[1] Another strand assumes costs are public and focuses on the issue of secret price cutting (Green and Porter, 1984, Aoyagi, 2002, Harrington and Skrzypacz, 2007, and Harrington and Skrzypacz, 2011). There is, however, no economic reason why a cartel must be able to observe either the costs, prices, or sales of the firms. Because a firm can lie and cheat at the same time, a cartel in general cannot use one scheme to elicit the private information and apply a separate scheme to

[1]Other related works include Cramton and Palfrey (1990), Athey, Bagwell, and Sanchirico (2004), Athey and Miller (2007), Miller (2012), Aoyagi (2003), Hörner and Jamison (2007), Escobar and Toikka (2013), Skrzypacz and Hopenhayn (2004), and Blume and Heidhues (2006).

induce the collusive prices. Our approach provides a natural solution to the twin problems of inducing the firms to reveal their costs and set the right prices.

Our paper is also related to the recent works that extend the Clarke-Groves approach to a dynamic environment. Bergemann and Välimäki (2010) introduce a dynamic pivot mechanism that generalizes the pivot mechanism to a dynamic setting. Athey and Segal (2013) introduce a dynamic generalization of the Arrow-d'Aspremont-Gerard-Varet mechanism (d'Aspremont and Gerard-Varet, 1979 and Arrow, 1979). Both Bergemann and Välimäki (2010) and Athey and Segal (2013) involve only private information, but in any period the distribution of an agent's types may depend on his previous types and past collective decisions.[2] In our model the cost distributions are stationary, but each firm can choose a hidden action that directly affects the profits of the other firms. Mezzetti (2004) considers a two-period mechanism-design problem. He shows that even when payoffs are interdependent, it is possible to enforce the efficient outcome by setting the transfer of each player equal to the reported realized payoffs of the other players. Our output-target mechanism is similar in that it uses the reported realized profits of the firms to enforce the efficient action profile. The difference is that in our model it is the action, rather than the type of a firm, that directly affects the profits of the other firms. Lastly, Hörner, Takahashi, and Vieille (2013) consider a general dynamic Bayesian game that incorporates both hidden actions and Markovian types. In the case of private independent types, they solve the private-information and imperfect-monitoring problems simultaneously by combining the Clarke-Groves approach of Athey and Segal (2013) and Mezzetti (2004) with s-

[2] In an early draft of Athey and Segal (2013), Athey and Segal (2007), each agent also chooses a private action but the action can affect only the distribution of his future types and not the types of the other agents.

tandard techniques in repeated games (Kandori and Matsushima, 1998). We show that under a fairly weak condition a truncated Clarke-Groves mechanism is still incentive compatible.

The rest of the paper is organized as follows. The next section introduces the model. Section 3 describes the standard Clarke-Groves approach and explains why it may lead to a large efficiency loss in a model with hidden actions. Section 4 introduces the output-target mechanism in a static setting and shows that it implements the efficient collusive outcome when the stochastic demand functions satisfy a monotone condition. Section 5 uses the mechanism to construct a perfect public equilibrium in a repeated game. Section 6 concludes.

2 Model

2.1 Demand

We consider an infinitely repeated oligopoly game. Let $\mathcal{N}$ denote a set of n firms, each with constant marginal cost. In each period $t = 1, \ldots, \infty$, each firm i chooses an "action" a_i from a compact interval $A_i \subset \Re_+$. Let $a \equiv (a_1, \ldots, a_n)$ denote an action profile and $A \equiv \Pi_{i=1}^n A_i$ the set of action profiles. The "output" for each firm i is denoted by an output function

$$y_i = y_i\left(a, \varepsilon_i\right) \in \Re_+,$$

where y_i is the output function of firm i, and ε_i a random shock that is distributed according to a smooth distribution function F_i on a support $\Omega_i \equiv [\underline{\varepsilon}_i, \overline{\varepsilon}_i]$. Write ε for $(\varepsilon_1, \ldots, \varepsilon_n)$ and ε_{-i} for ε with ε_i removed. Let F and F_{-i} denote the distributions of ε and ε_{-i}, respectively. We do not impose any restriction on the correlation between the demand shocks; they may be independent or correlated.

Our model encompasses both price and quantity competition. In the former case, a_i represents firm i's price and y_i its sales, and in the latter vice versa.[3]

For any $i, j \in \mathcal{N}$, $i \neq j$, we assume that y_i is continuous and increasing in ε_i and either increasing in a_j for all ε_i or decreasing in a_j for all ε_i. The latter means that the products of firms i and j are either always substitutes or always complements. For each firm i, y_i is integrable in ε_i. Let

$$\overline{y}_i\left(a\right) \equiv \int_{\varepsilon_i} y_i\left(a, \varepsilon_i\right) dF_i\left(\varepsilon_i\right)$$

denote firm i's expected output. We assume that for any $j \in \mathcal{N}$, $\partial y_i(a, \varepsilon_i)/\partial a_j$ exists for all a and almost all ε_i. Moreover, we assume $\overline{y}_i\left(a\right)$ is differentiable in a_j.[4] In particular,

$$\frac{\partial \overline{y}_i\left(a\right)}{\partial a_j} = \int_{\varepsilon_i} \frac{\partial y_i\left(a, \varepsilon_i\right)}{\partial a_j} dF_i\left(\varepsilon_i\right). \tag{1}$$

To rule out the pathological case where the marginal effect is concentrated on an arbitrarily small subset of ε_i, we assume there exists some finite $\eta > 0$ such that for any $i, j \in \mathcal{N}$, $i \neq j$, and any $a \in A$,

$$\left|\frac{\partial y_i\left(a, \varepsilon_i\right)}{\partial a_j}\right| \leq \eta \left|\frac{\partial \overline{y}_i\left(a\right)}{\partial a_j}\right| \quad \text{for almost all } \varepsilon_i \in \Omega_i. \tag{2}$$

2.2 Supply

In each period each firm i's marginal cost, c_i, is subject to a firm-specific shock that is independent of the demand shock and uncorrelated across firms. For each

[3] The assumption that y_i is positive means that a firm can dispose of its product freely in the case of quantity competition.

[4] A sufficient condition for $\overline{y}_i$ to be differentiable and (1) to hold is that there exists an integrable function ψ such that $|\partial y_i(a, \varepsilon_i)/\partial a_j| \leq \psi(\varepsilon_i)$ for all a. See, e.g., Theorem 20.4 of Aliprantis and Burkinshaw (1990).

firm i, the marginal cost c_i is distributed on a compact interval $C_i \subset \Re_+$ according to a distribution G_i. Firm i's profit is given by the function

$$\pi_i (a, c_i, \varepsilon_i) \equiv \Phi_1 (c_i, a_i) y_i (a, \varepsilon_i) + \Phi_2 (c_i, a_i) .$$

We consider two functional forms of Φ_1 and Φ_2. In the Bertrand case, $\Phi_1 (c_i, a_i) \equiv (a_i - c_i)$ and $\Phi_2 (c_i, a_i) \equiv 0$. In the Cournot case, $\Phi_1 (c_i, a_i) \equiv a_i$ and $\Phi_2 (c_i, a_i) \equiv -c_i a_i$.

Write C for $\Pi_{i=1}^n C_i$. Let $c = (c_1, \ldots, c_n) \in C$ denote a cost profile and c_{-i} a cost profile minus c_i, and let G and G_{-i} denote the distributions of c and c_{-i}, respectively. A cost-action profile $\alpha : C \to A$ is a function that maps each cost profile c into an action profile $\alpha (c) = (\alpha_1 (c), \ldots, \alpha_n (c)) \in A$. Let α^* denote a cost-action profile such that, for each $c \in C$, $\alpha^* (c)$ maximizes

$$\sum_{i=1}^{n} \int_{\varepsilon_i} \pi_i (a, c_i, \varepsilon_i) \, dF_i (\varepsilon_i) .$$

In the Bertrand case, we assume that $\Phi_1 (c_i, \alpha_i^*(c)) = (\alpha_i^*(c) - c_i) \geq 0$ for all c. For any cost-action profile α, let

$$\overline{\pi}_i (\alpha) \equiv \int_c \int_{\varepsilon_i} \pi_i (\alpha (c), c_i, \varepsilon_i) \, dF_i (\varepsilon_i) \, dG (c)$$

and

$$\overline{\Pi} (\alpha) \equiv \sum_{i=1}^{n} \overline{\pi}_i (\alpha) .$$

Given our assumptions, the incomplete-information oligopoly game where each firm i chooses a_i independently after observing c_i has a Bayesian-Nash equilibrium in distributional strategies.[5] To save notation, we assume that there is a pure-strategy equilibrium and denote it by the cost-action profile α^{NE}. Since the

[5] See Fudenberg and Tirole (1991), Theorem 6.3.

equilibrium action of firm i depends only on c_i, $\alpha_i^{NE}(c_i, c_{-i}) = \alpha_i^{NE}\left(c_i, c'_{-i}\right)$ for all $c_i \in C_i$ and $c_{-i}, c'_{-i} \in C_{-i}$. By definition, for any firm i and any $c \in C$,

$$\alpha_i^{NE}(c) \in \arg\max_{a_i \in A_i} \int_{c_{-i}} \int_{\varepsilon_i} \pi_i\left(a_i, \alpha_{-i}^{NE}(c), c_i, \varepsilon_i\right) dF_i(\varepsilon_i)\, dG_{-i}(c_{-i}).$$

Since enforcing collusion would be trivial if α^{NE} maximizes the expected cartel profit, we assume $\overline{\Pi}\left(\alpha^{NE}\right) < \overline{\Pi}(\alpha^*)$.

2.3 Timing and information

For any variable x we use x^t to denote the value of x in period t. Both the demand shocks ε^t and cost shocks c^t are uncorrelated over time. In each period t each firm i chooses a_i^t after observing c_i^t. Then each firm i observes its own output y_i^t after ε_i^t, which is not directly observable, is realized. We assume that c_i^t, a_i^t, and y_i^t are all private information. In every period the firms meet twice to exchange information, first after the marginal costs are drawn but before actions are chosen, and again after outputs are realized. Following Harrington and Skrzypacz (2011), we allow firms to exchange side-payments at the end of each period.

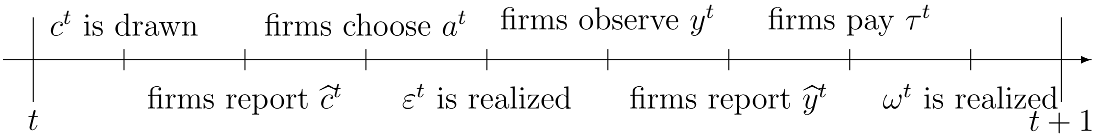


Figure 1: Timeline.

