RAND Journal of Economics
Vol. 51, No. 2, Summer 2020
pp. 421–446

Information exchange in cartels

Yu Awaya*

and

Vijay Krishna**

Antitrust authorities view the exchange of information among firms regarding costs, prices, or sales as anticompetitive. Such exchanges allow competitors to closely monitor each other, thereby facilitating collusion. But the exchange of aggregate *information, perhaps via a third party, is legal. The logic is that collusion is difficult if the* identity *of a price-cutting firm cannot be ascertained. Here, we examine this logic using Stigler's model of secret price cuts. We first identify circumstances such that when no information exchange is possible, collusion is difficult. We then show that if firms' aggregate sales are made public, nearly perfect collusion is possible.*

1. Introduction

■ Antitrust authorities in most countries try to restrict the exchange of information among competitors. In particular, the exchange of firm-specific data regarding costs, prices, and sales is viewed, at the very least, with suspicion. But the sharing of *aggregate* data, perhaps via a trade association or some other third party, is usually deemed to be legal. According to guidelines of the European Commission:

> "Exchanges of genuinely aggregated data, that is to say, where the recognition of individualised company level information is sufficiently difficult, are much less likely to lead to restrictive effects on competition than exchanges of company level data." (European Commission, 2011)

In the same vein, the US Federal Trade Commission suggests that competitors establish a "safety zone" in which "the shared statistics are sufficiently aggregated that no participant can discern the data of any other participant" (FTC, 2014). The Japan Fair Trade Commission too

* University of Rochester; yuawaya@gmail.com.
** Penn State University; vkrishna@psu.edu.
The research reported here was supported by a grant from the National Science Foundation (SES-1626783). Awaya and Krishna thank Reiko Aoki, Kentaro Inomata, Jun Nakabayashi, Hiroyuki Odagiri, and Yosuke Okada for helpful comments, especially concerning antitrust practice in Japan. We also thank Yossi Spiegel for his comments and the referees and the editor for their suggestions.

has ruled that the exchange of information "without clearly indicating the quantities, amounts, etc. of the individual constituent enterprises … " is permissible (JFTC, 2015).[1]

What is the logic of these guidelines? Individual data allow firms to monitor each others' activities better than aggregates and so restrictions on the exchange of individual data limit the possibility of coordination and collusion. Economic theory also seems to confirm this view. For instance, Radner, Myerson, and Maskin (1986) show that in a repeated partnership game, imperfect public monitoring—where each participant observes only an aggregate—severely restricts the prospects of cooperation relative to a situation in which individual choices are perfectly observed. When only an aggregate is observed, individual transgressors cannot be identified. In technical terms, the "pairwise identifiability" condition of Fudenberg, Levine, and Maskin (1994) fails.[2] Leading antitrust scholars, perhaps influenced by the theory, have also claimed that

> "[…] aggregating the data largely removes the value of information in facilitating collusion." (Carlton, Gertner, and Rosenfield, 1997)

This seemingly compelling logic notwithstanding, many cartels seem to operate quite successfully on the basis of aggregate information alone. Members of the copper plumbing tubes cartel

> "[…] set up a new data exchange system—initially on a monthly, later on a quarterly, basis through the World Bureau of Metal Statistics. WBMS statistics only contained aggregated figures and no company specific information." (European Commission, 2004)[3]

In this article, we study a repeated oligopoly with private monitoring in the style of Stigler (1964). Firms produce differentiated products and compete in prices. Firms' sales are determined by all the prices but are subject to individual and market shocks. In the baseline model, firms cannot observe the prices set by others and so secret price cuts are possible. Rather, firms observe only their own sales and, as in Stigler (1964), when sales are rather noisy, collusion is hard to sustain because of the difficulties firms face in monitoring each others' activities. We then consider an alternative situation in which a third party—say, a trade association—publicly reveals aggregate data on sales to all the participants. We assume that the sales data are accurate, perhaps having been verified by a sufficiently accurate audit.[4] We identify circumstances in which without any information exchange, industry profits in any equilibrium are low, whereas with the public availability of the aggregate, there exists an equilibrium in which firms' profits are close to those in a monopoly. This result suggests that the exchange of information among firms can be harmful even when it is in aggregate form and the data of any individual participant cannot be discerned.[5] Moreover, the potential increase in profits coming from the additional information available to the firms can be quite large.

How can firms collude using only aggregate data? The key is that whereas individual sales can be rather variable, aggregate sales are usually much less variable. Thus, if a firm cuts its price from an agreed upon level, this will cause, with high probability, an increase in aggregate sales

[1] In a case involving the titanium sponge industry, the JFTC did not allow the exchange of even aggregate information. The reason was that there were only *two* firms in the industry and so aggregate information could be perfectly disaggregated (JFTC, 1999).

[2] This article considers finite games. See Matsushima (1989) for a similar result that applies to games with continuous actions (as in the current article).

[3] Prior to this, the cartel engaged in the direct exchange of sales figures. Thus the cartel chose to move to a system where only aggregate information was exchanged. See Sugaya and Wolitzky (2018) for a rationalization of this change.

[4] In the well-studied lysine cartel case, AC-Treuhand, a Swiss consulting firm, actually audited sales figures (Marshall and Marx, 2008). The sorbates cartel verified reported sales by cross-checking against export figures reported to the government.

[5] Formally, as in Radner, Myerson, and Maskin (1986), the pairwise identifiability condition fails in our model.

observed by all. Even though the identity of the price cutter is not revealed, firms can use the aggregate as a signal to go into a punishment mode. This can be sufficient to deter price cuts. Moreover, the fact that the variability of the aggregate is low means that there is only a small chance of triggering a punishment if no one has cheated. Put another way, aggregate sales are sensitive to price cuts but relatively insensitive to demand shocks. In the absence of aggregate sales data, however, firms would have to rely only on their individual sales to detect such price cuts. This would lead to large type 1 and/or type 2 errors—not punishing when warranted and punishing when not—thereby, restricting equilibrium profits.

The key to our result is showing that *small* price cuts can be adequately deterred. If monitoring were perfect, then even the smallest price cut can, of course, be detected with probability one and punished severely. But with imperfect monitoring, a small price cut has only a negligible statistical effect on the aggregate and so only a small probability of triggering a punishment. Our analysis reveals, nevertheless, that if the variability of the aggregate is small yet positive, the trade-off is not worthwhile—the small potential gains from a small price cut are not enough to outweigh the small potential losses from being punished.

□ **Related literature.** Genesove and Mullin (2001) present a fascinating account of the internal functioning of a sugar-refining cartel from the 1920s. Much of our knowledge of how more recent cartels operate comes from reports of the European Commission which prosecuted over 20 industrial cartels in the 1990s—ranging from amino acids to zinc phosphate. Much of this material has been conveniently surveyed by Harrington (2006) and Marshall and Marx (2012). Both have numerous accounts of how cartels use aggregate information disseminated via trade associations or statistical bureaus. According to Marshall and Marx (2012), in at least 11 of the 22 European cases, collusion was facilitated by third parties.

Our basic model is that of a repeated game with *private* monitoring in which firms set price and observe only their own sales. When sales are noisy, this channel is a poor method of monitoring and so profits are limited. This, of course, was suggested by Stigler (1964) and we use the methodology developed in our earlier work, Awaya and Krishna (2019), to quantify Stigler's argument.

When aggregate sales information is available, the situation changes to one with *public* monitoring—all firms observe aggregate sales. There has been extensive interest in the question of when cooperation is or is not possible when monitoring is public. Radner, Myerson, and Maskin (1986) showed in an example that if deviators could not be identified via the public signals, then cooperation is impossible even when firms are arbitrarily patient. Identifiability plays a key role in the work of Fudenberg, Levine, and Maskin (1994), who showed that when public signals satisfy a pairwise identifiability condition, a "folk theorem" emerges. In finite games, identifiability requires that the number of public signals is large relative to the number of actions. In our model, there is a continuum of signals (aggregate sales) and actions (prices) and an analog of the identifiability condition due to Matsushima (1989) is not satisfied. To see this simply, consider a situation in which the public signal is just the total industry sales. When industry sales are high, it is likely that someone deviated but the identity of the deviator cannot be deduced. Nevertheless, our main result is that collusive equilibria are possible even based only on nonidentifiable signals. The reason is that, unlike earlier work, our result is not in the nature of a "folk theorem" in which the signal structure is held fixed and firms' discount factor is varied. We carry out a different exercise—the discount rate is held fixed and we ask what happens as the public signals become less noisy.

Formally, our results concern equilibria with two different monitoring structures. In the base case, firms rely only on their own sales to ascertain the situation and we compare this to a situation in which, in addition, aggregate sales data is available. Kandori (1992) has shown that in games with imperfect *public* monitoring, more informative signals (in the sense of Blackwell) weakly enlarge the set of public perfect equilibria. In our model, the base case is one of *private* monitoring and so Kandori's result cannot be used.

Marshall and Marx (2008) study how aggregate data can aid collusion but in a model with a *homogenous* product and large demand shocks. In such a setting, a price cut by a firm results in zero sales of other firms. Aggregate sales data is useful because if a firm's sales are zero whereas aggregate sales are positive, then it knows that a rival cut its price. On the other hand, if there is a large negative demand shock and aggregate sales are zero as well, then firms presume that no one deviated. This means that the chances that the cartel will break down even if no one has deviated are zero—there is no type 1 error.

In a recent article, Sugaya and Wolitzky (2018) also study the role of aggregate information in cartels where each firm is a local monopoly. Market-specific sales information may reveal to rivals that a particular market is attractive for entry. But such a deviation would not be possible if only aggregate information were available. For this reason, the exchange of aggregate information may be better for the cartel. In this model, the monitoring problem is trivial because if one firm enters another firm's market, this is detected immediately. In our model, the monitoring problem is paramount. Moreover, whereas we are comparing a situation with no information exchange to one where aggregate information is exchanged, Sugaya and Wolitzky (2018) compare a situation where only aggregate information is exchanged to one where firm-specific information is exchanged.

There are a number of articles that have looked at how *firm-specific* sales information can aid collusion. This line of research is pursued in Aoyagi (2002), Harrington and Skrzypacz (2011), and our earlier work, Awaya and Krishna (2016). Aryal, Ciliberto, and Leyden (2018) have shown how such communication affects price movements in the airline industry. In these models, sales information cannot be verified but equilibrium strategies guarantee that all firms have the incentive to report truthfully. In this article, we show that when only aggregate information is made public, it is difficult to achieve truthful reporting of sales. The sharing of firm-specific cost information is studied by Athey and Bagwell (2001).

We study the question of whether aggregate information facilitates collusion in a model with private monitoring, but the same question could be posed in the context of Green and Porter's (1984) model of *public* monitoring as well. In the Green–Porter model, firms choose quantities of a homogeneous product and the single market price is then determined by the total quantity placed on the market as well as by shocks to market demand. Thus, although the price is stochastic, total sales are not—these are completely determined by the choices of the firms. In the Green–Porter model, imperfect monitoring results from the fact that firms observe only the stochastic price and not total sales. If total sales figures were to be publicly announced, it would be rather simple to detect a deviation—firms would see an increase in total sales, whereas the deviator's identity would still be hidden, a strategy that called for firms to react by playing the one-shot equilibrium would be enough to deter any cheating.

