INTERNATIONAL ECONOMIC REVIEW
Vol. 48, No. 2, May 2007

COLLUSION, FLUCTUATING DEMAND, AND PRICE RIGIDITY*

BY MAKOTO HANAZONO AND HUANXING YANG[1]

Nagoya University, Japan; Ohio State University, U.S.A.

We study an infinitely repeated Bertrand game in which an i.i.d. demand shock occurs in each period. Each firm receives a private signal about the demand shock at the beginning of each period. At the end of each period, all information but the private signals becomes public. We consider the optimal symmetric perfect public equilibrium (SPPE) mainly for patient firms. We show that price rigidity arises in the optimal SPPE if the accuracy of the private signals is low. We also study the implications of more firms and firms' impatience on collusive pricing.

1. INTRODUCTION

Empirical evidence shows that prices in oligopolistic markets tend to be more rigid than prices in competitive markets (Dixon, 1983; Carlton, 1986). This suggests that collusion may result in firms' inflexible pricing behavior. It is not obvious, however, that collusion causes price rigidity. Why do colluding firms not adjust prices according to changing environments, by which they can potentially earn higher profits? What limits colluding firms' ability to coordinate, thereby leading to price rigidity?

In this article, we study how information asymmetry among firms limits colluding firms' ability to respond to *demand* shocks, and characterize the conditions under which prices are rigid. Specifically, we consider private information about demand in an infinitely repeated Bertrand game with two identical firms, in which an i.i.d. binary demand shock occurs in each period. Each firm receives a conditionally independent private signal about the underlying demand state at the beginning of each period, and then charges a price. The accuracy of private signals is measured by the probability of receiving a high (respectively, low) signal conditional on a high (respectively, low) underlying demand state. Therefore, the lower the accuracy, the higher the degree of informational asymmetry. At the end of each period, firms observe the underlying demand and the prices, but they never observe their rivals' signals. We adopt symmetric perfect public equilibria (SPPE) as our solution concept, which requires that, at each period, firms' pricing schemes

* Manuscript received August 2005; revised May 2006.

[1] The authors are grateful to Masaki Aoyagi, Kyle Bagwell, Shinsuke Kambe, George Mailath, Steven Matthews, Ichiro Obara, Rafael Rob, Tadashi Sekiguchi, two anonymous referees, and the seminar participants at Gakushuin, Keio, Kyoto, Ohio State, Osaka, Penn, Tohoku, and Tokyo for helpful discussions and valuable comments. Please address correspondence to: Huanxing Yang, Department of Economics, Ohio State University, 405 Arps Hall, 1945 N. High Street, Columbus, OH 43210, U.S.A. E-mail: *yang.1041@osu.edu.*

be symmetric after any public history.[2] Collusive firms are supposed to play an optimal SPPE, which yields the highest equilibrium payoff. We mainly deal with patient firms, so that firms have no incentive to commit the *off-schedule* deviations of charging prices that are not assigned for any signal.[3]

Firms can adopt two possible pricing schemes: sorting schemes in which firms charge different prices for different signals, or pooling (price rigidity) schemes in which they charge the same price regardless of signals. Under sorting schemes, firms potentially can exploit the information contained in private signals, and thus enjoy an *informational gain*. However, to support a sorting scheme in equilibrium, each firm should have no incentive to commit the *on-schedule* deviation where a firm receiving a specific signal charges a price that is assigned for another signal. Since this deviation cannot be detected, price distortions or future punishments on the equilibrium path are necessary to deter such deviation. This causes a loss in the equilibrium payoff that we call *coordination costs*. On the other hand, under pooling schemes firms incur no coordination costs since there is only one price on the equilibrium path and thus firms have no opportunity to undertake on-schedule deviation. At the same time, firms enjoy no informational gain since they always charge the same price regardless of signals. As a result, which scheme is optimal depends on the magnitude of the coordination costs relative to that of the informational gain.