Specifically, the timeline of each period t is as follows: (1) each firm i first privately draws a cost parameter c_i^t from C_i; (2) it then sends a cost report $\widehat{c}_i^t \in C_i$ to the other firms; (3) after receiving the cost reports from the other firms, firm i chooses a_i^t; (4) the demand shock ε_i^t is realized; (5) firm i observes its output y_i^t and updates its belief about ε_i^t; (6) it sends an output report $\widehat{y}_i^t$ to the other firms; (7) after receiving the output reports from the other firms, it makes a side-payment $\tau_{ij}^t \geq 0$ to each firm j; these payments are publicly observed by all firms; (8) finally, it observes the outcome ω^t of a public randomization device that is uniformly distributed between 0 and 1.[6] See Figure 1.

In period t each firm must choose a cost report, a private action, an output report and a vector of side-payments. Let $\widehat{c}^t = (\widehat{c}_1^t, \ldots, \widehat{c}_n^t)$ denote a profile of cost reports, and $\widehat{y}^t = (\widehat{y}_1^t, \ldots, \widehat{y}_n^t)$ a profile of output reports. A cost-reporting strategy for firm i is a function $\rho_i : C_i \to C_i$ that maps firm i's marginal cost c_i^t into a cost report $\widehat{c}_i^t \in C_i$. An action strategy is a function $\gamma_i : C \to A_i$ that determines firm i's period-t action, a_i^t, on the basis of c_i^t, firm i's own marginal cost and $\widehat{c}_{-i}^t$, the cost-report profile of all firms $j \neq i$. An output-reporting strategy is a function $r_i : C \times \Re_+ \to \Re_+$ that determines firm i's output report, $\widehat{y}_i^t$, on the basis of c_i^t, $\widehat{c}_{-i}^t$, and y_i^t. Finally, a transfer strategy is a function $b_i : C \times \Re_+^n \to \Re_+^n$ that maps c_i^t, $\widehat{c}_{-i}^t$, y_i^t, and $\widehat{y}_{-i}^t$ into a vector of side-payments $\tau_i^t = (\tau_{i1}^t, \ldots, \tau_{in}^t)$.[7] A reduced-normal-form stage-game strategy for firm i is a quadruple $(\rho_i, \gamma_i, r_i, b_i)$.[8]

A history of the game is an infinite sequence

$$\left\{\varepsilon^t, c^t, a^t, y^t, \widehat{c}^t, \widehat{y}^t, \tau^t, \omega^t\right\}_{t=1}^{\infty}.$$

At the beginning of period t (before observing the marginal cost of that period) each firm has observed a public history h_{pub}^t that includes the cost reports, output

[6]The existence of a public randomization device, a standard assumption in the literature, allows firms to coordinate their continuation strategies in future periods.

[7]We assume firm i also pays itself a side-payment to simplify notations. Its value obviously has no significance.

[8]Allowing γ_i to depend also on $\widehat{c}_i$ will lead to more tedious notations without changing the result. Similarly, it is not helpful to allow r_i and b_i to depend on $(a_i, \widehat{c}_i)$ and $(a_i, \widehat{c}_i, \widehat{y}_i)$, respectively.

reports, side-payments, outcomes of the randomization device in the first $(t-1)$ periods. In addition, firm i has observed a private history h_i^t of its own costs, actions, and outputs. Firm i's information at the beginning of period t includes both h_{pub}^t and h_i^t. A repeated-game strategy for firm i, denoted by σ_i, is a function that maps firm i's information in each period t into a stage-game strategy. The firms discount future profits by a factor $\delta < 1$. The average discounted profit for firm i under strategy profile $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ is

$$v_i(\sigma) = (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} E\left[\pi_i \left(a^t, c_i^t, \varepsilon_i^t\right) + \sum_{j \neq i} \left(\tau_{ji}^t - \tau_{ij}^t\right) \middle| \sigma \right],$$

where the expectation is taken over the distribution of histories induced by σ, F and G. The solution concept we use is perfect public equilibrium.[9] A repeated-game strategy of a firm is a public strategy if its stage-game strategy in any period depends only on the public history up to that period. A strategy profile is a perfect public equilibrium if the strategy of each firm is public and if the continuation strategy profile after any public history is a Nash equilibrium in the continuation game.[10]

3 The Clarke-Groves Approach

Although a standard Clarke-Groves mechanism involves only private information, it can be easily extended to allow hidden actions. Consider a one-shot collusion game with transfers defined by steps 1-5 of the stage game introduced in the

[9] It is common to focus on perfect public equilibrium in the literature. Notice, however, that this is not without loss of generality. For example, Kandori and Obara (2006) show that in a prisoners' dilemma game using private strategies can improve efficiency.

[10] See Definition 5.3 of Fudenberg and Tirole (1991) for a formal description of a perfect public equilibrium.

last section. But instead of making side-payments each firm i directly receives a transfer w_i that depends on $\widehat{c}$ and $\widehat{y}$.[11] The total payoff of firm i in this one-shot game is therefore

$$\pi_i(a, c_i, \varepsilon_i) + w_i(\widehat{c}, \widehat{y}).$$

The transfer function w_i is defined for all $(\widehat{c}, \widehat{y}) \in C \times \Re^n_+$. Since there is no external source of funding in the original repeated game, the total transfer must be negative. Hence,

$$\sum_{i=1}^{n} w_i(\widehat{c}, \widehat{y}) \leq 0 \quad \text{for all } (\widehat{c}, \widehat{y}) \in C \times \Re^n_+. \tag{3}$$

Our task is to design $w = (w_1, \ldots, w_n)$ to satisfy (3) such that it is a Nash equilibrium for each firm i to choose $\alpha_i^*(\widehat{c})$ and report c_i and y_i truthfully. As a cartel can destroy surplus through a price war, the total transfer can be strictly negative.

In a standard Clarke-Groves mechanism, each agent receives a transfer equal to the total payoffs of the other agents minus a term that is not manipulable by the agent. Let

$$\widetilde{\pi}_i(\widehat{y}_i, \widehat{c}) \equiv \Phi_1(\widehat{c}_i, \alpha_i^*(\widehat{c}))\,\widehat{y}_i + \Phi_2(\widehat{c}_i, \alpha_i^*(\widehat{c}))$$

denote the *reported* profit of firm i (the profit it will make if it chooses $\alpha_i^*(\widehat{c})$ and its cost and output reports are truthful). One way to implement the Clarke-Groves approach is to set

$$w_i^{CG}(\widehat{c}, \widehat{y}) = \sum_{j \neq i} \min\left(\widetilde{\pi}_j(\widehat{y}_j, \widehat{c}), \pi_j(\alpha^*(\widehat{c}), \widehat{c}_j, \bar{\varepsilon}_i)\right) + L_i(\widehat{c}), \tag{4}$$

where L_i is a function that satisfies the condition

$$\int_{c_{-i}} L_i(\widehat{c}_i, c_{-i})\, dG_{-i}(c_{-i}) = \int_{c_{-i}} L_i(\widehat{c}_i', c_{-i})\, dG_{-i}(c_{-i})$$

[11] As is common in the repeated-game literature, we present our result in a static model with transfers.

for all $\widehat{c}_i, \widehat{c}_i' \in C_i$. This condition ensures that firm i's cost report does not affect its own expected transfer when other firms are reporting truthfully. By (4), firm i's transfer is increasing in the profit of firm j up to $\pi_j(\alpha^*(\widehat{c}), \widehat{c}_j, \overline{\varepsilon}_j)$, the maximum equilibrium profit of firm j under the most favorable demand shock. The cap is needed to put an upper bound on the firms' transfer.[12]

A nice feature of the Clarke-Groves approach is that it solves the private monitoring and private information problem simultaneously. Since α^* maximizes the expected cartel profit, any cheating by a firm, whether by lying about cost or deviating from the equilibrium action, must hurt the other firms more than it helps itself. Furthermore, as w_i is independent of $\widehat{y}_i$, no firm could gain by lying about its output. It is therefore a Nash equilibrium for the firms to choose $\alpha^*(\widehat{c})$ and report truthfully. Note that the caps do not change the firms' incentives, because they do not bind in equilibrium and will never increase a firm's transfer if it deviates.

Unfortunately, the efficiency loss of a generalized Clarke-Groves mechanism is likely to be large. Unlike agents in a standard private-information environment, here each firm is choosing a private action that *directly* affects the payoffs of the other firms. Since a_i may affect $\widehat{y}_{-i}$ indirectly through y_{-i}, L_i cannot depend on $\widehat{y}_{-i}$. In general there exist no $(L_1, \ldots, L_n)$ such that

$$\sum_{i=1}^{n} w_i^{CG}(\widehat{c}, \widehat{y}) = 0 \quad \text{for all } (\widehat{c}, \widehat{y}) \in C \times \Re_+^n.$$

To conform with (3), we have to choose L_i, for each $i \in \mathcal{N}$, such that for any $c \in C$

$$\sum_{i=1}^{n} L_i(c) = -(n-1)\sum_{i=1}^{n} \pi_i(\alpha^*(c), c_i, \overline{\varepsilon}_i).$$

[12] An equivalent approach is to limit the report of y_j to the interval $[0, y_i(\alpha^*(\widehat{c}), \overline{\varepsilon}_i)]$.

As a result the firms as a group will be punished whenever the realized cartel profit is lower than the maximum cartel profit (i.e., the profit under the most favorable demand shock). The total expected transfer under $w^{CG} = \left(w_1^{CG}, \ldots, w_n^{CG}\right)$ is

$$W^{CG} = \sum_{i=1}^{n} \int_c \int_\varepsilon w_i^{CG} \left(c, y\left(\alpha^*\left(c\right), \varepsilon\right)\right) dF\left(\varepsilon\right) dG\left(c\right) \tag{5}$$

$$= (n-1)\left(\overline{\Pi}\left(\alpha^*\right) - \int_c \sum_{i=1}^{n} \pi_i\left(\alpha^*\left(c\right), c_i, \overline{\varepsilon}_i\right) dG\left(c\right)\right). \tag{6}$$

The efficiency loss is $\left|W^{CG}\right|$. If the maximum cartel profit is much higher than the average, $\left|W^{CG}\right|$ would be large even when the demand shocks are small.

4 Output-Target Mechanism

Inspired by the real-life cartel practice of requiring firms selling above quota to compensate firms selling below, we show that, for many stochastic demand functions, it is possible to reduce the efficiency loss of the Clarke-Groves approach by setting the profit caps strictly below the maximum profits.

Let

$$\widetilde{y}_j\left(a, y_j'\right) \equiv \int_{\varepsilon_j} \min\left(y_j\left(a, \varepsilon_j\right), y_j'\right) dF_j\left(\varepsilon_j\right)$$

denote firm j's expected output truncated from above at y_j'. For any i, $j \in \mathcal{N}$ with $j \neq i$ and any $a \in A$ such that $\partial \overline{y}_j(a)/\partial a_i \neq 0$, define

$$\lambda_{ij}\left(a, y_j'\right) \equiv \frac{\frac{\partial \widetilde{y}_j\left(a, y_j'\right)}{\partial a_i}}{\frac{\partial \overline{y}_j(a)}{\partial a_i}}$$

as the fraction of the marginal effect of a_i on y_j that is captured by $\widetilde{y}_j\left(a, y_j'\right)$. Since y_j is either increasing in a_i for all ε_j or decreasing in a_i for all ε_j, $\lambda_{ij}\left(a, y_j'\right)$ is between 0 and 1.