The remainder of this article is organized as follows. Section 2 outlines the basic model and formally defines the two repeated games we compare—one without aggregate information and one with. Section 3 then shows that without aggregate information firms have great difficulty in colluding when sales are rather noisy. The next section then identifies circumstances in which the dissemination of aggregate sales information, say by a trade association, facilitates collusion. Section 5 combines the results of the two previous sections to formally state our main finding. The model in Section 4 assumes that the aggregate information is accurate and in Section 6 we ask whether the firms in the cartel have the incentive to report their sales truthfully to the trade association. Section 7 concludes.

2. The market

■ Consider a market with $n \geq 2$ symmetric firms. The firms produce differentiated products at a constant cost, which we set to zero without loss of generality. The products are imperfect substitutes and firms compete in prices. Each firm i sets a price $p_i \in [0, p^{\max}]$, for its product and its sales, denoted by Y_i, are stochastically determined by the prices set by all firms $\mathbf{p} = (p_1, p_2, \ldots, p_n)$.[6] We assume that given the prices $\mathbf{p}$, firms' sales $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ are jointly distributed according to a multivariate *log-normal* density $f(\cdot \mid \mathbf{p})$, or equivalently, firms' log sales $\ln \mathbf{Y} = (\ln Y_1, \ln Y_2, \ldots, \ln Y_n)$ are normally distributed. Specifically, log-sales are jointly distributed according to a multivariate normal distribution of the form $\mathcal{N}(\mu(\mathbf{p}), \Sigma)$ where $\mu(\mathbf{p}) = E[\ln \mathbf{Y} \mid \mathbf{p}]$ and Σ is the variance–covariance matrix. Notice that although the prices $\mathbf{p}$ affect the expected log-sales $\mu(\mathbf{p})$, they do not affect Σ.

The expected sales of firm i's product are

$$E[Y_i \mid \mathbf{p}] = q_i(p_i, \mathbf{p}_{-i})$$

where the function q_i is decreasing in p_i and increasing in p_j for $j \neq i$. Furthermore, the function q_i is *symmetric* in the last $n-1$ components—that is, the prices of other firms can be interchanged without affecting i's expected sales.

The variance–covariance matrix is assumed to be of the form

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & & \rho \\ \vdots & & \ddots & \\ \rho & \rho & & 1 \end{bmatrix}$$

where σ^2 is the variance of individual log sales and $\rho < 0$ is the pairwise correlation coefficient. Firms' log sales are thus *negatively correlated* and the requirement that Σ be positive definite is equivalent to the condition that $\rho > -\frac{1}{n-1}$. The assumption that log sales are negatively correlated is equivalent to assuming that idiosyncratic demand shocks are larger in magnitude than industry-wide demand shocks. In other words, for a given price vector $\mathbf{p}$, market demand is relatively stable.[7]

Note that the expected sales of a firm can be written as

$$E[Y_i \mid \mathbf{p}] = \exp\left(\mu_i(\mathbf{p}) + \tfrac{1}{2}\sigma^2\right)$$

Firms maximize their *expected* profits:

$$\pi_i(p_i, \mathbf{p}_{-i}) = p_i q_i(p_i, \mathbf{p}_{-i})$$

and we suppose that π_i is *strictly concave* in p_i. In addition, we require

Condition 1. For all i, q_i is log-concave in p_j, $j \neq i$.

Condition 2.

$$\sum_{j=1}^{n} e_{ji} < 0$$

where $e_{ji} = \frac{\partial \ln q_j}{\partial \ln p_i}$ denotes the elasticity of j's expected demand with respect to i's price.

Both conditions are restrictions on the shape of demand. Condition 2 is just the elasticity version of the condition that the "own-price effect" is greater than the sum of the "cross-price effects." When demand is linear, that is,

$$q_i(p_i, \mathbf{p}_{-i}) = \alpha - \beta p_i + \frac{1}{n-1} \sum_{j \neq i} p_j \tag{1}$$

[6] Throughout, bold face characters denote vectors. As usual, if $\mathbf{p}$ is a vector of prices, we write $(p_i', \mathbf{p}_{-i})$ to denote the vector where the ith component of $\mathbf{p}$ has been replaced by p_i'.

[7] Harrington and Skrzypacz (2011) also study a model where firms' sales are negatively correlated. Indeed, their model total sales, although stochastic, are independent of prices.

Condition 1 is satisfied, of course, and Condition 2 holds if β is large enough.

□ **Stage game.** Let G denote the one-shot game where the firms choose prices p_i and the resulting expected profits are $\pi_i(p_i, \mathbf{p}_{-i})$.

We will assume that $\mathbf{p}^M = (p^M, p^M, \ldots, p^M)$ is the *unique* solution to the monopolist's problem:

$$\max_{\mathbf{p}} \sum_i \pi_i(\mathbf{p})$$

and let π^M be the resulting profits per firm. For technical reasons, we will also assume that a firm's expected sales are bounded away from zero. Of course, $p^{\max}$ is assumed to be no smaller than p^M. Let p^B be a firm's best response to a situation in which all the other firms set the monopoly price p^M, that is,

$$p^B = \arg\max_{p_1} \pi_1\left(p_1, \mathbf{p}^M_{-1}\right) \tag{2}$$

Our conditions guarantee that there exists a symmetric Nash equilibrium price p^N of G and let π^N be the resulting profits of a firm. [8] We assume that monopoly pricing does not constitute a Nash equilibrium, a property of most reasonable models of competition. We will also make use of the following notion introduced by Moulin and Vial (1978): a joint distribution χ over the prices of all firms is a *coarse correlated equilibrium* (*CCE*) if for all i and all p_i,

$$E_\chi[\pi_i(\mathbf{P})] \geq E_{\chi_{-i}}[\pi_i(p_i, \mathbf{P}_{-i})]$$

where $\mathbf{P}$ is the vector of random variables denoting prices and all expectations are taken with respect to the distribution χ. Every Nash equilibrium is, of course, a CCE with zero correlation—that is, one in which the distribution χ is a product of its marginals (the mixed strategies of firms).[9] Correlation in the choices of firms will be important when we consider the repeated version of G because firms' prices will become correlated via the information they receive.

For $\varepsilon > 0$, an ε-*coarse correlated equilibrium* (ε-*CCE*) is a distribution χ such that for all i and all p_i,

$$E_\chi[\pi_i(\mathbf{P})] \geq E_{\chi_{-i}}[\pi_i(p_i, \mathbf{P}_{-i})] - \varepsilon$$

□ **Repeated game.** We will study a repeated game G_δ where firms use the discount factor $\delta < 1$ to discount future profits. In each period t, firms choose prices $\mathbf{p}^t$ and given these prices, the sales $\mathbf{Y}^t$ of all the firms are realized as described above. Following Stigler (1964), each firm i observes *only* its own realized sales y_i^t ; it observes neither the prices set by other firms nor their sales.

Let $(p_i^1, y_i^1, p_i^2, y_i^2, \ldots, p_i^{t-1}, y_i^{t-1})$ denote the *private* history of firm i after $t-1$ periods. A *strategy* s_i for firm i is a collection of functions $(s_i^1, s_i^2, \ldots)$ such that in period t, s_i^t determines, for each private history, a distribution over the set of prices $[0, p^{\max}]$. Note that the possibility that firms may randomize is permitted. A *strategy profile* $\mathbf{s}$ is simply an n-tuple of strategies $(s_1, s_2, \ldots, s_n)$. A *Nash* equilibrium of G_δ is strategy profile $\mathbf{s}$ such that for each i, the strategy s_i is a best response to $\mathbf{s}_{-i}$.

□ **Repeated game with aggregate information.** We will compare equilibria of the repeated game G_δ with those of an alternative game G_δ^A in which, in addition to their own sales, firms

[8] If the one-shot game has multiple symmetric Nash equilibria, let p^N denote the Nash equilibrium price with the lowest equilibrium profits.

[9] More generally, every correlated equilibrium is also a CCE (and not necessarily vice versa).

also receive a public signal A which is an *aggregate* of industry sales. In what follows, we will suppose that

$$\ln A = \frac{1}{n}\sum_{j=1}^{n} \ln Y_j$$

that is, $\ln A$ is the arithmetic mean of log sales, or equivalently, A is the geometric mean of sales.[10]

When such aggregate information is available, a firm's strategy can depend not only on its private history of prices and sales but also on the realizations of all past aggregates $(a^1, a^2, \dots, a^{t-1})$. A special class of strategies are those in which firms' choices depend *only* on the history of aggregates $(a^1, a^2, \dots, a^{t-1})$ and as these are publicly known, such strategies are called *public strategies*.

The particular aggregate chosen here—mean of log sales—is particularly convenient because, at prices $\mathbf{p}$, if $(\ln Y_1, \ln Y_2, \dots, \ln Y_n)$ are jointly distributed according to the multivariate normal $\mathcal{N}(\mu(\mathbf{p}), \Sigma)$ then the aggregate $\ln A$ is distributed according to a *univariate* normal $\mathcal{N}(\mu(\mathbf{p}), \theta^2)$ where the expectation of $\ln A$ is

$$\mu(\mathbf{p}) = \frac{1}{n}\sum_{j=1}^{n} \mu_j(\mathbf{p}) \tag{3}$$

that is, the average of the firms' expected log sales, and the variance of $\ln A$ is

$$\theta^2 = \frac{1}{n}\sigma^2(1 + (n-1)\rho) \tag{4}$$

Later we will make use of the fact that for fixed σ^2, as $\rho \to -\frac{1}{n-1}$, $\theta^2 \to 0$.

3. Bound with no information exchange

■ In this section, we use the methodology developed in our earlier work, Awaya and Krishna (2019), to bound the equilibrium profits in the repeated game G_δ in which no information is exchanged and firms observe only their own sales. The next section will then show how this bound can be overcome once aggregate sales information is available.

Specifically, in Awaya and Krishna (2019) we showed that the set of Nash equilibrium (NE) payoffs of the repeated game G_δ is contained in the set of ε-CCE of the stage game G (see Section 2 for a definition) for a value of ε that depends on the discount factor δ, the quality of monitoring η and the payoffs in G. Precisely,

$$NE(G_\delta) \subseteq \varepsilon\text{-}CCE(G) \tag{5}$$

where $\varepsilon = \frac{\delta^2}{1-\delta}\eta\|\pi\|_\infty$.

The key parameter determining ε is η—the quality of monitoring. We measure this as follows. Consider two price vectors $\mathbf{p}$ and $\mathbf{p}'$ and the resulting joint distributions of the sales of firms other than i, given by the multivariate log-normal densities $f_{-i}(\cdot \mid \mathbf{p})$ and $f_{-i}(\cdot \mid \mathbf{p}')$. If these two distributions are close together, then it will be difficult for firms $j \neq i$ to detect the change from $\mathbf{p}$ to $\mathbf{p}'$. Thus, the quality of monitoring can be measured by the maximum "distance" between any two such distributions. As in Awaya and Krishna (2019), we use the so-called total variation metric to measure this distance. Because f is symmetric, the quality of monitoring is the same for all firms.

[10] A model where firms' sales are themselves normally distributed and the aggregate is just the sum of these is virtually isomorphic to the one studied here. The drawback of this alternative model, not present in the log-normal model, is that sales would be negative with positive probability.