An output-target function $y_j^K : C \rightarrow \Re_+$ is a function that maps each cost-report profile to an output target. Let

$$K_j(\widehat{c}) \equiv \Phi_1 \left(\widehat{c}_j, \alpha_j^* (\widehat{c})\right) y_j^K (\widehat{c}) + \Phi_2 \left(\widehat{c}_j, \alpha_j^* (\widehat{c})\right) \tag{7}$$

denote the profit target associated with $y_j^K (\widehat{c})$; that is, the profit of firm j when its cost is $\widehat{c}_j$, action is $\alpha_j^* (\widehat{c})$, and output is $y_j^K (\widehat{c})$. We say that y_j^K partially captures the effect of a_{-i} on y_j if

$$\lambda_{ij} \left(\alpha^* (\widehat{c}) , y_j^K (\widehat{c})\right) > 0, \quad \text{for each } \widehat{c} \in C \text{ and each } i \neq j. \tag{8}$$

To simplify exposition, we will assume $\partial \overline{y}_i (\alpha^* (c)) / \partial a_j \neq 0$ for any firms i, j and any $c \in C$ so that an output-target function that satisfies (8) always exists. The assumption, however, is not crucial. See footnote 13 for a general version of the mechanism.

Given a vector of output-target functions $y^K \equiv \left(y_1^K, \ldots, y_n^K\right)$ that satisfies (8) for each y_j^K, we set the transfer of each firm i to

$$w_i^* (\widehat{c}, \widehat{y}) \equiv \widehat{w}_i^1 (\widehat{c}, \widehat{y}_{-i}) + \widehat{w}_i^2 (\widehat{c}_i) - \frac{1}{n-1} \sum_{j \neq i} \widehat{w}_j^2 (\widehat{c}_j) , \tag{9}$$

where

$$\widehat{w}_i^1 (\widehat{c}, \widehat{y}_{-i}) \equiv \sum_{j \neq i} \frac{\min \left(\widetilde{\pi}_j (\widehat{y}_j, \widehat{c}) , K_j (\widehat{c})\right) - K_j (\widehat{c})}{\lambda_{ij} \left(\alpha^* (\widehat{c}) , y_j^K (\widehat{c})\right)}; \tag{10}$$

$$\begin{aligned} \widehat{w}_i^2 (\widehat{c}_i) \equiv & \sum_{j \neq i} \int_{c_{-i}} \int_{\varepsilon_j} \pi_j \left(\alpha^* (\widehat{c}_i, c_{-i}) , c_j, \varepsilon_j\right) dF_j (\varepsilon_j) \, dG_{-i} (c_{-i}) \\ & - \int_{c_{-i}} \int_{\varepsilon_{-i}} \widehat{w}_i^1 \left(\widehat{c}_i, c_{-i}, y_{-i} \left(\alpha^* (\widehat{c}_i, c_{-i}) , \varepsilon_{-i}\right)\right) dF_{-i} (\varepsilon_{-i}) \, dG_{-i} (c_{-i}) . \end{aligned} \tag{11}$$

The first component, $\widehat{w}_i^1$, requires firm i pay a penalty equal to the profit shortfall of each firm $j \neq i$, multiplied by λ_{ij}^{-1}. The second component of the transfer, $\widehat{w}_i^2$, captures the difference between the expected value of $\widehat{w}_i^1$ and the expected profits of firms $j \neq i$. Since the second and third components of (9) sum to zero across firms, the total transfer is

$$\sum_{i=1}^{n} \widehat{w}_i^1 \left(\widehat{c}, \widehat{y}_{-i}\right) \leq 0$$

for all $(\widehat{c}, \widehat{y}) \in C \times \Re_+^n$.

We call $w^* \equiv (w_1^*, \ldots, w_n^*)$ an output-target mechanism. Note that if we set each $y_j^K(\widehat{c})$ to $y_j(\alpha^*(\widehat{c}), \overline{\varepsilon}_j)$, then $K_j(\widehat{c})$ would be equal to $\pi_j(\alpha^*(\widehat{c}), \widehat{c}_j, \overline{\varepsilon}_j)$, and the mechanism would become a generalized Clarke-Groves mechanism. The two mechanisms, however, are different when some $y_j^K(\widehat{c})$ is strictly less than $y_j(\alpha^*(\widehat{c}), \overline{\varepsilon}_j)$. Our main result is that despite the truncation, an output-target mechanism may still implement α^*. We proceed in two steps. First, we show that an output-target mechanism implements α^* under a monotone condition. We then show that the monotone condition is satisfied by a wide class of stochastic demand functions.

Definition 1. The demand functions of the firms satisfy the *monotone λ condition* if for any i, $j \in \mathcal{N}$ with $i \neq j$, $a_{-i} \in A_{-i}$, $a_i', a_i'' \in A_i$, and $y_j' \in \Re_+$,

$$\lambda_{ij}\left(a_i', a_{-i}, y_j'\right) \geq \lambda_{ij}\left(a_i'', a_{-i}, y_j'\right)$$

whenever $\overline{y}_j(a_i', a_{-i}) \leq \overline{y}_j(a_i'', a_{-i})$ and both $\lambda_{ij}\left(a_i', a_{-i}, y_j'\right)$ and $\lambda_{ij}\left(a_i'', a_{-i}, y_j'\right)$ are defined.

The monotone λ condition says that the truncated output function of firm j captures a larger fraction of the marginal effect of firm i's action when firm i chooses an action that lowers firm j's output.[13] Since the output target of each firm $j \neq i$ captures a strictly positive fraction of the external effect of firm i's action, firm i's transfer can be scaled up so that the marginal effect of a_i on the

[13]Given the monotone λ condition, we can dispense with the assumption that

transfer exactly matches its marginal effect on the profits of the other firms. The monotone λ condition implies that the punishment for a deviation that hurts firm j is greater than the loss of firm j, while the reward for a deviation that helps firm j is smaller than the gain of firm j. This, combined with the fact that α^* is efficient, implies the following result.

Proposition 1. In the static collusion game with transfers, it is a Nash equilibrium for each firm i to report c_i and y_i truthfully and choose $\alpha_i^*(\widehat{c})$ under an output-target mechanism if the demand functions satisfy the monotone λ condition.

Proof. Let ρ_i^* denote the truth-telling cost-reporting strategy and r_i^* the truth-telling output-reporting strategy. For any ρ_i and γ_i, let α^{ρ_i,γ_i} denote the cost-action profile induced by firm i choosing (ρ_i, γ_i) and firms $j \neq i$ choosing $\left(\rho_{-i}^*, \alpha_{-i}^*\right)$. That is, for each $j \in \mathcal{N}$ and $c \in C$,

$$\alpha_j^{\rho_i,\gamma_i}(c) \equiv \begin{cases} \gamma_i(c) & \text{if} \quad j = i \\ \alpha_j^*(\rho_i(c_i), c_{-i}) & \text{if} \quad j \neq i \end{cases}.$$

$\partial \overline{y}_j(\alpha^*(c))/\partial a_i \neq 0$. Let

$$B_{ij}(a) \equiv \left\{ a_i' \in A_i \,\middle|\, \overline{y}_j(a_i', a_{-i}) \leq \overline{y}_j(a) \text{ and } \frac{\partial \overline{y}_j(a_i', a_{-i})}{\partial a_i} \neq 0 \right\},$$

and

$$\widetilde{\lambda}_{ij}\left(a, y_j'\right) \equiv \begin{cases} \inf_{a_i' \in B_{ij}(a)} \dfrac{\frac{\partial \widetilde{y}_j\left(a_i', a_{-i}, y_j'\right)}{\partial a_i}}{\frac{\partial \overline{y}_j\left(a_i', a_{-i}\right)}{\partial a_i}} & \text{if } B_{ij}(a) \text{ is non-empty;} \\ \infty & \text{otherwise.} \end{cases} \tag{12}$$

We can always pick $y_j^K(\widehat{c})$ such that $\widetilde{\lambda}_{ij}\left(\alpha^*(\widehat{c}), y_j^K(\widehat{c})\right) > 0$. All results will go through by replacing $\lambda_{ij}\left(\alpha^*(\widehat{c}), y_j^K(\widehat{c})\right)$ by $\widetilde{\lambda}_{ij}\left(\alpha^*(\widehat{c}), y_j^K(\widehat{c})\right)$ in the definition of $\widehat{w}_i^1$ in (10).

If firms $j \neq i$ choose $\left(\rho_{-i}^{*}, \alpha_{-i}^{*}\right)$, firm i would receive an expected profit equal to

$$
\begin{aligned}
& u_i\left(\rho_i, \gamma_i, r_i ; w_i^*\right) \equiv \bar{\pi}_i\left(\alpha^{\rho_i, \gamma_i}\right) \\
& +\int_c \int_{\varepsilon} w_i^*\left(\rho_i\left(c_i\right), c_{-i}, r_i\left(c, y_i\left(\alpha^{\rho_i, \gamma_i}(c), \varepsilon_i\right)\right), y_{-i}\left(\alpha^{\rho_i, \gamma_i}(c), \varepsilon_{-i}\right)\right) d F(\varepsilon) d G(c)
\end{aligned}
$$

in the static collusive game if it chooses cost-reporting strategy ρ_i, action strategy γ_i, and output-reporting strategy r_i. Let Σ_i denote the set of all cost-reporting, action, and output-reporting strategies. The strategy profile $\left(\rho^*, \alpha^*, r^*\right)$ is a Nash equilibrium of the static collusive game if for each firm i

$$
\left(\rho_i^*, \alpha_i^*, r_i^*\right) \in \arg \max_{\left(\rho_i, \gamma_i, r_i\right) \in \Sigma_i} u_i\left(\rho_i, \gamma_i, r_i ; w_i^*\right). \tag{13}
$$

Since w_i^* does not depend on $\widehat{y}_i$, it is a best response for firm i to report its output truthfully. It is therefore sufficient to prove that $\left(\rho_i^*, \alpha_i^*\right)$ maximizes $u_i\left(\rho_i, \gamma_i, r_i^* ; w_i^*\right)$ with respect to all $\left(\rho_i, \gamma_i\right)$.