Definition 1. The *quality* of monitoring of f is defined as

$$\eta = \max_{\mathbf{p},\mathbf{p}'} \left\| f_{-i}(\cdot \mid \mathbf{p}) - f_{-i}(\cdot \mid \mathbf{p}') \right\|_{TV}$$

where f_{-i} is the marginal of f on Y_{-i} and $\|f - g\|_{TV}$ denotes the *total variation* distance between densities f and g.[11]

We now argue that as $\sigma^2 \to \infty$, $\eta \to 0$. Because the total variation distance is unaffected by monotonic transformations, we also have

$$\eta = \max_{\mathbf{p},\mathbf{p}'} \left\| \mathcal{N}(\mu_{-i}, \Sigma_{-i}) - \mathcal{N}\big(\mu'_{-i}, \Sigma_{-i}\big) \right\|_{TV}$$

where $\mu_{-i} = \mu_{-i}(\mathbf{p})$ and $\mu'_{-i} = \mu_{-i}(\mathbf{p}')$ and Σ_{-i} is the $(n-1) \times (n-1)$ submatrix of Σ obtained by deleting the ith row and ith column of Σ.

Although the total variation distance between two multivariate normal distributions cannot be written in closed form, a useful upper bound to the total variation distance between two distributions f and g is given by

$$\|f - g\|_{TV} \leq \sqrt{2} H(f, g)$$

where H denotes that Hellinger distance between the distributions. For normal distributions with the same variance–covariance matrix, an explicit formula is

$$H\big(\mathcal{N}(\mu_{-i}, \Sigma_{-i}), \mathcal{N}\big(\mu'_{-i}, \Sigma_{-i}\big)\big) = \sqrt{1 - \exp\left(-\tfrac{1}{8}(\Delta\mu_{-i})^T (\Sigma_{-i})^{-1} \Delta\mu_{-i}\right)}$$

where $\Delta\mu_{-i} = \mu_{-i} - \mu'_{-i}$ is the difference in means. The determinant of the $(n-1) \times (n-1)$ matrix Σ_{-i} is

$$\det \Sigma_{-i} = \sigma^2 (1 - \rho)^{n-2} (1 + (n-2)\rho)$$

which is an increasing function of ρ over the interval $(-\frac{1}{n-1}, 0)$. Thus, for any $\rho > -\frac{1}{n-1}$,

$$\det \Sigma_{-i} > \sigma^2 \left(\frac{n}{n-1}\right)^{n-2} \left(\frac{1}{n-1}\right)$$

Thus, as $\sigma^2 \to \infty$, $\det \Sigma_{-i} \to \infty$ and hence the quadratic form in the expression for H goes to zero. This implies for any $\overline{\eta}$, there exists a $\underline{\sigma}^2$ such that for all $\sigma^2 > \underline{\sigma}^2$, $\eta < \overline{\eta}$ for all $\rho \in (-\frac{1}{n-1}, 0)$. Note that $\underline{\sigma}^2$ can be chosen independently of ρ.

It is easy to see that monopoly profits cannot be attained in any CCE of G. This is because $\mathbf{p}^M$ is the *unique* maximizer of joint profits and so the only way to attain these profits in a *CCE* would be to put probability one on this price vector. But then a firm has a profitable deviation as monopoly prices do not constitute a Nash equilibrium of G. This result then implies,

Proposition 1. For σ^2 large enough (and hence η small enough), the profits in any equilibrium of G_δ are bounded away from monopoly profits.

□ **Linear demand.** Proposition 1 establishes that when sales are quite variable, profits in any equilibrium of the repeated game without information exchange G_δ are bounded away from monopoly profits. This relies on the fact that for the one-shot price-setting game G we study, the set of CCE is itself bounded away from monopoly profits. We now show that when (expected) demands q_i are linear in prices, no CCE can result in greater profits than those in the unique Nash equilibrium of the game. Thus with linear demand, the gap between monopoly profits and the set of CCE profits is as large as possible. For the case of a duopoly, this is implied by a result of

[11] The total variation distance between two densities f and g on X is defined as $\|f - g\|_{TV} = \frac{1}{2}\int_X |f(x) - g(x)| dx$.

Gérard-Varet and Moulin (1978).[12] The following result extends this to any number of firms and, more important for our purposes, generalizes to include ε-*CCE* (allowing for the possibility of $\varepsilon = 0$ as a special case).

Lemma 1. Suppose that for all i,

$$q_i(p_i, \mathbf{p}_{-i}) = \alpha - \beta p_i + \frac{1}{n-1}\sum_{j \neq i} p_j$$

where $\beta > 1$. Then the ε-CCE of the one-shot price-setting game G that maximizes industry profits is a pure strategy ε-Nash equilibrium.

Proof. The set of ε-CCE distributions is, by definition, convex. Because the price game is symmetric, this means that there is a *symmetric* ε-*CCE* distribution χ that maximizes industry profits.

The expected profits of firm i in this ε-*CCE* is

$$E[\pi_i(\mathbf{P})] = E\left[P_i\left(\alpha - \beta P_i + \tfrac{1}{n-1}\sum_{j \neq i} P_j\right)\right] \tag{6}$$

where the expectation is taken over the random prices $\mathbf{P} = (P_1, P_2, \ldots, P_n)$ with respect to χ. Now as χ is symmetric, routine calculations show that

$$E[\pi_i(\mathbf{P})] = E[P](\alpha - \beta E[P] + E[P]) - (\beta - Corr)Var[P]$$

where $E[P]$ is the common expectation of all the P_i, $Var[P]$ is the common variance and $Corr$ is the common pairwise correlation coefficient between any two P_i and P_j. This can be rewritten as

$$E[\pi_i(\mathbf{P})] = \pi_i(E[\mathbf{P}]) - (\beta - Corr)Var[P]$$

that is, the profits at the expected prices less a term that depends only on the correlation and variance of the distribution χ.

The symmetric ε-*CCE* that maximizes industry profits is the solution to the problem: choose $E[P]$, $Var[P]$, and $Corr$ to

$$\max \pi_i(E[\mathbf{P}]) - (\beta - Corr)Var[P]$$

subject to: for all i,

$$\pi_i(E[\mathbf{P}]) - (\beta - Corr)Var[P] + \varepsilon \geq \max_{p_i} \pi_i(p_i, E[\mathbf{P}_{-i}])$$

Because $\beta > 1 \geq Corr$, any feasible, symmetric χ with $Var[P] > 0$ is inferior to a degenerate distribution on $E[\mathbf{P}]$ with $Var[P] = 0$. This is because the degenerate distribution would still satisfy the constraints and lead to a higher value for the objective function. But any degenerate distribution that is feasible is a pure-strategy ε-Nash equilibrium. □

Using Lemma 1 we obtain the following strengthening of Proposition 1.

Corollary 1. Suppose demand is linear and $\beta > 1$. Total industry profits in any equilibrium of G_δ approach the total profits in the unique Nash equilibrium of the one-shot game G as σ^2 becomes large.

[12] Gérard-Varet and Moulin (1978) have also shown that with nonlinear demand, there may be CCE with profits that exceed those in any one-shot Nash equilibrium. Moulin, Ray, and Sen Gupta (2014) exhibit other examples of this in two-person games with quadratic payoffs.

4. Equilibrium with aggregate sales information

■ Having shown that firms' profits are bounded away from monopoly profits when no information is exchanged, we now ask whether aggregate information can, in fact, help firms collude to near-monopoly levels. In this section, we will establish that there are circumstances—pertaining to the correlation in sales—such that aggregate information can facilitate collusion even though it does not identify transgressors. Indeed, we will show that near-perfect collusion is possible using rather simple "trigger strategies" in which firms start by charging the monopoly price and permanently revert to punish via the one-shot Nash equilibrium if the aggregate of sales ever exceeds a predetermined threshold. We emphasize that our result is not a "folk theorem" in which the monitoring structure (the stochastic relationship between prices and sales) is held fixed and the question is whether it is possible for firms to collude as the discount factor goes to one. Indeed because the information available to firms is in aggregate form, the identifiability problem does not allow a high degree of collusion even for high discount factors. Our analysis proceeds along a different path. We hold the discount factor to be fixed and then instead change the monitoring structure. Of course, the discount factor cannot be arbitrarily small. At the very least, it should be high enough so that collusion is possible with perfect monitoring and we first determine this minimum level.

□ **Perfect monitoring benchmark.** Consider an artificial situation with *perfect monitoring* in which firms can observe each other's actions—the prices set by all firms—perfectly. In this case, a simple "grim trigger" strategy—charge p^M and continue to do so if all firms have charged p^M in all past periods; otherwise, charge p^N—suffices to sustain monopoly prices as long as δ is high enough. Specifically, with perfect monitoring, the grim trigger strategy constitutes an equilibrium as long as

$$\pi^M \geq (1-\delta)\pi_1\left(p^B\right) + \delta\pi^N$$

where p^B is a firm's stage game *best response* to all other firms setting the monopoly price p^M and to save on notation, we write $\pi_1(p^B) = \pi_1(p^B, \mathbf{p}^M_{-1})$. The inequality above is equivalent to

$$\delta \geq \underline{\delta} \equiv \frac{\pi_1(p^B) - \pi^M}{\pi_1(p^B) - \pi^N} \tag{7}$$

We will refer to $\underline{\delta}$ as the perfect monitoring benchmark.

□ **A collusive equilibrium.** The main result of this section is

Proposition 2. Fix $\delta > \underline{\delta}$. When the variance of the aggregate, θ^2, is small enough, there exists a collusive equilibrium of the repeated game G^A_δ with aggregate sales information. As θ^2 approaches zero, the profits from these equilibria approach monopoly levels.

To establish the proposition, we will propose an explicit strategy for the firms. As mentioned above, firms' behavior will be dictated by a trigger value for the aggregate. We will then show that the proposed strategies constitute an equilibrium as θ^2 becomes small, or equivalently, for fixed σ^2, ρ approaches $-\frac{1}{n-1}$. Moreover, the equilibrium profits will approach monopoly levels when the variability of total sales is small.[13]

□ **Strategies.** Consider the following analog of the "grim trigger" strategy with a publicly known trigger—a specific value a for the aggregate A (the exact value of the trigger a will be determined later):

[13] This is in contrast to the model of Harrington and Skrzypacz (2011) where total sales vary but independently of prices.

- In period 1, set the monopoly price p^M.
- In any period $t > 1$, set the monopoly price if and only if in all past periods $\tau < t$ the aggregate $A^\tau \leq a$. Otherwise, set the Nash price p^N.

We will now show that for an appropriate choice of the trigger, a, the strategies given above constitute an equilibrium.

Suppose all firms follow the suggested strategy and for all $\tau < t - 1$ the condition $A^\tau \leq a$ has been satisfied so that in period $t - 1$ all firms set p^M. If all firms charge p^M, then the aggregate $\ln A$ is normally distributed with mean $\mu^M \equiv \mu(\mathbf{p}^M)$ and variance θ^2. The probability that the firms will then set p^M in period t is just

$$\begin{aligned}\Pr\left[A^{t-1} \leq a\right] &= \Pr\left[\ln A^{t-1} \leq \ln a\right] \\ &= \Phi\left(\frac{\ln a - \mu^M}{\theta}\right)\end{aligned}$$

where Φ is the cumulative distribution function of the standard normal $\mathcal{N}(0, 1)$. If all firms follow the suggested strategy, the resulting discounted average payoff of any firm must satisfy

$$v = (1-\delta)\pi^M + \delta\left[\Phi\left(\frac{\ln a - \mu^M}{\theta}\right)v + \left(1 - \Phi\left(\frac{\ln a - \mu^M}{\theta}\right)\right)\pi^N\right]$$

and so

$$v = \frac{(1-\delta)\pi^M + \delta\left(1 - \Phi\left(\frac{\ln a - \mu^M}{\theta}\right)\right)\pi^N}{1 - \delta\Phi\left(\frac{\ln a - \mu^M}{\theta}\right)} \tag{8}$$

If firm 1, say, deviates and charges a price p_1 in period t, then its per-period profits are

$$(1-\delta)\pi_1(p_1) + \delta\left[\Phi\left(\frac{\ln a - \mu(p_1)}{\theta}\right)\left(v - \pi^N\right) + \pi^N\right]$$

where, to save on notation, we write

$$\pi_1(p_1) = \pi_1\left(p_1, \mathbf{p}_{-1}^M\right)$$

as the expected per-period profits of firm 1 when it charges p_1 and all other firms charge the monopoly price p^M and

$$\mu(p_1) = \frac{1}{n}\sum_j \mu_j\left(p_1, \mathbf{p}_{-1}^M\right)$$

is the industry mean of the expected log sales in the same circumstances.