Define, for any firms i, $j \in \mathcal{N}$ with $i \neq j$ and for any $c \in C$,

$$
\begin{aligned}
H_{ij}\left(c, a_i\right) \equiv & \frac{1}{\lambda_{ij}\left(\alpha^*(c), y_j^K(c)\right)} \int_{\varepsilon_j} \min \left(\pi_j\left(a_i, \alpha_{-i}^*(c), c_j, \varepsilon_j\right)-K_j(c), 0\right) d F_j\left(\varepsilon_j\right) \\
& -\int_{\varepsilon_j} \pi_j\left(a_i, \alpha_{-i}^*(c), c_j, \varepsilon_j\right) d F_j\left(\varepsilon_j\right).
\end{aligned} \tag{14}
$$

Substituting (9)-(11) and (14) into $u_i\left(\rho_i, \gamma_i, r_i^* ; w_i^*\right)$, we have

$$
\begin{aligned}
u_i\left(\rho_i, \gamma_i, r_i^* ; w^*\right)= & \sum_{j=1}^n \bar{\pi}_j\left(\alpha^{\rho_i, \gamma_i}\right)+\sum_{j \neq i} \int_c H_{ij}\left(\rho_i\left(c_i\right), c_{-i}, \gamma_i(c)\right) d G(c) \\
& +\int_{c_i} \widehat{w}_i^2\left(\rho_i\left(c_i\right)\right) d G_i\left(c_i\right)-\frac{1}{n-1} \sum_{j \neq i} \int_{c_j} \widehat{w}_j^2\left(c_j\right) d G_j\left(c_j\right).
\end{aligned} \tag{15}
$$

The first term of (15) is the total cartel profit when firm i chooses $\left(\rho_i, \gamma_i\right)$ and firms $-i$ choose $\left(\rho_{-i}^*, \alpha_{-i}^*\right)$. Since α^* is efficient, this term is maximized when firm i chooses $\left(\rho_i^*, \alpha_i^*\right)$.

We now turn to the second term on the right-hand side of (15). For each $c \in C$ and $a_i \in A_i$,

$$\frac{\partial H_{ij}}{\partial a_i} = \Phi_1\left(c_j, \alpha_j^*(c)\right)\left(\frac{1}{\lambda_{ij}\left(\alpha^*(c), y_j^K(c)\right)}\frac{\partial \widetilde{y}_j}{\partial a_i} - \frac{\partial \overline{y}_j}{\partial a_i}\right).$$

By the definition of λ_{ij}, $\partial H_{ij}/\partial a_i = 0$ when $a_i = \alpha_i^*(c)$. This, together with the monotone λ condition, implies that

$$\frac{\partial H_{ij}(c, a_i)}{\partial a_i} \geq 0 \quad \text{if and only if} \quad a_i \leq \alpha_i^*(c).$$

Since H_{ij} is continuous in a_i, $\alpha_i^*(c) \in \arg\max_{a_i \in A_i} H_{ij}(c, a_i)$ for any $c \in C$. Note that by definition, for any $\widehat{c}_i \in C_i$,

$$\widehat{w}_i^2(\widehat{c}_i) = -\sum_{j \neq i} \int_{c_{-i}} H_{ij}\left(\widehat{c}_i, c_{-i}, \alpha_i^*(\widehat{c}_i, c_{-i})\right) dG_{-i}(c_{-i}).$$

Thus, for any ρ_i and γ_i,

$$\begin{aligned}\int_{c_i \in C_i} \widehat{w}_i^2(\rho_i(c_i))\, dG_i(c_i) &= -\sum_{j \neq i} \int_c H_{ij}\left(\rho_i(c_i), c_{-i}, \alpha_i^*(\rho_i(c_i), c_{-i})\right) dG(c) \\ &\leq -\sum_{j \neq i} \int_c H_{ij}\left(\rho_i(c_i), c_{-i}, \gamma_i(c_i, c_{-i})\right) dG(c).\end{aligned}$$

Thus, the sum of the second and third terms on the right-hand side of (15) is maximized by any strategy (ρ_i, γ_i) that satisfies the condition that $\gamma_i(c_i, c_{-i}) = \alpha_i^*(\rho_i(c_i), c_{-i})$, including (ρ_i^*, α_i^*) in particular. Since (ρ_i^*, α_i^*) also maximizes the first term of (15), $u_i(\rho_i, \gamma_i, r_i^*; w^*)$ is maximized when $(\rho_i, \gamma_i) = (\rho_i^*, \alpha_i^*)$. □

Proposition 2. The monotone λ condition holds if one of the following conditions is satisfied for each firm $i \in \mathcal{N}$:

1. there exists a differentiable function $\chi_i(a)$ with non-zero partial derivatives with respect to a_{-i} such that for any ε_i and any $j \neq i$, $\frac{\partial y_i(a, \varepsilon_i)}{\partial a_j} = \frac{\partial \chi_i(a)}{\partial a_j}$;

2. there exists a differentiable function $\chi_i(a)$ with non-zero partial derivatives with respect to a_{-i} such that for any ε_i and any $j \neq i$, $\frac{\partial y_i(a,\varepsilon_i)}{\partial a_j} = \varepsilon_i \frac{\partial \chi_i(a)}{\partial a_j}$;

3. the density of ε_i is differentiable and log-concave, and there exist a differentiable function $\chi_i(a)$ with non-zero partial derivatives with respect to a_{-i}, an increasing and differentiable function h_i, and a function ν_i such that for any ε_i and any $j \neq i$, $\frac{\partial y_i(a,\varepsilon_i)}{\partial a_j} = \nu_i(a_i)\, h_i'(\chi_i(a) + \varepsilon_i) \frac{\partial \chi_i(a)}{\partial a_j}$.

Proposition 2 identifies conditions under which the monotone λ condition is satisfied. The conditions limit the interaction between the random component (ε_i) and the deterministic component (χ_i) of firm i's output function, but impose no restrictions on the functional form of χ_i. As the following examples show, these conditions are satisfied by many stochastic demand functions.[14]

Example 1 (Linear Demand).

$$y_i(a, \varepsilon_i) = \sum_{j=1}^{n} l_{ij} a_j + \varepsilon_i.$$

Example 2 (Log-linear demand).

$$y_i(a, \varepsilon_i) = \varepsilon_i \Pi_{j=1}^{n} a_j^{l_{ij}}.$$

Example 3 (Logit demand).

$$y_i(a, \varepsilon_i) = \frac{\exp(-l_{ii} a_i)}{\overline{\varepsilon}_i - \varepsilon_i + \sum_{j=1}^{n} \exp(-l_{ij} a_j)}.$$

A formal proof of Proposition 2 is provided in the appendix. To understand Proposition 2, let

$$\Gamma_j(y_j', a) \equiv \{\varepsilon_j \in \Omega_j | y_j(a, \varepsilon_j) \leq y_j'\}$$

[14]In the examples, l_{ij} and l_{ii} are constants. It is clear that Examples 1 and 2 satisfy conditions 1 and 2 of Proposition 2, respectively. To see that Example 3 satisfies the second part of condition 3, set $h_i(x) = 1/(\overline{\varepsilon}_i - x)$, $\chi_i(a) = -\sum_{j=1}^{n} \exp(-l_{ij} a_j)$, and $\nu_i(a_i) = \exp(-l_{ii} a_i)$.

denote the set of ε_j for which y_j is less than y'_j. We can then write

$$\lambda_{ij}\left(a, y'_j\right) = \frac{\int_{\varepsilon_j \in \Gamma_j\left(y'_j, a\right)} \frac{\partial y_j(a,\varepsilon_j)}{\partial a_i} dF_j\left(\varepsilon_j\right)}{\int_{\varepsilon_j \in \Omega_j} \frac{\partial y_j(a,\varepsilon_j)}{\partial a_i} dF_j\left(\varepsilon_j\right)}. \tag{16}$$

Note that a_i affects λ_{ij}, first, through Γ_j and, second, through the variation of $\partial y_j / \partial a_i$ across ε_j. We call the first channel the cutoff effect and the second the distribution effect. Since $y_j\left(a, \varepsilon_j\right)$ is either increasing in a_i for all ε_j or decreasing in a_i for all ε_j, for any $a_i, a'_i \in A_i$ and any $a_{-i} \in A_{-i}$, $\overline{y}_j\left(a_i, a_{-i}\right) \leq \overline{y}_j\left(a'_i, a_{-i}\right)$ implies $y_j\left(a_i, a_{-i}, \varepsilon_j\right) \leq y_j\left(a'_i, a_{-i}, \varepsilon_j\right)$ for all ε_j. Hence,

$$\Gamma_j\left(y'_j, a'_i, a_{-i}\right) \subseteq \Gamma_j\left(y'_j, a_i, a_{-i}\right)$$

if $\overline{y}_j\left(a_i, a_{-i}\right) \leq \overline{y}_j\left(a'_i, a_{-i}\right)$. Intuitively, any action by firm i that lowers firm j's expected output would increase the probability that y_j is less than y'_j. Since the truncated output function captures the external effect of a_i only when y_j is below y'_j, the cutoff effect always raises λ_{ij} when firm i chooses an action that harms firm j. The distribution effect of a_i, however, is ambiguous. The conditions in Proposition 2 ensures that the distribution effect does not overwhelm the cutoff effect.

The total expected transfer under an output-target mechanism with a target profile y^K is

$$W\left(y^K\right) \equiv \sum_{i=1}^{n} \sum_{j \neq i} \int_c \int_{\varepsilon_j} \frac{\min\left(\pi_j\left(\alpha^*\left(c\right), c_j, \varepsilon_j\right) - K_j(c), 0\right)}{\lambda_{ij}\left(\alpha^*\left(c\right), y_j^K\left(c\right)\right)} dF_j\left(\varepsilon_j\right) dG\left(c\right). \tag{17}$$

Recall that an output-target mechanism with $y_j^K\left(c\right) = y_j\left(\alpha^*\left(c\right), \overline{\varepsilon}_j\right)$ for all $j \in \mathcal{N}$ and $c \in C$ is equivalent to a generalized Clarke-Groves mechanism. It is straightforward to check that in that case $\lambda_{ij}\left(\alpha^*\left(c\right), y_j^K\left(c\right)\right) = 1$ and $W\left(y^K\right) = W^{CG}$. Lowering $y_j^K\left(c\right)$ lowers K_j but also λ_{ij}, as a truncated output function does not fully captures the effect of a_i on y_j.