Firm 1 does not have an incentive to deviate as long as: for all p_1

$$\begin{aligned}&(1-\delta)\pi^M + \delta\left[\Phi\left(\frac{\ln a - \mu^M}{\theta}\right)\left(v - \pi^N\right) + \pi^N\right] \\ &\geq (1-\delta)\pi_1(p_1) + \delta\left[\Phi\left(\frac{\ln a - \mu(p_1)}{\theta}\right)\left(v - \pi^N\right) + \pi^N\right]\end{aligned}$$

After substituting for v from (8) and rearranging terms, this becomes: for all p_1

$$\frac{1 - \delta\Phi\left(\frac{\ln a - \mu(p_1)}{\theta}\right)}{1 - \delta\Phi\left(\frac{\ln a - \mu^M}{\theta}\right)} - \frac{\pi_1(p_1) - \pi^N}{\pi^M - \pi^N} \geq 0 \tag{9}$$

It will be convenient to rewrite the incentive condition (9) by changing the variable from prices $p_1 \in (0, p^{\max}]$ to "log quantities," specifically, industry mean expected log sales $m \in [\mu(p^{\max}), \mu(0))$. Because all other firms are charging p^M, every price p_1 of firm 1 induces a value of m uniquely and vice versa. To see the latter, notice that Conditions 1 and 2 (see Section 2) imply that $\mu(p_1)$ is a decreasing and concave function of p_1 over the interval $(0, p^{\max}]$. With this change of variable, (9) is equivalent to the statement that for all $m \in [\mu(p^{\max}), \mu(0))$,

$$L_\theta(m) \equiv \frac{1 - \delta\Phi\left(\frac{\ln a - m}{\theta}\right)}{1 - \delta\Phi\left(\frac{\ln a - \mu^M}{\theta}\right)} - \frac{\Pi_1(m) - \pi^N}{\pi^M - \pi^N} \geq 0 \tag{10}$$

where $\Pi_1(m) = \pi_1(\mu^{-1}(m))$. Because μ is decreasing and concave, so is μ^{-1} and the concavity of π_1 implies that Π_1 is concave as well. Of course, $L_\theta(\mu^M) = 0$.

We will show below that when θ is small enough, for an appropriate choice of the trigger threshold $a = a_\theta$, $L_\theta(m) \geq 0$ for all $m \in [\mu(p^{\max}), \mu(0))$.

□ **Choice of trigger.** The trigger is chosen so that no *local* deviation around monopoly is profitable. In other words, the trigger is chosen to satisfy $L'_\theta(\mu^M) = 0$ and this can be done when θ is below a threshold. Later we will show that when θ is small enough, this choice of a will also make large deviations unprofitable.

Lemma 2. For all δ, there exists a θ^* such that for all $\theta < \theta^*$ there is an a_θ such that when $a = a_\theta$, L_θ attains a local minimum at μ^M.

Proof. Note that

$$L'_\theta\left(\mu^M\right) = \frac{1}{\theta} \frac{\delta\phi\left(\frac{\ln a - \mu^M}{\theta}\right)}{1 - \delta\Phi\left(\frac{\ln a - \mu^M}{\theta}\right)} - \frac{\Pi'_1(\mu^M)}{\pi^M - \pi^N}$$

where $\phi \equiv \Phi'$ is the density of the standard normal $\mathcal{N}(0, 1)$.

Consider the function

$$\lambda(z) \equiv \frac{\delta\phi(z)}{1 - \delta\Phi(z)}$$

This has a unique maximum at $z^* > 0$, is decreasing for all $z > z^*$, and approaches 0 as $z \to \infty$ (see Appendix A). Note that z^* depends only on δ. Let θ^* be such that

$$\frac{1}{\theta^*} \frac{\delta\phi(z^*)}{1 - \delta\Phi(z^*)} - \frac{\Pi'_1(\mu^M)}{\pi^M - \pi^N} = 0$$

For any $\theta < \theta^*$, there exists a unique $z_\theta > z^*$ that solves

$$\frac{1}{\theta} \frac{\delta\phi(z_\theta)}{1 - \delta\Phi(z_\theta)} - \frac{\Pi'_1(\mu^M)}{\pi^M - \pi^N} = 0$$

and now let the trigger a_θ be chosen to satisfy

$$\frac{\ln a_\theta - \mu^M}{\theta} = z_\theta > 0 \tag{11}$$

Moreover, when $a = a_\theta$

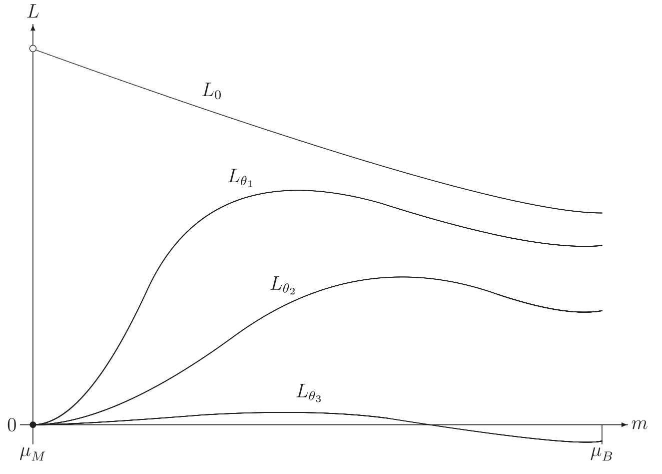
$$L''_\theta\left(\mu^M\right) = \frac{1}{\theta^2} \frac{\delta\phi(z_\theta) z_\theta}{1 - \delta\Phi(z_\theta)} - \frac{\Pi''_1(\mu^M)}{\pi^M - \pi^N} > 0$$

because $\phi'(z) = -z\phi(z)$ and $\Pi''_1(\mu^M) < 0$. □

□ **Preventing large deviations.** The choice of the trigger a_θ as in Lemma 2 prevents small deviations around μ^M. We now complete the proof that the strategies constitute an equilibrium by showing that, in fact, they deter large deviations as well. Precisely, once θ is small enough, for all m, $L_\theta(m) \geq 0$. Specifically,

FIGURE 1

A DEPICTION OF THE CONVERGENCE OF L_θ TO L_0



Lemma 3. For any $\delta > \underline{\delta}$, the perfect monitoring benchmark, there exists a $\overline{\theta}$ such that for all $\theta < \overline{\theta}$,

$$L_\theta(m) > 0$$

for all $m \in [\mu(p^{\max}), \mu(0))$.

The proof of Lemma 3 is rather involved and so we begin by outlining the main elements. First, it is useful to divide the potential deviations into three cases depending on the value of m.

(1) $\mu(p^{\max}) \leq m \leq \mu^M$
(2) $\mu^M < m \leq \mu^B$
(3) $\mu^B < m < \mu(0)$

Case 1 considers deviations such that m is smaller than μ^M (or equivalently, deviations in which a firm raises its price above the monopoly price p^M). Some routine calculations show that $L'_\theta < 0$ for all $m \leq \mu^M$. Moreover, it attains a minimum value of 0 at $m = \mu^M$ and so $L_\theta > 0$ for all $m < \mu^M$. Note that although charging such high prices may not be profitable in the short run, there is a long-run benefit in that this increases the probability that the cartel will continue to charge p^M in the future. The argument above then establishes that the short-run losses outweigh any long-run benefits.

Case 3 considers deviations such that m is larger than even μ^B (or equivalently, deviations in which a firm undertakes a severe price cut below p^B) In this case, it is easy to verify that $L'_\theta > 0$. Now there is no trade-off between the short run and the long run—an increase in m (or a decrease in p_1) decreases both current profits and makes punishment more likely.

Although it is relatively simple to rule out deviations in Cases 1 and 3, doing so for Case 2 is complicated and constitutes the bulk of the proof. This is because Case 2 considers *moderate* increases in m above μ^M. Equivalently, it considers moderate price cuts from p^M which affect the aggregate only slightly. For such deviations, the short-run versus long-run trade-off causes L'_θ to change sign, perhaps more than once (see Figure 1). Thus, it is difficult to identify a "best" deviation and so a different approach is needed. The main idea here is to show that as θ gets small, L_θ monotonically approaches L_0, the corresponding function for $\theta = 0$. The latter is equivalent to a situation with perfect monitoring. The difficulty here is that L_0 is a discontinuous function (formally defined below in (12)). This is because, with perfect monitoring, even the smallest deviation triggers a large punishment with probability one. The main argument rests on the fact that L_θ converges *pointwise* to L_0 (Lemma 9). But because L_0 is discontinuous at μ^M, the *pointwise* convergence of L_θ to L_0 by itself does not guarantee that $L_\theta(m) \geq 0$ for m close to μ^M. In other words, it is possible that small price cuts from p^M may have a such a small probability of being detected that they turn out to be profitable. To rule out this possibility, a closer examination of the local behavior of L_θ in a neighborhood of μ^M (resulting from small price cuts) is needed and taking care of all these details account for the somewhat lengthy proof.

Figure 1 illustrates the functions L_θ and L_0 for the case of linear demand.[14] The figure depicts the functions L_θ for values of θ such that $\theta_1 < \theta_2 < \theta_3$ as well as the (discontinuous) function L_0. When θ is large (that is, $\theta = \theta_3$), aggregate log sales are quite variable and the incentive condition fails. As θ decreases to θ_2, and then to θ_1, aggregate sales become less variable and deviations become unprofitable.

□ **Proof of Lemma 3.**

Case 1. $\mu(p^{\max}) \leq m \leq \mu^M$. Note that

$$L''_\theta(m) = \frac{1}{\theta^2} \frac{\delta\phi\left(\frac{\ln a_\theta - m}{\theta}\right)}{1 - \delta\Phi\left(\frac{\ln a_\theta - \mu^M}{\theta}\right)} \left(\frac{\ln a_\theta - m}{\theta}\right) - \frac{\Pi''_1(m)}{\pi^M - \pi^N}$$

again using the fact that for the standard normal density $\phi'(z) = -z\phi(z)$. As $\Pi''_1(m) < 0$, we have that for all $m \leq \mu^M$

$$L''_\theta(m) > 0$$

and so L_θ is strictly convex in the interval $[\mu(p^{\max}), \mu^M]$. Moreover, because the trigger is chosen so that $L'_\theta(\mu^M) = 0$ (see Lemma 2) this means that L_θ reaches a minimum at μ^M; that is, for all $m \in [\mu(p^{\max}), \mu^M)$, $L_\theta(m) > L_\theta(\mu^M) = 0$.

Case 2. $\mu^M \leq m \leq \mu^B$. Define

$$L_0(m) = \begin{cases} 0 & \text{if } m = \mu^M \\ \frac{1}{1-\delta} - \frac{\Pi_1(m) - \pi^N}{\pi^M - \pi^N} & \text{if } m \neq \mu^M \end{cases} \tag{12}$$

Unlike the case for $\theta > 0$, L_0 is discontinuous at μ^M and in fact,

$$\lim_{m \to \mu^M} L_0(m) = \frac{\delta}{1-\delta} \tag{13}$$

and by comparing (10) to (12), it is easily verified that for all $\theta > 0$, $L_\theta < L_0$. Moreover,

$$L'_\theta(m) > L'_0(m)$$

a fact that will play a key role below.

Define $\widehat{m}_\theta$ to be the maximizer of L_θ in the interval $[\mu^M, \mu^B]$. From Lemma 10 in Appendix B, as $\theta \to 0$, $\lim \widehat{m}_\theta = \mu^M$ and $\lim L_\theta(\widehat{m}_\theta) = \frac{\delta}{1-\delta}$.

Three subcases have to be considered separately. Recall from (11) that $\mu^M < \ln a_\theta$.

First, if $\mu^M < m \leq \ln a_\theta$, the same argument as in Case 1 shows that both $L_\theta(m) > 0$ and $L'_\theta(m) > 0$. Because $\widehat{m}_\theta$ maximizes L_θ in $[\mu^M, \mu^B]$, this implies that $\ln a_\theta < \widehat{m}_\theta$.

Second, if $\ln a_\theta < m \leq \widehat{m}_\theta$, then

[14] Specifically, $n = 3$ and $\delta = 0.7$. The expected demand of firm i is $q_i = 10 - 2p_i + \frac{1}{2}(p_j + p_k)$.

$$\begin{aligned} L_\theta(m) &= L_\theta(\ln a_\theta) + \int_{\ln a_\theta}^{m} L'_\theta(s)ds \\ &\geq L_\theta(\ln a_\theta) + \int_{\ln a_\theta}^{m} L'_0(s)ds \\ &\geq L_\theta(\ln a_\theta) + \int_{\ln a_\theta}^{\widehat{m}_\theta} L'_0(s)ds \end{aligned}$$

as $L'_\theta > L'_0$ and $L'_0 < 0$. In the limit we have

$$\begin{aligned} \lim_{\theta\to 0} L_\theta(m) &\geq \lim_{\theta\to 0} L_\theta(\ln a_\theta) + \int_{\lim(\ln a_\theta)}^{\lim \widehat{m}_\theta} L'_0(s)ds \\ &= \frac{1}{2}\frac{\delta}{1-\delta} \\ &> 0 \end{aligned}$$

using Lemma 11 in Appendix B and the fact that $\lim \widehat{m}_\theta = \lim(\ln a_\theta) = \mu^M$.

Finally, suppose that $\widehat{m}_\theta \leq m \leq \mu^B$. Then,

$$\begin{aligned} L_\theta(m) &= L_\theta(\widehat{m}_\theta) + \int_{\widehat{m}_\theta}^{m} L'_\theta(s)ds \\ &\geq L_\theta(\widehat{m}_\theta) + \int_{\widehat{m}_\theta}^{m} L'_0(s)ds \end{aligned}$$

and taking limits and using (13), we have

$$\begin{aligned} \lim_{\theta\to 0} L_\theta(m) &\geq \frac{\delta}{1-\delta} + \lim_{\theta\to 0}\int_{\widehat{m}_\theta}^{m} L'_0(s)ds \\ &= L_0\left(\mu^M\right) + \int_{\mu^M}^{m} L'_0(s)ds \\ &= L_0(m) \\ &> 0 \end{aligned}$$

where $L_0(\mu^M) = \lim_{s\to\mu^M} L_0(s)$ and $L_0(m) > 0$ by our choice of δ.

Case 3. $\mu^B \leq m \leq \mu(0)$. In this region

$$L'_\theta(m) = \frac{1}{\theta}\frac{\delta\phi\left(\frac{\ln a_\theta - m}{\theta}\right)}{1-\delta\Phi\left(\frac{\ln a_\theta - \mu^M}{\theta}\right)} - \frac{\Pi'_1(m)}{\pi^M - \pi^N}$$

is positive because $\Pi'_1(m) < 0$ for $m > \mu^B$. Now because $L_\theta(\mu^B) > 0$ for small θ, then so is $L_\theta(m)$.

This concludes the proof of Lemma 3.

5. Main result

■ The main result then combines the findings of the previous two sections. In particular, Proposition 1 shows that there is a $\underline{\sigma}$ (which can be chosen independently of ρ) such that for all $\sigma > \underline{\sigma}$, the profits from any equilibrium without aggregate information are bounded away from monopoly levels. Proposition 2 shows that for any fixed σ, when ρ is small enough (and hence θ is small as well), there exists an equilibrium with aggregate information whose profits approximate monopoly.

Thus, we have

Theorem 1. Fix $\delta > \underline{\delta}$, the perfect monitoring benchmark. There exists a $\underline{\sigma}$ such that for every $\sigma > \underline{\sigma}$ there exists a $\overline{\rho}_\sigma$ such that for all $\rho < \overline{\rho}_\sigma$, there is an equilibrium with aggregate sales information whose profits exceed that from any equilibrium without such information. In the limit, for fixed σ, as $\rho \downarrow -\frac{1}{n-1}$, profits from the equilibrium with aggregate sales approach monopoly profits.

□ **Linear demand.** An upper bound to the gains from the exchange of aggregate sales information over no exchange is $\pi^M - \pi^N$. In the case of linear demand, this bound can be arbitrarily approached when σ is large and ρ is small enough. Using Corollary 1 we have

Corollary 2. Fix $\delta > \underline{\delta}$, the perfect monitoring benchmark. Given any $\kappa > 0$, there exists a $\underline{\sigma}$ such that for every $\sigma > \underline{\sigma}$ there exists a $\overline{\rho}_\sigma$ such that for all $\rho < \overline{\rho}_\sigma$, the profit gains from the exchange of aggregate sales information exceeds $\pi^M - \pi^N - \kappa$.

6. Why Swiss accountants are needed

■ We have supposed throughout that accurate aggregate sales information is disseminated by a third party, say a trade association or a consulting firm employed by the industry. But this requires that firms report their individual sales figures truthfully to the third party. There are many instances in which firms verify sales figures. According to a European Commission report in a cartel case involving amino acids:

> "ADM [Archer Daniels Midland] stated that the way for them to communicate is through a trade association. ADM explained by way of example that ADM reported its citric acid sales every month to a trade association, and every year, Swiss accountants audited those figures." (European Commission, 2001b)

Similarly, in the carbonless paper cartel an employee of one of the firms

> "[…] who doubted the figures supplied by Sarrió (Torraspapel), had asked and received permission to audit the information on Sarrió's sales volumes on Sarrió's premises. (European Commission, 2001a)

We now consider a variant of our model in which firms report their log sales to a trade association but these reports are *not* verifiable—that is, audits are not possible. In other words, firms can tailor their reports strategically if they wish—these are merely "cheap talk." Once again the trade association publishes the average of log sales *as reported* by firms. In such a setting, can the exchange of aggregate information facilitate collusion? We show below that if the aggregate is composed of the unverified reports of individual firms, then there is always an incentive to misreport, typically to underreport, sales. The reason is that as only aggregate sales are made public, a firm can cheat by secretly cutting its price and then reporting sales in a way that the distribution of reported aggregate sales after the price cut is identical to the distribution of actual aggregate sales without the price cut. This result sheds some light on why cartels find it difficult to share information accurately and so have to resort to external auditing or to other verification methods.

Formally, suppose that as in the original model of Section 2, firms have private knowledge of their own sales Y_i. At the end of each period, each firm sends a report Z_i to a third party—a trade association—about its sales. These reports are unconstrained—a firm is free to report any amount Z_i it wishes—and cannot be verified. The trade association publishes the average of the *reported* log sales

$$\ln A = \frac{1}{n}\sum_{j=1}^{n} \ln Z_j$$

and in a public perfect equilibrium (PPE), firms' pricing strategies depend only on the history of these aggregates.[15] We ask whether there are collusive public perfect equilibria—in which the profits exceed those in the one-shot Nash equilibrium—with the property that every firm has the incentive to truthfully and accurately report its sales to the trade association. In other words, are there collusive equilibria in which it is optimal for every firm to report $Z_i = Y_i$ provided that all other firms are doing the same.

Proposition 3. No collusive pure strategy *PPE* with aggregate information can induce truthful reporting.

Proof. Suppose $\mathbf{s}$ is a collusive pure strategy *PPE* with aggregate information A and suppose that t is a period in which, given the public history of (reported) aggregates, $h^{t-1} = (a^1, \ldots, a^{t-1})$,$\mathbf{s}^t(h^{t-1}) = \mathbf{p} \neq \mathbf{p}^N$. Then there is a firm, say 1, for whom a price, say p_1', is such that in period t, $\pi_1(p_1', \mathbf{p}_{-1}) > \pi_1(\mathbf{p})$. Note that because the history of aggregates is public, the prices that firms will charge in period t, $\mathbf{p}$ is commonly known.

If firm 1 follows the strategy s_1 and charges p_1 as prescribed, then the log sales in period t, $\ln \mathbf{Y}$ are jointly distributed according to a multivariate normal distribution $\mathcal{N}(\mu(\mathbf{p}), \Sigma)$ and so with truthful reporting, the aggregate $\ln A$ (average log sales) is distributed according to a univariate normal distribution $\mathcal{N}(\mu(\mathbf{p}), \theta^2)$, where $\mu(\mathbf{p}) = \frac{1}{n}\sum_i \mu_i(\mathbf{p})$ and $\theta^2 = \frac{1}{n}\sigma^2(1 + (n - 1)\rho)$.

Now suppose firm 1 deviates to a price p_1' in period t. Then, the log sales $\ln \mathbf{Y}$ are jointly distributed according to $\mathcal{N}(\mu(p_1', \mathbf{p}_{-1}), \Sigma)$ and the actual average log sales in period t, $\ln A$ are distributed according to $\mathcal{N}(\mu(p_1', \mathbf{p}_{-1}), \theta^2)$.

Suppose firm 1 misreports its log sales by a constant amount $n(\mu(\mathbf{p}) - \mu(p_1', \mathbf{p}_{-1}))$. Then the reported average log sales, denoted by, $\ln A'$ are such that

$$\begin{aligned} E[\ln A'] &= \mu\big(p_1', \mathbf{p}_{-1}\big) + \frac{1}{n} \times n\big(\mu(\mathbf{p}) - \mu\big(p_1', \mathbf{p}_{-1}\big)\big) \\ &= \mu(\mathbf{p}) \end{aligned}$$

Thus, the distribution of the aggregate with no deviation and truthful reporting is the same as the distribution after a deviation and false reporting. This means that firm 1 can make a short-term gain with no long-term consequences. □

Although the proposition is quite general in that it makes no assumptions about the demand functions, it concerns only *pure* strategy public perfect equilibria. Can randomized strategies induce truthful reporting? We show in an example—with multiplicatively separable demand functions—that even randomization cannot induce firms to report truthfully to the trade association.