To get a better idea of the overall welfare effect, suppose, for each firm j, $y_j = \varepsilon_j \chi_j(a)$, $\underline{\varepsilon}_j \geq 0$, and $\int \varepsilon_j dF_j(\varepsilon_j) = \left(\overline{\varepsilon}_j + \underline{\varepsilon}_j\right)/2$ (e.g., when the distribution of ε_j is symmetric). Let

$$W_{ijc}\left(y_j^K\right) \equiv \int_{\varepsilon_j} \frac{\min\left(\pi_j\left(\alpha^*(c), c_j, \varepsilon_j\right) - \left(\Phi_1\left(c_j, \alpha_j^*(c)\right) y_j^K + \Phi_2\left(c_j, \alpha_j^*(c)\right)\right), 0\right)}{\lambda_{ij}\left(\alpha^*(c), y_j^K\right)} dF_j(\varepsilon_j)$$

denote the part of firm i's "punishment" that is tied to firm j's reported output for a specific cost profile c. The left derivative of W_{ijc} evaluated at $y_j\left(\alpha^*(c), \overline{\varepsilon}_j\right)$ has the same sign as

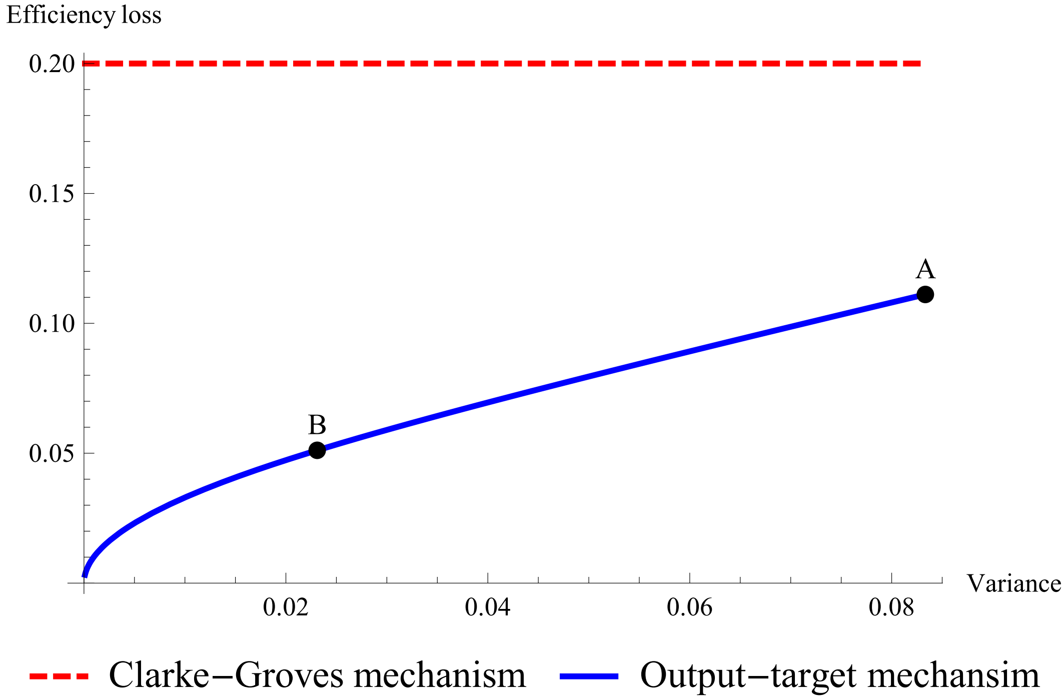
$$\frac{\left(\overline{\varepsilon}_j - \underline{\varepsilon}_j\right)\overline{\varepsilon}_j}{\overline{\varepsilon}_j + \underline{\varepsilon}_j} \frac{dF_j\left(\overline{\varepsilon}_j\right)}{d\varepsilon_j} - 1.$$

Hence, the left derivative is negative (meaning that lowering y_j^K below $y_j\left(\alpha^*(c), \overline{\varepsilon}_j\right)$ reduces $\left|W_{ijc}\left(y_j^K\right)\right|$) if $dF_j\left(\overline{\varepsilon}_j\right)/d\varepsilon_j$ is less than $1/\left(\overline{\varepsilon}_j - \underline{\varepsilon}_j\right)$, the density of a uniform distribution on the same support. Thus, when the demand functions are log-linear and the distributions of random shocks are symmetric and hump-shaped, there is always an output-target mechanism (with output targets below the maximum) that strictly outperforms the generalized Clarke-Groves mechanism. Intuitively, when large random shocks are unlikely, a slight reduction of the output targets has little impact on the incentives of the firms and, hence, relatively small effect on λ_{ij}.

Figure 2 depicts the efficiency loss as a fraction of the monopoly profit of an output-target mechanism of a symmetric duopoly with log-linear demand and truncated-normal random shocks.[15] The output target is set equal to the average monopoly demand. The maximum monopoly demand is twenty percent higher than the mean. The efficiency loss of the output-target mechanism is about 11 percent when the demand shock is uniformly distributed (point A in Figure 2). As the variance of the random shock decreases, the efficiency loss further decreases. If

[15]For simplicity we assume there is no cost shocks.

Figure 2: Comparison of efficiency loss for a symmetric duopoly. Demand is log-linear $y_i(a, \varepsilon_i) = \varepsilon_i \Pi_{j=1}^{2} a_j^{l_{ij}}$, random shock ε_i follows a normal distribution with mean 2.5 and truncated between $[2, 3]$, and costs are common knowledge. Variance in the graph refers to the actual variance of ε_i, rather than that of the underlying normal distribution.



the probability that the demand is within ten percent of the mean is greater than 0.9 (point B in Figure 2), the efficiency loss is less than 5.1 percent. By contrast, the efficiency loss of the generalized Clarke-Groves mechanism, $|W^{CG}|$, is always twenty percent of the monopoly profit.

In general the optimal output targets depend on the functional form of the demand functions, as well as on the distribution of the random shocks. Choosing the optimal targets, however, may not be crucial. In the above example, setting the output target at the mean is not optimal.[16] Nevertheless, the efficiency loss converges to zero when the variance of the demand shock is small. In fact, for any demand functions that satisfy the monotone λ condition, there always exists an output-target mechanism that is nearly efficient when the random shock is sufficiently small.

Proposition 3. Suppose the demand system satisfies the monotone λ condition. Then, given any $\zeta > 0$, there exists κ such that if for all $i \in \mathcal{N}$

$$F_i\left(\int \varepsilon_i dF_i(\varepsilon_i) + \kappa\right) - F_i\left(\int \varepsilon_i dF_i(\varepsilon_i) - \kappa\right) \geq 1 - \kappa, \tag{18}$$

then there exists an output-target mechanism with efficiency loss less than ζ.

The proof of Proposition 3 is provided in the appendix. The inequality (18) says that the probability that the demand shock ε_i is different from its mean by κ is less than κ. Since y_i is continuous in ε_i, the distribution of each firm's output would be concentrated around the mean when κ is small. In this case output targets that are slightly above the mean would capture a large fraction of the external effects of the firms' actions. Each λ_{ij} would therefore be close to one. Since the equilibrium outputs of the firms are unlikely to be significantly different from the mean, the numerator in the integrand of the right-hand side of (17) would be close to zero.

[16] In this example because the support of output shifts with the action of the firms, the optimal cutoff should be set as low as possible (subject to the constraint that the maximum side-payment does not exceed some upper bound). In general the optimal cutoff could be in the interior when the support does not shift. The example is chosen to illustrate the welfare effect of truncation. For that purpose whether the support is shifting is not important.

5 Implementation

In this section we apply the results in the last section to construct a perfect public equilibrium in the original repeated game. The equilibrium trigger-strategy profile, denoted by $\mathcal{S}(\beta, \mu)$, is characterized by a probability function $\mu : C \times \Re_+^n \to [0,1]$ and a matrix of side-payment functions $\beta \equiv \{\beta_{ij}\}_{i,j \in \mathcal{N}}$, where $\beta_{ij} : C \times \Re_+^n \to \Re_+$. There are two states: collusive and non-collusive. The equilibrium starts off in the collusive state in period 1. In the collusive state each firm i reports its cost c_i truthfully, chooses the action $\alpha_i^* (c_i, \widehat{c}_{-i})$, and reports the output y_i truthfully. Each firm i then pays $\beta_{ij}(\widehat{c}, \widehat{y})$ to each firm j. If all firms make the required side-payments, the equilibrium will stay in the collusive state in the next period with probability $1 - \mu(\widehat{c}, \widehat{y})$ and switch to the non-collusive state with probability $\mu(\widehat{c}, \widehat{y})$. If some firms do not make the required payments, then the game will switch to the non-collusive state in the next period with probability one. In the non-collusive state firm i sends the same cost report c_i'', chooses the action strategy α_i^{NE}, sends the same output report y_i'', and makes no side-payments.[17] The non-collusive state is absorbing. Once the equilibrium enters the non-collusive state, it stays there forever.

Proposition 4. Suppose the demand functions satisfy the monotone λ condition, and there is an output-target mechanism $w^* = (w_1^*, \ldots, w_n^*)$ with efficiency loss $|W| < \overline{\Pi}(\alpha^*) - \overline{\Pi}\left(\alpha^{NE}\right)$. Then there exists $\delta^* < 1$ such that for any $\delta \geq \delta^*$, there exist β and μ so that the trigger-strategy profile $\mathcal{S}(\beta, \mu)$ is a perfect public equilibrium with total average discounted profit of the firms equal to $\overline{\Pi}(\alpha^*) - |W|$.

It is obvious that no firm can deviate profitably in the non-collusive state. We can therefore focus on the collusive state. In the Appendix we show that it is

[17] Any $c_i'' \in C_i$ and $y_i'' \in \Re_+$ will suffice.

possible to set (β, μ) such that the continuation profit of each firm i (including side-payments) equals w_i^* plus a constant d_i. Intuitively, when δ is sufficiently large, we can first choose μ so that $|W|$ is destroyed through the price war, and then use β to adjust the continuation profit of each firm. The constant d_i is chosen so that firm i prefers making the required side-payments to switching to the non-collusive state.

The following corollary follows immediately from Propositions 3 and 4.

Corollary 1. Suppose the demand functions satisfy the monotone λ condition. Then, for any $\zeta > 0$, there exists κ such that if the demand shock for each firm i is smaller than κ in the sense of (18), then there exists a perfect public equilibrium with total average discounted profit of the firms greater than $\overline{\Pi}(\alpha^*) - \zeta$.

The main features of our equilibrium are consistent with transfer schemes used by many cartels discussed in Harrington and Skrzypacz (2011). Harrington and Skrzypacz (2011) propose a different equilibrium to explain these features. Corollary 1 is the counterpart to their Theorem 1, which show that the equilibrium cartel profit converges to the monopoly level when the demand shocks go to zero. In the equilibrium of Harrington and Skrzypacz (2011) the transfers serve as an output tax. To figure out the right tax rate, the cartel would need to know the cost of the firm. It is therefore important in their model that costs are public. Since in their model a firm is required to pay more when it reports higher sales, it would have an incentive to under-report. Harrington and Skrzypacz (2011) show that when the industry demand is stochastic but completely inelastic with respect to the prices of the firms, truth-telling is incentive compatible when the probability of the industry demand reaching the maximum value is sufficiently high. In our equilibrium since the output report of each firm affects only the continuation profits of the other firms, there is no incentive to mis-report. However, since a firm is punished only when the sales of another firm is below quota, the punishment does not fully capture the external effects of a firm's action in our model. Our contribution is to show that this problem can be solved when the demand functions satisfy the monotone λ condition.

6 Concluding Remarks

In this paper we show that for a wide class of stochastic demand functions, it is possible to enforce the cartel-profit-maximizing outcome by tying each firm's continuation profit to the truncated profits of the other firms. We show that when the firms are patient and the demand shocks are small there is a perfect public equilibrium with cartel profit close to the monopoly level. One lesson we can draw from the analysis is that without active antitrust enforcement, a cartel can operate successfully even when it cannot directly monitor the costs, prices, or sales of the firms.

Our mechanism can be viewed as a truncated Clarke-Groves mechanism. Despite the truncation, the mechanism inherits many nice properties of a standard Clarke-Groves mechanism. Existing models of repeated games assume either costs or actions are public. In reality firms are likely to set price secretly and possess private information about costs and sales. Our mechanism solves the private-monitoring and private-information problems simultaneously. The mechanism is very robust. While we have assumed that firms are choosing either prices or quantities, the approach also applies when firms collude on other dimensions. In an earlier draft of this paper, we also show that the mechanism can be easily extended to allow for multi-market collusion.