Proposition 4. Suppose that expected demand functions are multiplicatively separable, that is, the expected demand

$$q_i(\mathbf{p}) = \beta(p_i)\prod_{j\neq i}\gamma(p_j)$$

[15] Firms' overall strategies are private, however, because although their actions depend on public histories, their reports may depend on private histories. Kandori (2003) calls such strategies semipublic.

Then no collusive PPE, pure or mixed, can induce truthful reporting.

Proof. As above, suppose $\mathbf{s}$ is a collusive *PPE* that is possibly mixed and suppose that t is a period in which given the public history of (reported) aggregates, $h^{t-1} = (a^1, \ldots, a^{t-1})$, the distribution over prices $\mathbf{s}^t(h^{t-1})$ is not degenerate on $\mathbf{p}^N$. Then there is a firm, say 1, for whom a price, say p_1', is such that in period t, $\pi_1(p_1', \mathbf{s}_{-1}^t(h^{t-1})) > \pi_1(\mathbf{s}^t(h^{t-1}))$.

Note that when (expected) demand functions are multiplicatively separable, the effect of a price change by firm 1 on the log sales of any firm is independent of the prices of firms other than 1. Precisely, for all p_1 and p_1', the cross-price effect on log sales

$$\ln q_i(\mathbf{p}) - \ln q_i(p_1', \mathbf{p}_{-1}) = \gamma(p_1) - \gamma(p_1')$$

and the own-price effect

$$\ln q_1(\mathbf{p}) - \ln q_1(p_1', \mathbf{p}_{-1}) = \beta(p_1') - \beta(p_1)$$

do not depend on $\mathbf{p}_{-1}$.

Because

$$q_i(\mathbf{p}) = E[Y_i \mid \mathbf{p}] = \exp\left(\mu_i(\mathbf{p}) + \tfrac{1}{2}\sigma^2\right)$$

we have that

$$\mu_i(\mathbf{p}) = \ln q_i(\mathbf{p}) - \tfrac{1}{2}\sigma^2$$

and so

$$\mu(\mathbf{p}) = \frac{1}{n}\sum_{i=1}^{n} \ln q_i(\mathbf{p}) - \tfrac{1}{2}\sigma^2$$

Similarly,

$$\mu(p_1', \mathbf{p}_{-1}) = \frac{1}{n}\sum_{i=1}^{n} \ln q_i(p_1, \mathbf{p}_{-1}) - \tfrac{1}{2}\sigma^2$$

$$\mu(\mathbf{p}) - \mu(p_1', \mathbf{p}_{-1}) = \frac{1}{n}\left(\beta(p_1') - \beta(p_1)\right) + \frac{n-1}{n}\left(\gamma(p_1) - \gamma(p_1')\right)$$

Now suppose that firm 1 deviates and charges p_1' instead of $s_1(h_1^{t-1})$. Suppose further that the firm then reports its log sales to be:

$$\ln Z_1 = \ln Y_1 + \left(\beta(p_1') - \beta(p_1)\right) + (n-1)\left(\gamma(p_1) - \gamma(p_1')\right)$$

In other words, the reported log sales $\ln Z_1$ equal the actual log sales $\ln Y_1$ plus an amount that depends neither on the prices of other firms $\mathbf{p}_{-1}$ nor on the sales of other firms $\mathbf{y}_{-1}$. Thus, this constitutes a feasible "lying" strategy. Misreporting in this manner has the property that, for any $\mathbf{p}_{-1}$, the resulting (false) aggregate $\ln A'$ satisfies

$$\begin{aligned}
E\left[\ln A' \mid p_1', \mathbf{p}_{-1}\right] &= \frac{1}{n}E\left[\ln Z_1 + \sum_{i=2}^{n} \ln Y_i \mid p_1', \mathbf{p}_{-1}\right] \\
&= \frac{1}{n}E\left[\sum_{i=1}^{n} \ln Y_i \mid p_1', \mathbf{p}_{-1}\right] \\
&\quad + \frac{1}{n}\left(\beta(p_1') - \beta(p_1)\right) + \frac{n-1}{n}\left(\gamma(p_1) - \gamma(p_1')\right) \\
&= \mu(p_1', \mathbf{p}_{-1}) + \mu(\mathbf{p}) - \mu(p_1', \mathbf{p}_{-1})
\end{aligned}$$

$$= \mu(\mathbf{p})$$
$$= E[\ln A \mid \mathbf{p}]$$

Suppose that $p_1 \in \text{supp}\, s_1(h^{t-1})$ and p_1' is a price that is profitable in the short run against $\mathbf{s}_{-1}(h^{t-1})$. Because the deviation is such that for all $\mathbf{p}_{-1}$, the distribution of the reported aggregate after the deviation is the same as the distribution of the aggregate in the absence of a deviation, no matter how the other firms are randomizing, this deviation will result in the same distribution of the aggregate. Thus, we have shown that there is a deviating strategy for firm 1 which is profitable in the short term and has no long-term consequences. □

A cautionary note. The results of this section should not be interpreted to say that there is little harm in allowing firms to self-report sales to each other. Our analysis pertains only to situations in which firms report their sales to a trade association which then publicly releases aggregate information only. If individual sales figures were to become commonly known, it might be possible to induce truthtelling.

7. Conclusion

■ Communication is central to antitrust law—tacit collusion is not unlawful per se whereas explicit collusion is. Kaplow (2013) has argued forcefully that this focus on "communication-based prohibition" is misguided. First, communication that facilitates collusion can take many forms and does not always have to take place in smoke-filled rooms. For instance, firms could make public announcements of their pricing intentions. Second, one should focus on the welfare effects of the conduct rather than how it was achieved. This aspect of communication—to *reach* a collusive agreement—is, however, different from that considered here. In this article, we study interfirm communication that may help firms monitor each other better to *sustain* a collusive agreement. Antitrust authorities are well aware of this aspect of communication and have sought to curb cartelization by imposing restrictions on the kinds of information that firms can exchange. Our findings point out that these restrictions may be less effective than suggested by economic theory.

Appendix A: Proof of Proposition 2

□ **Properties of λ.** In this section we study the properties of the function

$$\lambda(z) \equiv \frac{\delta\phi(z)}{1 - \delta\Phi(z)}$$

where ϕ and Φ are the density and distribution functions of the standard normal. Note that $\lambda(z)$ is like a "discounted hazard rate"—for $\delta = 1$ it is just the hazard rate of the normal. Many arguments in the proof of the main result hinge on the shape of λ.

Lemma 4. Suppose $\delta < 1$. Then λ has a unique maximum at $z^* > 0$.

Proof. Notice that

$$\lambda' = \frac{(1-\delta\Phi)\delta\phi' + \delta^2\phi^2}{(1-\delta\Phi)^2}$$
$$= \frac{-z(1-\delta\Phi)\delta\phi + \delta^2\phi^2}{(1-\delta\Phi)^2}$$
$$= \lambda(\lambda - z)$$

where the second equality follows from the fact that $\phi' = -z\phi$ for the standard normal density ϕ. Thus, for all $z \leq 0$, $\lambda'(z) > 0$.

Next, because $\lambda(z) < \frac{\delta}{1-\delta}\phi(z)$, and $\lim_{z\to\infty}\phi(z) = 0$, we have that $\lim_{z\to\infty}\lambda(z) = 0$ as well. Together with the fact that $\lambda(0) > 0$ and for all $z \le 0$, $\lambda'(z) > 0$, this implies that there exists an $z^* > 0$ where λ reaches a maximum. Let $z^* > 0$ be the smallest maximizer of λ. Then

$$0 = \lambda'(z^*) = \lambda(z^*)(\lambda(z^*) - z^*)$$

and thus $\lambda(z^*) = z^*$. Now for any $z > z^*$, we have $z > z^* = \lambda(z^*) \ge \lambda(z)$ because z^* maximizes λ and so for any $z > z^*$,

$$\lambda'(z) = \lambda(z)(\lambda(z) - z) < 0$$

Thus, z^* is the unique maximizer of λ. □

Lemma 5. For all z, $\lambda'(z) < 1$.

Proof. Recall that

$$\lambda' = \lambda(\lambda - z)$$

For all $z \ge z^*$, $\lambda'(z) \le 0$ and so the condition trivially holds.

Now suppose, by contradiction, that there exists an $z < z^*$ such that $\lambda'(z) \ge 1$. Now notice that

$$\begin{aligned}\lambda''(z) &= \lambda'(z)(\lambda(z) - z) + \lambda(z)(\lambda'(z) - 1)\\ &= \lambda(z)(\lambda(z) - z)^2 + \lambda(z)(\lambda'(z) - 1)\end{aligned}$$

Thus, we have that at any $z < z^*$ such that $\lambda'(z) \ge 1$ it is the case that

$$\lambda''(z) > 0$$

and so λ' is increasing. This implies that if there is an $z < z^*$ such that $\lambda'(z) \ge 1$, then $\lambda'(z^*) > 1$ which is a contradiction. □

Lemma 6. For all $z < z' < z^*$ or $z^* < z < z'$

$$\frac{1}{\lambda(z) - z} < \frac{1}{\lambda(z') - z'}$$

Moreover, $\frac{1}{\lambda(z)-z} \to \infty$ as $z \uparrow z^*$ and $\frac{1}{\lambda(z)-z} \to -\infty$ as $z \downarrow -\infty$.

Proof. The derivative of the given function is

$$-\frac{\lambda'(z) - 1}{(\lambda(z) - z)^2} > 0$$

because $\lambda'(z) < 1$. □

Recall that for $\theta < \theta^*$, the normalized trigger $z_\theta = \frac{\ln a_\theta - \mu^M}{\theta} > z^*$ is chosen to satisfy

$$\begin{aligned}0 &= L'_\theta\left(\mu^M\right)\\ &= \frac{1}{\theta}\frac{\delta\phi(z_\theta)}{1 - \delta\Phi(z_\theta)} - \frac{\Pi'_1\left(\mu^M\right)}{\pi^M - \pi^N}\\ &= \frac{1}{\theta}\lambda(z_\theta) - \frac{\Pi'_1\left(\mu^M\right)}{\pi^M - \pi^N}\end{aligned} \tag{A1}$$

Lemma 7. Suppose z_θ is the normalized trigger as defined in (A1). Then,

$$\lim_{\theta\to 0} z_\theta = \infty \text{ and } \lim_{\theta\to 0}(\theta z_\theta) = 0$$

Proof. From (A1) it must be that as $\theta \to 0$, $\lambda(z_\theta) \to 0$ or equivalently, that $z_\theta \to \infty$.

Differentiating (A1) with respect to θ, we get

$$\begin{aligned}\frac{dz_\theta}{d\theta} &= \frac{1}{\theta}\frac{\lambda(z_\theta)}{\lambda'(z_\theta)}\\ &= \frac{1}{\theta}\frac{1}{\lambda(z_\theta) - z_\theta}\\ &< 0\end{aligned} \tag{A2}$$

because $z_\theta > z^*$. Using L'Hospital's rule,

$$\lim_{\theta\to 0}(\theta z_\theta) = \lim_{\theta\to 0}\left(\frac{\frac{dz_\theta}{d\theta}}{\frac{d}{d\theta}\left(\frac{1}{\theta}\right)}\right) = \lim_{\theta\to 0}\frac{\frac{1}{\theta}\frac{1}{\lambda(z_\theta)-z_\theta}}{-\frac{1}{\theta^2}} = \lim_{\theta\to 0}\left(\frac{\theta}{z_\theta-\lambda(z_\theta)}\right) = 0$$

because $\lambda(z_\theta)\to 0$ and $z_\theta\to\infty$. □

□ **Properties of L_θ.** This section contains results on various properties of the function $L_\theta(m)$ in the region $\mu^M \le m \le \mu^B$. Thus, it studies the incentives of a firm to deviate to prices p_1 which represent small price cuts from the monopoly price p^M. The first lemma shows that the incentive to deviate in such a manner falls as the variability of the aggregate, measured by θ, decreases.

Lemma 8. For all $m \in (\mu^M, \mu^B]$,

$$\frac{dL_\theta(m)}{d\theta} < 0$$

Proof. Because

$$L_\theta(m) = \frac{1-\delta\Phi\left(\frac{\ln a_\theta - m}{\theta}\right)}{1-\delta\Phi\left(\frac{\ln a_\theta-\mu^M}{\theta}\right)} - \frac{\Pi_1(m)-\pi^N}{\pi^M-\pi^N}$$

differentiating this with respect to θ shows that the sign of $\frac{dL_\theta(m)}{d\theta}$ is the same as the sign of

$$\frac{\delta\phi\left(\frac{\ln a_\theta-\mu^M}{\theta}\right)}{1-\delta\Phi\left(\frac{\ln a_\theta-\mu^M}{\theta}\right)}\frac{\theta\frac{d\ln a_\theta}{d\theta}-\left(\ln a_\theta-\mu^M\right)}{\theta^2} - \frac{\delta\phi\left(\frac{\ln a_\theta-m}{\theta}\right)}{1-\delta\Phi\left(\frac{\ln a_\theta-m}{\theta}\right)}\frac{\theta\frac{d\ln a_\theta}{d\theta}-(\ln a_\theta-m)}{\theta^2}$$

Writing $z=\frac{\ln a_\theta-m}{\theta}$ and recalling that $z_\theta=\frac{\ln a_\theta-\mu^M}{\theta}$, the expression above becomes

$$\frac{\delta\phi(z_\theta)}{1-\delta\Phi(z_\theta)}\frac{\theta\frac{d\ln a_\theta}{d\theta}-\left(\ln a_\theta-\mu^M\right)}{\theta^2} - \frac{\delta\phi(z)}{1-\delta\Phi(z)}\frac{\theta\frac{d\ln a_\theta}{d\theta}-(\ln a_\theta-m)}{\theta^2}$$

which can be rewritten as

$$\lambda(z_\theta)\frac{1}{\theta}\left(\frac{d\ln a_\theta}{d\theta}-z_\theta\right)-\lambda(z)\frac{1}{\theta}\left(\frac{d\ln a_\theta}{d\theta}-z\right)$$

But we have shown that (see (A2))

$$\frac{dz_\theta}{d\theta}=\frac{1}{\theta}\frac{1}{\lambda(z_\theta)-z_\theta}$$

Equivalently,

$$\begin{aligned}\frac{d\left(\frac{\ln a_\theta-\mu^M}{\theta}\right)}{d\theta} &= \frac{\theta\frac{d\ln a_\theta}{d\theta}-\left(\ln a_\theta-\mu^M\right)}{\theta^2}\\ &= \frac{1}{\theta}\left(\frac{d\ln a_\theta}{d\theta}-z_\theta\right)\end{aligned}$$

and so

$$\frac{d\ln a_\theta}{d\theta}=\frac{1}{\lambda(z_\theta)-z_\theta}+z_\theta$$

Substituting in the expression above, we obtain that the sign of $\frac{dL_\theta(m)}{d\theta}$ is the same as the sign of

$$\lambda(z_\theta)\left(\frac{1}{\lambda(z_\theta)-z_\theta}\right)-\lambda(z)\left(\frac{1}{\lambda(z_\theta)-z_\theta}+z_\theta-z\right)$$

For $z_\theta > z^*$, define

$$J(z)=\frac{\lambda(z_\theta)}{\lambda(z_\theta)-z_\theta}-\lambda(z)\left(\frac{1}{\lambda(z_\theta)-z_\theta}+z_\theta-z\right)$$

Observe that

$$J(z_\theta)=0$$

The monotonicity of $L_\theta(m)$ in θ is implied by the statement that $J(z) < J(z_\theta)$ for all z.

First, note that

$$\lim_{|z|\to\infty} J(z) = \frac{\lambda(z_\theta)}{\lambda(z_\theta) - z_\theta} < 0$$

because both $\lambda(z)$ and $z\lambda(z)$ converge to 0 as $|z| \to \infty$.

For any $z \neq z^*$, we can write

$$J'(z) = \lambda'(z)\left(\frac{1}{\lambda(z) - z} + z - \frac{1}{\lambda(z_\theta) - z_\theta} - z_\theta\right)$$

because $\lambda'(z) = \lambda(z)(\lambda(z) - z)$.

Case 1: $z < z^*$.

For $z < z^*$, $\lambda'(z) > 0$ and so the sign of $J'(z)$ is the same as that of

$$\frac{1}{\lambda(z) - z} + z - \frac{1}{\lambda(z_\theta) - z_\theta} - z_\theta$$

Lemma 6 shows that the function $\frac{1}{\lambda(z)-z} + z$ is increasing over $(-\infty, z^*)$ and is onto the whole real line. This means that (i) for small z (a large negative number), $J'(z) < 0$, (ii) for some $z_0 < z^*$, $J'(z_0) = 0$, and (iii) for all $z \in (z_0, z^*)$, $J'(z) > 0$.

Case 2: $z^* < z < z_\theta$.

For $z \in (z^*, z_\theta)$, $J'(z) > 0$ because $\lambda'(z) < 0$ but

$$\frac{1}{\lambda(z) - z} + z - \frac{1}{\lambda(z_\theta) - z_\theta} - z_\theta < 0$$

Case 3: $z > z_\theta$.

Finally, for $z > z_\theta$ we have that $J'(z) < 0$.

Thus the behavior of the function $J(z)$ is as follows:

$$J'(z) = \begin{cases} < 0 & z < (-\infty, z_0) \\ = 0 & z = z_0 \\ > 0 & z \in (z_0, z_\theta) \\ = 0 & z = z_\theta \\ < 0 & z > z_\theta \end{cases}$$

Combined with the fact that

$$\lim_{|z|\to\infty} J(z) = \frac{\lambda(z_\theta)}{\lambda(z_\theta) - z_\theta} < 0$$

this means that J is maximized at $z = z_\theta$ and hence for all $z \neq z_\theta$,

$$J(z) < J(z_\theta) = 0$$

□

The next result establishes the pointwise convergence of the function L_θ to its perfect monitoring counterpart L_0.

Lemma 9. For all $m \in (\mu^M, \mu^B]$,

$$\lim_{\theta\to 0} L_\theta(m) = L_0(m)$$

Proof. Using the definitions of the two functions, for $m \in (\mu^M, \mu^B]$

$$L_0(m) - L_\theta(m) = \frac{1}{1-\delta} - \frac{1 - \delta\Phi\left(\frac{\ln a_\theta - m}{\theta}\right)}{1 - \delta\Phi\left(\frac{\ln a_\theta - \mu^M}{\theta}\right)} \tag{A3}$$

From Lemma 7 above, $z_\theta = \frac{\ln a_\theta - \mu^M}{\theta} \to \infty$ as $\theta \to 0$ and this implies that

$$\Phi\left(\frac{\ln a_\theta - \mu^M}{\theta}\right) \to 1$$

Also, from Lemma 7 $\theta z_\theta = (\ln a_\theta - \mu^M) \to 0$ as $\theta \to 0$. Now fix any $m \in (\mu^M, \mu^B]$. For small enough θ, $\ln a_\theta < m$ and so as $\theta \to 0$,

$$\lim_{\theta\to 0}\left(\frac{\ln a_\theta - m}{\theta}\right) = \lim_{\theta\to 0}\left(\frac{\mu^M - m}{\theta}\right) = -\infty$$

and hence

$$\Phi\left(\frac{\ln a_\theta - m}{\theta}\right) \to 0$$

Using these in (A3) completes the proof. □

Although the previous lemma established the pointwise convergence of L_θ to L_0, recall that L_0 is discontinuous at μ^M. This is because with perfect monitoring arbitrarily small price cuts can be detected although no matter how small θ is, there is a small enough price cut that is detected only with small probability. The next two lemmas study the limiting behavior of L_θ in the vicinity of μ^M, that is, precisely where L_0 is discontinuous.

Lemma 10. Let $\widehat{m}_\theta$ be the maximizer of L_θ in the interval $[\mu^M, \mu^B]$. For all $\varepsilon > 0$, there exists a $\theta(\varepsilon)$ such that for all $\theta < \theta(\varepsilon)$, (i) $\widehat{m}_\theta - \mu^M < \varepsilon$; and (ii) $\frac{\delta}{1-\delta} - L_\theta(\widehat{m}_\theta) < \varepsilon$.

Proof. (i) Suppose to the contrary that there exists a sequence $\theta(n) \downarrow 0$ such that $\lim \widehat{m}_{\theta(n)} = m' > \mu^M$. By Lemma 8, $L_{\theta(n)}$ is increasing in n, and thus the sequence $L_{\theta(n)}(\widehat{m}_{\theta(n)})$ is also increasing and so converges. We also know from that for all n, $L_{\theta(n)}(\widehat{m}_{\theta(n)}) < L_0(\widehat{m}_{\theta(n)})$ and so as $n \to \infty$,

$$\begin{aligned}\lim L_{\theta(n)}\big(\widehat{m}_{\theta(n)}\big) &\leq \lim L_0\big(\widehat{m}_{\theta(n)}\big)\\ &= L_0(m')\end{aligned}$$

Let $m'' \in (\mu^M, m')$. Then because L_0 is decreasing, we have

$$\lim L_{\theta(n)}\big(\widehat{m}_{\theta(n)}\big) < L_0(m'')$$

But because $L_{\theta(n)}(m'') \to L_0(m'')$, for large n we have $L_{\theta(n)}(m'') > L_{\theta(n)}(\widehat{m}_{\theta(n)})$. This contradicts the fact that $\widehat{m}_{\theta(n)}$ maximizes $L_{\theta(n)}$.

(ii) Again, suppose to the contrary that

$$L_\theta(\widehat{m}_\theta) \uparrow K < \frac{\delta}{1-\delta}$$

Thus, there exists $\widetilde{m} > \mu^M$ such that $L_0(\widetilde{m}) > K$. By pointwise convergence Lemma 9, $\lim L_\theta(\widetilde{m}) = L_0(\widetilde{m}) > K$. Thus for small enough θ, we have a contradiction to the fact that $\widehat{m}(\theta)$ maximized L_θ. □

Lemma 11.

$$\lim_{\theta\to 0} L_\theta(\ln a_\theta) = \frac{1}{2}\frac{\delta}{1-\delta}$$

Proof. By definition,

$$L_\theta(\ln a_\theta) = \frac{1-\delta\Phi(0)}{1-\delta\Phi\left(\frac{\ln a_\theta - \mu^M}{\theta}\right)} - \frac{\Pi_1(\ln a_\theta) - \pi^N}{\pi^M - \pi^N}$$

Now as $\theta \to 0$, $\ln a_\theta \to \mu^M$ and $\frac{\ln a_\theta - \mu^M}{\theta} = z_\theta \to \infty$ (see Appendix A). This implies that

$$\begin{aligned}\lim L_\theta(\ln a_\theta) &= \frac{1-\frac{1}{2}\delta}{1-\delta\Phi(\infty)} - 1\\ &= \frac{1}{2}\frac{\delta}{1-\delta}\end{aligned}$$

□

Appendix B: Discrete choice example

The main result of this article relies on the fact that when firms' sales are negatively correlated, aggregate sales carry more information about deviations than do individual sales. The log-normal specification then allows us to calculate the information content of aggregate sales explicitly and to bound the information content of individual sales.[16] A drawback

[16] Recall that we measure the information content via the total variation metric.

of the log-normal specification is that the term "aggregate" refers to the arithmetic mean of log sales or equivalently, the geometric mean of actual sales.