One common criticism against models of repeated games is that the strategies are unrealistically complex. An advantage of our mechanism is that its underlying idea—forcing each firm to internalize the externalities of its action—is simple and intuitive. As Harrington and Skrzypacz (2011) have documented, compensation schemes whereby firms that sell above quotas are required to compensate firms that sell below were common among many recent cartels. Suslow (2005) finds that cartels that used some type of self-imposed penalty schemes were more stable. Compared to the equilibrium proposed by Harrington and Skrzypacz (2011), our mechanism require neither costs be public nor the industry demand be inelastic with respect to prices. The robustness helps explain why these schemes were employed by cartels in different industries.

In our equilibrium the continuation profits of the firms in the collusive state depend solely on the current-period reports. It is well-known since the seminal works of Rubinstein (1979), Rubinstein and Yaari (1983), Radner (1983), and Abreu, Milgrom, and Pearce (1991) that it is possible to reduce the efficiency loss due to imperfect monitoring by linking punishment decisions across periods. We show how those techniques can be combined with the Clarke-Groves approach in a private-monitoring setting (without cost shocks) in a separate paper (Chan and Zhang, forthcoming). Finally, the cost shocks in our model are uncorrelated over time. Recent works (Athey and Bagwell, 2008, Athey and Segal, 2013, among others) have considered a more general dynamic environment with Markovian types. Those models, however, do not allow private actions that directly affect other players' payoffs. Extending our model to the case of Markovian types would be a worthwhile extension.

A Proof of Proposition 2

In Conditions 1 and 2, the marginal effect of a_j on y_i can be written as

$$\frac{\partial y_i}{\partial a_j} = \xi_{ij}(a)\,\zeta_{ij}(a_{-j}, \varepsilon_i), \tag{19}$$

where ξ_{ij} is independent of ε_i, and ζ_{ij} is positive and independent of a_j. Factoring $\xi_{ij}(a)$ out of the integrals, we have

$$\lambda_{ji}(a, y_i') = \frac{\frac{\partial \widetilde{y}_i\left(a_j, a_{-j}, y_i'\right)}{\partial a_j}}{\frac{\partial \overline{y}_i(a_j, a_{-j})}{\partial a_j}} = \frac{\int_{\varepsilon_i \in \Gamma_i\left(y_i', a_j, a_{-j}\right)} \zeta_{ij}(a_{-j}, \varepsilon_i)\, dF_i(\varepsilon_i)}{\int_{\varepsilon_i \in \Omega_i} \zeta_{ij}(a_{-j}, \varepsilon_i)\, dF_i(\varepsilon_i)}. \tag{20}$$

Note that firm j's action a_j affects the right-hand side of (20) only through $\Gamma_i(y_i', a_j, a_{-j})$. For any $a_j', a_j'' \in A_j$, $\Gamma_i\left(y_i', a_j', a_{-j}\right) \subseteq \Gamma_i\left(y_i', a_j'', a_{-j}\right)$ if $y_i\left(a_j', a_{-j}, \varepsilon_i\right) > y_i\left(a_j'', a_{-j}, \varepsilon_i\right)$ for all ε_i. Hence, the ratio on the left-hand side of (20) would be smaller under a_j' than a_j''. Intuitively, because firm i's output is more likely to be lower than y_i' when firm j is choosing an action that lowers firm i's output, firm i's truncated output function $\widetilde{y}_i$ captures a bigger fraction of the marginal effect of firm j's action when firm j's action reduces firm i's output.

It remains to consider Condition 3.

Let f_i denote the density of ε_i. Recall that we assume f_i is strictly positive, differentiable, and log-concave for any $\varepsilon_i \in [\underline{\varepsilon}_i, \overline{\varepsilon}_i]$. By assumption,

$$\begin{aligned}\lambda_{ji}(a, y_i') = \frac{\frac{\partial \widetilde{y}_i\left(a, y_i'\right)}{\partial a_j}}{\frac{\partial \overline{y}_i(a)}{\partial a_j}} &= \frac{\int_{\varepsilon_i \in \Gamma_i\left(y_i', a\right)} h_i'(\chi_i(a) + \varepsilon_i)\, dF_i(\varepsilon_i)}{\int_{\varepsilon_i \in \Omega_i} h_i'(\chi_i(a) + \varepsilon_i)\, dF_i(\varepsilon_i)} \\ &= \frac{\int_{\varepsilon_i \in \Gamma_i\left(y_i', a\right)} f_i(\varepsilon_i)\, dh_i(\chi_i(a) + \varepsilon_i)}{\int_{\varepsilon_i \in \Omega_i} f_i(\varepsilon_i)\, dh_i(\chi_i(a) + \varepsilon_i)}.\end{aligned}$$

Suppose $\partial \widetilde{y}_i(a, y_i')/\partial a_j \neq 0$. Then $\Gamma_i(y_i', a)$ is nonempty. Let $\varepsilon_i' \equiv \sup \Gamma_i(y_i', a)$ and let $z' \equiv h_i(\chi_i(a) + \varepsilon_i')$. Through changing variables, we have

$$\lambda_{ji}(a, y_i') = \frac{\int_{\underline{\varepsilon}_i}^{\varepsilon_i'} f_i(\varepsilon_i)\, dh_i(\chi_i(a) + \varepsilon_i)}{\int_{\underline{\varepsilon}_i}^{\overline{\varepsilon}_i} f_i(\varepsilon_i)\, dh_i(\chi_i(a) + \varepsilon_i)} = \frac{\int_{h_i(\chi_i(a) + \underline{\varepsilon}_i)}^{z'} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz}{\int_{h_i(\chi_i(a) + \underline{\varepsilon}_i)}^{h_i(\chi_i(a) + \overline{\varepsilon}_i)} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz}.$$

We can write

$$\lambda_{ji}\left(a, y_i'\right) = \frac{1}{1 + \widehat{\lambda}_{ji}\left(a, y_i'\right)},$$

where

$$\widehat{\lambda}_{ji}\left(a, y_i'\right) = \frac{\int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz}{\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz}.$$

Differentiating $\widehat{\lambda}_{ji}\left(a, y_i'\right)$ with respect to a_j, we obtain

$$\begin{aligned} \frac{\partial \widehat{\lambda}_{ji}\left(a, y_i'\right)}{\partial a_j} =& \frac{\partial \chi_i(a)}{\partial a_j} \widehat{\lambda}_{ji}\left(a, y_i'\right) \left(\frac{\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} f_i'\left(h_i^{-1}(z) - \chi_i(a)\right) dz}{\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz} \right. \\ & - \frac{\int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} f_i'\left(h_i^{-1}(z) - \chi_i(a)\right) dz}{\int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz} + \frac{h_i'\left(\chi_i(a) + \underline{\varepsilon}_i\right) f_i\left(\underline{\varepsilon}_i\right)}{\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz} \\ & \left. + \frac{h_i'\left(\chi_i(a) + \overline{\varepsilon}_i\right) f_i\left(\overline{\varepsilon}_i\right)}{\int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz} \right). \end{aligned} \tag{21}$$

Let

$$\begin{aligned} \gamma_1(z, a) &\equiv \frac{f_i\left(h_i^{-1}(z) - \chi_i(a)\right)}{\int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz}; \\ \gamma_2(z, a) &\equiv \frac{f_i\left(h_i^{-1}(z) - \chi_i(a)\right)}{\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} f_i\left(h_i^{-1}(z) - \chi_i(a)\right) dz}. \end{aligned}$$

Since h_i is increasing and f_i is log-concave, $f_i'\left(h_i^{-1}(z) - \chi_i(a)\right)/f_i\left(h_i^{-1}(z) - \chi_i(a)\right)$ is decreasing in z. Hence the first two terms inside the bracket on the right-hand side of (21) is equal to

$$\frac{\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} f_i'\left(h_i^{-1}(z)-\chi_i(a)\right)dz}{\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} f_i\left(h_i^{-1}(z)-\chi_i(a)\right)dz} - \frac{\int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} f_i'\left(h_i^{-1}(z)-\chi_i(a)\right)dz}{\int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} f_i\left(h_i^{-1}(z)-\chi_i(a)\right)dz}$$

$$=\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} \gamma_2(z,a)\frac{f_i'\left(h_i^{-1}(z)-\chi_i(a)\right)}{f_i\left(h_i^{-1}(z)-\chi_i(a)\right)}dz - \int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} \gamma_1(z,a)\frac{f_i'\left(h_i^{-1}(z)-\chi_i(a)\right)}{f_i\left(h_i^{-1}(z)-\chi_i(a)\right)}dz$$

$$\geq \frac{f_i'(\varepsilon_i')}{f_i(\varepsilon_i')}\left(\int_{h_i(\chi_i(a)+\underline{\varepsilon}_i)}^{z'} \gamma_2(z,a)dz - \int_{z'}^{h_i(\chi_i(a)+\overline{\varepsilon}_i)} \gamma_1(z,a)dz\right)$$

$$=0.$$

Since the other two terms inside the bracket on the right-hand side of (21) are nonnegative, we have $\partial\widehat{\lambda}_{ji}(a,y_i')/\partial a_j \geq 0$ if and only if $\partial\chi_i(a)/\partial a_j \geq 0$. Since $\partial\lambda_{ji}(a,y_i')/\partial a_j$ and $\partial\widehat{\lambda}_{ji}(a,y_i')/\partial a_j$ have opposite signs, $\partial\lambda_{ji}(a,y_i')/\partial a_j \leq 0$ if and only if $\partial\chi_i(a)/\partial a_j \geq 0$.

We need to prove that for any a_j' and a_j such that $\lambda_{ji}\left(a_{-j},a_j',y_i'\right)$ and $\lambda_{ji}(a_{-j},a_j,y_i')$ are well-defined, $\lambda_{ji}\left(a_{-j},a_j',y_i'\right) \leq \lambda_{ji}(a_{-j},a_j,y_i')$ if and only if $\overline{y}_i\left(a_{-j},a_j'\right) \geq \overline{y}_i(a_{-j},a_j)$. We focus on the case where $\partial\chi_i(a)/\partial a_j \geq 0$. The case for $\partial\chi_i(a)/\partial a_j \leq 0$ is similar and omitted. Suppose $\partial\chi_i(a)/\partial a_j \geq 0$ and $a_j' \geq a_j$. Suppose that $\lambda_{ji}(a_{-j},a_j,y_i') = 0$. Since $f_i(\varepsilon_i)$ is strictly positive for any $\varepsilon_i \in [\underline{\varepsilon}_i,\overline{\varepsilon}_i]$, we must have either $\Gamma_i(y_i',a_{-j},a_j) = \emptyset$ or $h_i(\chi_i(a_{-j},a_j)+\underline{\varepsilon}_i) = z'$. Since both $y_i(a_{-j},a_j,\varepsilon_i)$ and $h_i(\chi_i(a_{-j},a_j)+\underline{\varepsilon}_i)$ are nondecreasing in a_j, we must also have either $\Gamma_i(y_i',a_{-j},a_j') = \emptyset$ or $h_i\left(\chi_i(a_{-j},a_j')+\underline{\varepsilon}_i\right) = z'$. Hence $\lambda_{ji}\left(a_{-j},a_j',y_i'\right) = 0$. Suppose that $\lambda_{ji}(a_{-j},a_j,y_i') > 0$ and $\lambda_{ji}\left(a_{-j},a_j',y_i'\right) = 0$. Then clearly $\lambda_{ji}\left(a_{-j},a_j',y_i'\right) < \lambda_{ji}(a_{-j},a_j,y_i')$. Finally, suppose that both $\lambda_{ji}(a_{-j},a_j,y_i') > 0$ and $\lambda_{ji}\left(a_{-j},a_j',y_i'\right) > 0$. Then $\lambda_{ji}\left(a_{-j},a_j'',y_i'\right) > 0$ for all $a_j'' \in \left[a_j,a_j'\right]$. From the argument in the last paragraph, $\partial\lambda_{ji}\left(a_{-j},a_j'',y_i'\right)/\partial a_j$ is defined and negative for all $a_j'' \in \left[a_j,a_j'\right]$. Hence, in this case $\lambda_{ji}\left(a_{-j},a_j',y_i'\right)$ must also be less than $\lambda_{ji}(a_{-j},a_j,y_i')$.

B Proof of Proposition 3

To prove Proposition 3, it suffices to show that for any $\zeta > 0$ we can find κ small enough such that if the size of the demand shocks is smaller than κ, then we can choose the output target $y_j^K(c)$ for each $j \in \mathcal{N}$ and each $c \in C$ such that $\lambda_{ij}\left(\alpha^*(c), y_j^K(c)\right) > 0$ for each $i \neq j$ (so that the mechanism is well-defined), and $|W| \leq \zeta$.

Since $y_i(a, \varepsilon_i)$ is continuous in ε_i, for any $\kappa_1 > 0$ we can choose κ small enough such that for all $i \in \mathcal{N}$

$$\Pr\left(\varepsilon_i \in \Omega_i \text{ s.t. } \left|\left(y_i\left(a^*(c), \varepsilon_i\right) - \overline{y}_i\left(a^*(c)\right)\right)/\overline{y}_i\left(a^*(c)\right)\right| \geq \kappa_1\right) \leq \kappa_1, \qquad (22)$$

whenever the size of the demand shocks is smaller than κ in the sense of (18). Inequality (22) says that the probability that firm i's output is different from the mean output by more than a factor of κ_1 is less than κ_1. Then we choose the output target to be slightly above the mean by letting

$$y_j^K(c) = \overline{y}_j\left(\alpha^*(c)\right)(1 + \kappa_1)$$

for each $j \in \mathcal{N}$ and $c \in C$. Hence, by the definition of $K_j(c)$ we have

$$\pi_j\left(\alpha^*(c), c_j, \varepsilon_j\right) < K_j(c) - 2\kappa_1 \Phi_1\left(c_j, \alpha_j^*(c)\right)\overline{y}_j\left(\alpha^*(c)\right)$$

if and only if

$$y_j\left(\alpha^*(c), \varepsilon_j\right) < \overline{y}_j\left(\alpha^*(c)\right)(1 - \kappa_1),$$

which by assumption occurs with probability less than or equal to κ_1.

We first show that $\lambda_{ij}\left(\alpha^*(c), y_j^K(c)\right) > 0$ when $\kappa_1 < 1/\eta$. We consider the case that $\overline{y}_j(a)$ is increasing in a_i. The decreasing case is similar. Since $\frac{\partial \overline{y}_j(\alpha^*(c))}{\partial a_i} > 0$

(see footnote 13), we have

$$
\begin{aligned}
\frac{\partial \widetilde{y}_j\left(\alpha^*(c), y_j^K(c)\right)}{\partial a_i} &= \frac{\partial \overline{y}_j\left(\alpha^*(c)\right)}{\partial a_i} - \int_{\varepsilon_j \notin \Gamma_j\left(y_j^K(c), \alpha^*(c)\right)} \frac{\partial y_j\left(\alpha^*(c), \varepsilon_j\right)}{\partial a_i} dF_j\left(\varepsilon_j\right) \\
&\geq \frac{\partial \overline{y}_j\left(\alpha^*(c)\right)}{\partial a_i} - \Pr\left(\varepsilon_j \notin \Gamma_j\left(y_j^K(c), \alpha^*(c)\right)\right) \eta \frac{\partial \overline{y}_j\left(\alpha^*(c)\right)}{\partial a_i} \\
&= \left(1 - \Pr\left(\varepsilon_j \notin \Gamma_j\left(y_j^K(c), \alpha^*(c)\right)\right) \eta\right) \frac{\partial \overline{y}_j\left(\alpha^*(c)\right)}{\partial a_i} \\
&\geq (1 - \kappa_1 \eta) \frac{\partial \overline{y}_j\left(\alpha^*(c)\right)}{\partial a_i},
\end{aligned}
$$

where the first inequality follows from condition (2) and the last inequality follows from the assumption that the size of the demand shocks is smaller than κ. Hence

$$
\lambda_{ij}\left(\alpha^*(c), y_j^K(c)\right) = \frac{\frac{\partial \widetilde{y}_j\left(\alpha^*(c), y_j^K(c)\right)}{\partial a_i}}{\frac{\partial \overline{y}_j\left(\alpha^*(c)\right)}{\partial a_i}} \geq 1 - \kappa_1 \eta > 0 \tag{23}
$$

if $\kappa_1 < 1/\eta$.

Next we show that the efficiency loss $|W|$ can be made arbitrarily small by choosing κ_1 small enough. Let

$$
\begin{aligned}
L_1 &= \max_{j,c} \Phi_1\left(c_j, \alpha_j^*(c)\right) \overline{y}_j\left(\alpha^*(c)\right), \\
L_2 &= \min_{j,c,\varepsilon_j}\left(\pi_j\left(\alpha^*(c), c_j, \varepsilon_j\right) - K_j(c)\right),
\end{aligned}
$$

and

$$
\begin{aligned}
E_1 &= \left\{\varepsilon_j \in \Omega_j | \pi_j\left(\alpha^*(c), c_j, \varepsilon_j\right) \geq K_j(c) - 2\kappa_1 \Phi_1\left(c_j, \alpha_j^*(c)\right) \overline{y}_j\left(\alpha^*(c)\right)\right\}, \\
E_2 &= \left\{\varepsilon_j \in \Omega_j | \pi_j\left(\alpha^*(c), c_j, \varepsilon_j\right) < K_j(c) - 2\kappa_1 \Phi_1\left(c_j, \alpha_j^*(c)\right) \overline{y}_j\left(\alpha^*(c)\right)\right\}.
\end{aligned}
$$

Then, we have for each firm j and each $c \in C$,

$$
\begin{aligned}
& \int_{\varepsilon_j} \min\left(\pi_j\left(\alpha^*(c), c_j, \varepsilon_j\right) - K_j(c), 0\right) dF_j\left(\varepsilon_j\right) \\
\geq\ & \int_{\varepsilon_j \in E_1} -2\kappa_1 L_1 dF_j\left(\varepsilon_j\right) + \int_{\varepsilon_j \in E_2} L_2 dF_j\left(\varepsilon_j\right) \\
\geq\ & -2\kappa_1 L_1 + \Pr\left(\varepsilon_j \in \Omega_j \text{ s.t. } y_j\left(\alpha^*(c), \varepsilon_j\right) < \overline{y}_j\left(\alpha^*(c)\right)(1 - \kappa_1)\right) L_2 \\
\geq\ & -2\kappa_1 L_1 + \kappa_1 L_2.
\end{aligned}
$$

From (23) we have for all $c \in C$

$$\frac{1}{\lambda_{ij}\left(\alpha^*(c), y_j^K(c)\right)} \leq \frac{1}{1-\kappa_1\eta}.$$

It follows that

$$\begin{aligned} W &= \sum_{i=1}^{n}\sum_{j\neq i}\int_c\int_{\varepsilon_j}\frac{\min\left(\pi_j\left(\alpha^*\left(c\right),c_j,\varepsilon_j\right)-K_j\left(c\right),0\right)}{\lambda_{ij}\left(\alpha^*(c),y_j^K(c)\right)}dF_j\left(\varepsilon_j\right)dG\left(c\right) \\ &\geq \frac{n(n-1)}{1-\eta\kappa_1}\left(-2\kappa_1 L_1+\kappa_1 L_2\right), \end{aligned}$$

which tends to zero as κ_1 tends to zero. Hence, we can choose κ_1 small enough such that $\lambda_{ij}\left(\alpha^*(c), y_j^K(c)\right) > 0$ and $|W| \leq \zeta$.

C Proof of Proposition 4

Firm i's ex ante average discounted profit in the non-collusive state of $\mathcal{S}\left(\beta,\mu\right)$ is

$$v_i^N \equiv \overline{\pi}_i\left(\alpha^{NE}\right).$$

Its ex ante average discounted profit in the collusive state is

$$v_i^* \equiv \overline{\pi}_i\left(\alpha^*\right)+\sum_{j\neq i}\left(\overline{\beta}_{ji}-\overline{\beta}_{ij}\right)+\frac{\delta\overline{\mu}\left(v_i^N-v_i^*\right)}{1-\delta}, \tag{24}$$

where

$$\overline{\beta}_{ij} \equiv \int_c\int_\varepsilon \beta_{ij}\left(c, y\left(\alpha^*\left(c\right),\varepsilon\right)\right)dF\left(\varepsilon\right)dG\left(c\right),$$

and

$$\overline{\mu} \equiv \int_c\int_\varepsilon \mu(c, y\left(\alpha^*\left(c\right),\varepsilon\right))dF\left(\varepsilon\right)dG\left(c\right).$$

Let

$$\overline{w}_i^* \equiv \int_c\int_\varepsilon w_i^*\left(c, y\left(\alpha^*\left(c\right),\varepsilon\right)\right)dF\left(\varepsilon\right)dG\left(c\right)$$

denote firm i's expected transfer under w^*. Since $\overline{\Pi}(\alpha^*) - \overline{\Pi}\left(\alpha^{NE}\right) > |W|$, we can pick a vector $d = (d_1, \ldots, d_n)$ with $\sum_{i=1}^n d_i = 0$ such that for each firm i

$$\overline{\pi}_i(\alpha^*) + \overline{w}_i^* + d_i > v_i^N. \tag{25}$$

It follows that there exists $\delta^* < 1$ such that for any $\widehat{c} \in C$ and $\widehat{y} \in \Re_+^n$

$$w_i^*(\widehat{c}, \widehat{y}) + d_i \geq \delta^* (1-\delta^*)^{-1} \left(v_i^N - (\overline{\pi}_i(\alpha^*) + \overline{w}_i^* + d_i)\right). \tag{26}$$