In this appendix we study a different model of demand than in the body of the article. This is the familiar discrete choice model in which the stochastic nature of firms' sales arises from the random choices of the consumers. In the discrete choice model (see Anderson, de Palma, and Thisse, 1992), outlined below, we compare two situations: one in which firms know only their own sales and the other, in which aggregate sales (now the *sum* of individual firms' sales) is publicly announced. Unlike the log-normal model in the body of the article, the discrete choice model does not allow an explicit comparison of the information content of aggregate sales versus individual sales. Here we do not carry out a complete equilibrium analysis but only compare the information contents of the two scenarios.

Specifically, there are three firms (labelled 1, 2, and 3) that produce differentiated (indivisible) products and four consumers (labelled 0, 1, 2, and 3). Each consumer i has a privately known reservation value v_i for the product and so will not purchase a product whose price exceeds this value. For consumer 0, v_0 is either v^L or v^H with probabilities $1-\varepsilon$ and ε, respectively. For consumer $i = 1, 2, 3$, v_i is again either v^L or v^H but now with probabilities ε and $1-\varepsilon$, respectively. Define

$$h_i(p) = \begin{cases} \exp(-p) & \text{if } p < v_i \\ 0 & \text{otherwise} \end{cases}$$

The probability that consumer 0 will purchase firm j's product is given by the familiar "logit" form

$$r_j^0(\mathbf{p}) = \frac{h_0(-p_j)}{h_0(-p_i) + h_0(-p_{i+1}) + h_0(-p_{i+2})}$$

We will use the convention that $\frac{0}{0} = 0$ so that if all three prices are above v_0, then for all j, $r_j^0(\mathbf{p}) = 0$.

For $i = 1, 2, 3$, consumer i has a special affinity for the product of firm i. Given prices $\mathbf{p} = (p_1, p_2, p_3)$, the probability that consumer $i = 1, 2, 3$ will purchase firm j's product is

$$r_j^i(\mathbf{p}) = \begin{cases} \frac{bh_i(-p_i)}{bh_i(-p_i)+h_i(-p_{i+1})+h_i(-p_{i+2})} & \text{if } j = i \\ \frac{h_i(-p_j)}{bh_i(-p_i)+h_i(-p_{i+1})+h_i(-p_{i+2})} & \text{otherwise} \end{cases}$$

where $b > 1$ and $i + k$ is understood to be mod 3. Again, if all three prices are above v_i, then for all j, $r_j^i(\mathbf{p}) = 0$. Thus, consumer i is biased in favor of firm i's product and $b > 1$ is a measure of the bias.

Here, as in the body of the article, firms' sales are negatively correlated and total sales are stochastic.

Consider a case where ε is small and $\frac{4}{3}v^L < v^H \leq 2v^L$. Because selling to three customers at a price of v^H is more profitable than selling to four customers at a price of v^L, the monopoly price is

$$p^M = v^H$$

and it is easily verified that firm 1's myopic best response when the other firms charge p^M is

$$p^B = v^L$$

At the monopoly price, with high probability, only customers 1, 2, and 3 buy the product, with customer i favoring firm i's product. What happens when firm 1, say, cuts its price to p^B? There are two effects. It sells to customers 2 and 3 with higher probability but more important, now also sells to customer 0 (with high probability).

When ε is small and the bias b is large, a price cut by firm 1 affects the marginal distributions of firm 2 or 3's sales only slightly. Formally, the total variation distance between the marginal distributions at monopoly prices and when firm 1 cuts its price is small. Thus, the individual sales data are rather uninformative about price cuts. The same is true of the joint distribution of the sales of firm 2 and firm 3.

In this same situation, total sales are quite informative. At monopoly prices total sales are 3 with high probability and if firm 1 cuts its price, they are 4 with high probability. The total variation distance between the two distributions of total sales is large.

□ **Numerical example.** Consider the following numerical example. Specifically, suppose that $v^L = 1$ and $v^H = 2$. For simplicity, we suppose that $\varepsilon = 0$. Further, suppose that $b = 10$.

If all firms charge the monopoly price $p^M = 2$, the marginal distribution of an individual firm's sales $Y_j = 0, 1, 2, 3, 4$ is

$f_j(Y_j \mid \mathbf{p}^M) =$

0	1	2	3	4
0.140	0.726	0.128	0.006	0

and note that with high probability a firm's sales are 1

Now suppose firm 1, say, cuts its price to p^B. At prices (p^B, p^M, p^M), the marginal distribution of sales Y_j of firm $j = 2, 3$ is

$f_j(Y_j \mid p^B, \mathbf{p}_{-1}^M) =$

0	1	2	3	4
0.243	0.680	0.075	0.002	0

The total variation distance between the two distributions is

$$\left\| f_j\left(Y_j \mid \mathbf{p}^M\right) - f_j\left(Y_j \mid p^B, \mathbf{p}^M_{-1}\right) \right\|_{TV} = 0.103$$

A similar calculation, omitted here, shows that the TV distance between the joint distributions of firm 2's and 3's sales

$$\left\| f_{2,3}\left(Y_2, Y_3 \mid \mathbf{p}^M\right) - f_{2,3}\left(Y_2, Y_3 \mid p^B, \mathbf{p}^M_{-1}\right) \right\|_{TV} = 0.204$$

The distributions of total sales are trivial. At monopoly prices $\mathbf{p}^M$, the total sales are 3 and if firm 1 cuts its price to p^B, then the total sales are 4. The total variation distance between the two distributions is 1.

Thus, in this example, the distribution of total sales is substantially more responsive to a price cut than the distribution of individual sales or the joint distribution of other firms' sales.

More generally, as b increases, the TV distance between the joint (or marginal) distributions of individual sales goes to 0. For small ε, the TV distance between the distributions of total sales is close to 1.

References

Anderson, S., de Palma, A. and Thisse, J.-F. *Discrete Choice Theory of Product Differentiation*, Cambridge, Mass.: MIT Press, 1992.

Aoyagi, M. "Collusion in Dynamic Bertrand Oligopoly with Correlated Private Signals and Communication." *Journal of Economic Theory*, Vol. 102 (2002), pp. 229–248.

Aryal, G., Ciliberto, F. and Leyden, B. "Public Communication and Collusion in the Airline Industry," SSRN Working Paper 3122560, 2018.

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

Awaya, Y. and Krishna, V. "On Communication and Collusion." *American Economic Review*, Vol. 106 (2016), pp. 285–315.

———. "Communication and Cooperation in Repeated Games." *Theoretical Economics*, Vol. 14 (2019), 513–553.

Carlton, D., Gertner, R. and Rosenfield, A. "Communication among Competitors: Game Theory and Antitrust." *George Mason Law Review*, Vol. 5 (1997), pp. 423–440.

European Commission. "Case COMP/36.212/E1—Carbonless Paper," https://ec.europa.eu/competition/elojade/isef/case_details.cfm?proc_code=1_36212, (2001a).

———. "Case COMP/36.545/F3—Amino Acids," https://ec.europa.eu/competition/elojade/isef/case_details.cfm?proc_code=1_36545, (2001b).

———. "Case COMP/38.069/E1—Copper Plumbing Tubes," https://ec.europa.eu/competition/elojade/isef/case_details.cfm?proc_code=1_38069, (2004).

———. "Guidelines on the Applicability of Article 101 of the Treaty on the Functioning of the European Union to Horizontal Co-operation Agreements." *Official Journal of the European Union*, Vol. IV(C) (2011), pp. 11–19.

Federal Trade Commission. "Information Exchange: Be Reasonable," www.ftc.gov/news-events/blogs/competition-matters/2014/12/information-exchange-be-reasonable, 2014.

Fudenberg, D., Levine, D. and Maskin, E. "The Folk Theorem with Imperfect Public Information." *Econometrica*, Vol. 62 (1994), pp. 997–1039.

Genesove, D. and Mullin, W. "Rules, Communication, and Collusion: Narrative Evidence from the Sugar Institute Case." *American Economic Review*, Vol. 91 (2001), pp. 379–398.

Green, E. and Porter, R. "Noncooperative Collusion under Imperfect Price Information." *Econometrica*, Vol. 52 (1984), pp. 87–100.

Gérard-Varet, L-A. and Moulin, H. "Correlation in Duopoly." *Journal of Economic Theory*, Vol. 19 (1978), pp. 123–139.

Harrington, J. "How do Cartels Operate?" *Foundations and Trends in Microeconomics*, Vol. 2 (2006), pp. 1–105.

———. "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices." *American Economic Review*, Vol. 101 (2011), pp. 2425–2449.

Japan Fair Trade Commission. "Information on Production Volume and Demand Trend [Group guidelines 9-1, 9-4]," www.jftc.go.jp/dk/soudanjirei/johokatudo/johokatsudo01.html, 1999.

———. "Collection and Provision of Information by Business Organizations," www.jftc.go.jp/dk/soudanjirei/h28/h27nendomokuji/h27nendo09.html, 2015.

Kandori, M. "The Use of Information in Repeated Games with Imperfect Monitoring." *Review of Economic Studies*, Vol. 59 (1992), pp. 581–593.

———. "Randomization, Communication, and Efficiency in Repeated Games with Imperfect Public Monitoring." *Econometrica*, Vol. 71 (2003), pp. 345–353.

Kaplow, L. *Competition Policy and Price Fixing*, Princeton: Princeton University Press, 2013.

Marshall, R. and Marx, L. "Explicit Collusion and Market Share Allocations," http://faculty.fuqua.duke.edu/∼marx/bio/papers/marketshares.pdf, 2008.

———. *The Economics of Collusion*. Cambridge, Mass.: MIT Press, 2012.

Matsushima, H. "Efficiency in Repeated Games with Imperfect Monitoring." *Journal of Economic Theory*, Vol. 48 (1989), pp. 428–442.

Moulin, H. and Vial, J.-P. "Strategically Zero-Sum Games: The Class of Games Whose Mixed Strategy Equilibria Cannot be Improved Upon." *International Journal of Game Theory*, Vol. 7 (1979), pp. 201–221.

Moulin, H., Ray, I. and Sen Gupta, S. "Improving Nash by Coarse Correlation." *Journal of Economic Theory*, Vol. 150 (2014), pp. 852–865.

Radner, R., Myerson, R. and Maskin, E. "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria." *Review of Economic Studies*, Vol. 53 (1986), pp. 59–69.

Stigler, G. "A Theory of Oligopoly." *Journal of Political Economy*, Vol. 74 (1964), pp. 44–61.

Sugaya, T. and Wolitzky, A. "Maintaining Privacy in Cartels." *Journal of Political Economy*, Vol. 126 (2018), pp. 2569–2607.