For any $\delta \geq \delta^*$, define μ and β as follows. First, for any $(\widehat{c}, \widehat{y}) \in C \times \Re_+^n$, set

$$\mu(\widehat{c}, \widehat{y}) \equiv \frac{(1-\delta)\delta^{-1} \sum_{i=1}^n w_i^*(\widehat{c}, \widehat{y})}{\overline{\Pi}(\alpha^{NE}) - \overline{\Pi}(\alpha^*) - W}. \tag{27}$$

Since $\overline{\Pi}(\alpha^*) + W > \overline{\Pi}\left(\alpha^{NE}\right)$, (26) implies $\mu(\widehat{c}, \widehat{y}) \in [0,1]$ for all $(\widehat{c}, \widehat{y}) \in C \times \Re_+^n$. Next, set

$$\beta_{ij}(\widehat{c}, \widehat{y}) \equiv \begin{cases} \beta_i^{net}(\widehat{c}, \widehat{y}) \frac{\min\left(\beta_j^{net}(\widehat{c}, \widehat{y}), 0\right)}{\sum_{k=1}^n \min\left(\beta_k^{net}(\widehat{c}, \widehat{y}), 0\right)} & \text{if } \beta_i^{net}(\widehat{c}, \widehat{y}) > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{28}$$

where

$$\beta_i^{net}(\widehat{c}, \widehat{y}) \equiv \delta(1-\delta)^{-1} \mu(\widehat{c}, \widehat{y}) \left(v_i^N - (\overline{\pi}_i(\alpha^*) + \overline{w}_i^* + d_i)\right) - (w_i^*(\widehat{c}, \widehat{y}) + d_i).$$

When $\beta_i^{net} > 0$ firm i pays each firm j with a negative β_j^{net} an amount in proportion to firm j's share of $\sum_{k=1}^n \min(\beta_k^{net}(\widehat{c}, \widehat{y}), 0)$. When $\beta_i^{net} \leq 0$, firm i makes no side-payments.[18] It is straightforward to check that for all $(\widehat{c}, \widehat{y}) \in C \times \Re_+^n$

$$\sum_{j \neq i} (\beta_{ji}(\widehat{c}, \widehat{y}) - \beta_{ij}(\widehat{c}, \widehat{y})) + \delta(1-\delta)^{-1} \mu(\widehat{c}, \widehat{y}) \left(v_i^N - (\overline{\pi}_i(\alpha^*) + \overline{w}_i^* + d_i)\right) = w_i^*(\widehat{c}, \widehat{y}) + d_i. \tag{29}$$

[18]When there are only two firms, this reduces to having the firm with a positive net payment pay the other firm the amount it owes.

We can think of the left-hand side of (29), firm i's continuation profit including side-payments at the end of a period in the collusive state, as the implicit transfer generated by the trigger strategy profile. By Proposition 1, conditional on firm i making the required side-payments, it would be optimal for it to report c_i and y_i truthfully and choose action $\alpha_i^*(\widehat{c})$ (assuming all firms $-i$ follow the trigger-strategy profile).

Hence, to show that the trigger-strategy profile is a perfect public equilibrium, we only need to show that it is optimal for firm i to make the side-payments. Substituting (29) into (24) and rearranging terms, we have

$$v_i^* = \overline{\pi}_i(\alpha^*) + \overline{w}_i^* + d_i. \tag{30}$$

Hence, by (25) $v_i^* > v_i^N$ for any firm i. Intuitively, we choose d_i to redistribute profits so that each firm is better off in the collusive state than in the non-collusive state.

Conditional on any $(\widehat{c}, \widehat{y}) \in C \times \Re_+^n$, firm i would receive a discounted profit

$$\sum_{j\neq i} (\beta_{ji}(\widehat{c}, \widehat{y}) - \beta_{ij}(\widehat{c}, \widehat{y})) + \delta \left(\mu(\widehat{c}, \widehat{y}) \frac{v_i^N - v_i^*}{1-\delta} + \frac{v_i^*}{1-\delta} \right) \tag{31}$$

in the continuation game if it makes the required side-payments and follows the trigger strategy thereafter. If it deviates, it would receive at most

$$\frac{\delta v_i^N}{1-\delta} + \sum_{j\neq i} \beta_{ji}(\widehat{c}, \widehat{y}). \tag{32}$$

It is therefore optimal for firm i to make the side-payments required if

$$-\sum_{j\neq i} \beta_{ij}(\widehat{c}, \widehat{y}) + \delta \left(\mu(\widehat{c}, \widehat{y}) \frac{v_i^N - v_i^*}{1-\delta} + \frac{v_i^*}{1-\delta} \right) \geq \frac{\delta v_i^N}{1-\delta}. \tag{33}$$

There are two cases to consider. When $\beta_i^{net}(\widehat{c}, \widehat{y}) \leq 0$, (33) is satisfied as $\sum_{j\neq i} \beta_{ij}(\widehat{c}, \widehat{y}) = 0$, $\mu(\widehat{c}, \widehat{y}) \geq 0$, and $v_i^* \geq v_i^N$. When $\beta_i^{net}(\widehat{c}, \widehat{y}) > 0$, it follows from (29) and (30) that the left-hand side of (33) is equal to $w_i^*(\widehat{c},\widehat{y}) + d_i + \delta\,(1-\delta)^{-1}\,v_i^*$, which, by (26) and (30), is greater than the right-hand side of (33).

Finally, by (30), $\sum_{i=1}^{n} v_i^* = \overline{\Pi}\,(\alpha^*) - |W|$.

Acknowledgements. We thank Joseph Harrington, Andrzej Skrzypacz, an associate editor, two anonymous referees, and participants at various seminars and conferences for helpful comments. This research is funded by the Chinese National Science Foundation (Project No. 71171125). Chan is supported by the Shanghai Dongfang Xuezhe Program.

References

Abreu, Dilip, Paul Milgrom, and David Pearce, 1991, Information and timing in repeated partnerships, *Econometrica* 59, 1713–1733.

Aliprantis, Charalambos D., and Owen Burkinshaw, 1990, *Principles of Real Analysis* (Academic Press) 2nd edn.

Aoyagi, Masaki, 2002, Collusion in dynamic Bertrand oligopoly with correlated private signals and communication, *Journal of Economic Theory* 102, 229 – 248.

——— , 2003, Bid rotation and collusion in repeated auctions, *Journal of Economic Theory* 112, 79–105.

Arrow, Kenneth Joseph, 1979, The property rights doctrine and demand revelation under incomplete information, in M Boskin, ed.: *Economics and Human Welfare* (Academic Press: New York).

Athey, Susan, and Kyle Bagwell, 2001, Optimal collusion with private information, *The RAND Journal of Economics* 32, 428–465.

——— , 2008, Collusion with persistent cost shocks, *Econometrica* 76, 493–540.

——— , and Chris Sanchirico, 2004, Collusion and price rigidity, *The Review of Economic Studies* 71, 317–349.

Athey, Susan, and David A. Miller, 2007, Efficiency in repeated trade with hidden valuations, *Theoretical Economics* 2.

Athey, Susan, and Ilya Segal, 2007, An efficient dynamic mechanism, Discussion paper.

——— , 2013, An efficient dynamic mechanism, *Econometrica* 81, 2463–2485.

Bergemann, Dirk, and Juuso Välimäki, 2010, The dynamic pivot mechanism, *Econometrica* 78, 771–789.

Blume, Andreas, and Paul Heidhues, 2006, Private monitoring in auctions, *Journal of Economic Theory* 131, 179–211.

Chan, Jimmy, and Wenzhang Zhang, forthcoming, Approximate efficiency in repeated games with side-payments and correlated signals, *Theoretical Economics.*

Clarke, Edward H, 1971, Multipart pricing of public goods, *Public Choice* 11, pp. 17–33.

Compte, Olivier, 1998, Communication in repeated games with imperfect private monitoring, *Econometrica* 66, 597–626.

Cramton, Peter C, and Thomas R Palfrey, 1990, Cartel enforcement with uncertainty about costs, *International Economic Review* 31, 17–47.

d'Aspremont, Claude, and Louis-Andre Gerard-Varet, 1979, Incentives and incomplete information, *Journal of Public Economics* 11, 25–45.

Escobar, Juan F., and Juuso Toikka, 2013, Efficiency in games with Markovian private information, *Econometrica* 81, 1887–1934.

Fudenberg, Drew, and David K. Levine, 2007, The Nash-threats folk theorem with communication and approximate common knowledge in two player games, *Journal of Economic Theory* 132, 461 – 473.

Fudenberg, Drew, and Jean Tirole, 1991, *Game Theory* (The MIT Press).

Green, Edward J., and Robert H. Porter, 1984, Noncooperative collusion under imperfect price information, *Econometrica* 52, 87–100.

Groves, Theodore, 1973, Incentives in teams, *Econometrica* 41, pp. 617–631.

Harrington, Joseph E., and Andrzej Skrzypacz, 2011, Private monitoring and communication in cartels: Explaining recent collusive practices, *American Economic Review* 101, 2425–49.

Harrington, Joseph E. Jr., and Andrzej Skrzypacz, 2007, Collusion under monitoring of sales, *The RAND Journal of Economics* 38, 314–331.

Hörner, Johannes, and Julian Jamison, 2007, Collusion with (almost) no information, *The RAND Journal of Economics* 38, 804–822.

Hörner, Johannes, Satoru Takahashi, and Nicolas Vieille, 2013, Truthful Equilibria in Dynamic Bayesian Games, Working paper, Yale University.

Kandori, Michihiro, and Hitoshi Matsushima, 1998, Private observation, communication and collusion, *Econometrica* 66, 627–652.

Kandori, Michihiro, and Ichiro Obara, 2006, Efficiency in repeated games revisited: The role of private strategies, *Econometrica* 74, 499–519.

Mezzetti, Claudio, 2004, Mechanism design with interdependent valuations: Efficiency, *Econometrica* 72, 1617–1626.

Miller, David A., 2012, Robust collusion with private information, *The Review of Economic Studies* 79, 778–811.

Obara, Ichiro, 2009, Folk theorem with communication, *Journal of Economic Theory* 144, 120 – 134.

Radner, Roy, 1983, Repeated principal-agent games with discounting, *Econometrica* 53, 1173–1198.

Rubinstein, Ariel, 1979, An optimal policy for offenses that may have been committed by accident, in S. Brams, A. Schotter, and G. Schwodiauer, ed.: *Applied Game Theory* (Wurzberg: Physica-Verlag).

——— , and Menahem E. Yaari, 1983, Repeated insurance contracts and moral hazard, *Journal of Economic Theory* 30, 74–97.

Skrzypacz, Andrzej, and Hugo Hopenhayn, 2004, Tacit collusion in repeated auctions, *Journal of Economic Theory* 114, 153–169.

Suslow, Valerie Y., 2005, Cartel contract duration: Empirical evidence from inter-war international cartels, *Industrial and Corporate Change* 14, 705 – 744.

Vickrey, William, 1961, Counterspeculation, auctions, and competitive sealed tenders, *The Journal of Finance* 16, pp. 8–37.

Zheng, Bingyong, 2008, Approximate efficiency in repeated games with correlated private signals, *Games and Economic Behavior* 63, 406 – 416.