Our first result is that if the accuracy of private signals is low, a pooling scheme is optimal; thus price rigidity optimally arises on the equilibrium path. Intuitively, as the accuracy of signals decreases, while the informational gain of a sorting scheme decreases, the coordination costs increase. Roughly speaking, the less accurate the private signals, the more likely that miscoordination will occur on the equilibrium path, leading to bigger coordination costs. Therefore, a pooling scheme becomes relatively more profitable as the accuracy of private signals decreases. Our second result is that when the accuracy of private signals is close to perfect, some sorting scheme strictly dominates any price rigidity scheme. This is because when signals are nearly perfect, coordination costs vanish but informational gain becomes large; thus some sorting scheme is optimal.

These results contribute to our understanding of which industries, and under what conditions, should exhibit rigid prices. Our model predicts that a collusive industry with less predictable demand is more likely to exhibit rigid prices, and a collusive industry with highly predictable demand changes prices more frequently. This prediction is consistent with an empirical finding by Weiss (1993). These results also have macroeconomic implications. For example, as the aggregate demand becomes more difficult to predict, the demand of each industry becomes more difficult to predict as well; as a result, prices in collusive industries tend to be more rigid. This also implies that shocks in aggregate demand may impact the relative prices between collusive industries and competitive industries.

Our third result is that the price war behavior under the optimal sorting scheme exhibits different patterns from those in Green and Porter (1984) and Rotemberg

[2] The above symmetry requirement implies that firms share future punishments or rewards together. This solution concept is adopted by Abreu et al. (1991) and Athey et al. (2004).

[3] Any off-schedule deviation is always detected and thus can be deterred by imposing severe future punishments, which do not affect the equilibrium payoff.

and Saloner (1986).[4] In our model only the price for the low signal is distorted downward (to deter on-schedule deviation), whereas Rotemberg and Saloner predict that only the price in higher demand states needs to be distorted downward (to deter off-schedule deviation). In Green and Porter, future equilibrium punishments are only triggered in low demand states, whereas in our optimal sorting scheme they are only triggered in high demand states. The general conclusion is that price war behavior is sensitive to the information structure and firms' patience level.

Our model can be extended to the setting of more firms, with the results qualitatively unchanged. Interestingly, in contrast to a common wisdom in industrial organization, we find that as the number of firms increases from two to three in a collusive industry, a price rigidity scheme is more likely to be optimal. Intuitively, as the number of firms increases, coordination among firms becomes more difficult, which reduces the profitability of sorting schemes relative to pooling schemes.

Our model is closely related to Athey et al. (2004, ABS henceforth), who provide the first rigorous explanation about the relationship between collusion and price rigidity in a repeated game framework when firms receive private *cost* shocks. Although our model shares several modeling features with ABS, there are reasons why it is of independent interest to study the relationship between collusion and price rigidity in the presence of demand shocks. First, empirical evidence suggests that price rigidity with respect to cost shocks is different from that with respect to demand shocks, both in terms of magnitude and in relations to concentration.[5] Second, cost shocks and demand shocks affect firms' payoffs differently; whereas individual cost shocks only affect an individual firm's own profit, demand shocks commonly affect all firms' profits. This leads to another difference: Private information in our model is inherently correlated across firms, whereas it is independent in ABS. Moreover, ABS mainly deal with the case of inelastic demand, whereas elastic demand is essential for our model.

These differences generate a distinction on the role of future punishments; in our setting, a sorting scheme with future equilibrium punishments might be optimal, whereas in ABS future equilibrium punishments have no value.[6] One advantage to our approach lies in comparative statics. Our model predicts that the rigidity of prices depends on the predictability about demand. Moreover, a price rigidity scheme is more likely to be optimal as the number of firms increases in a collusive industry. These predictions make our model more conducive to empirical testing than ABS. In ABS, whether prices are rigid depends on the cost distribution, which

[4] Both papers analyze repeated interactions of firms in an oligopolistic market in which the demand state stochastically fluctuates over time. However, there is no private information in either model.

[5] Geroski (1992) and Weiss (1993) find that prices are more rigid with respect to demand shocks than to cost shocks. According to Weiss, price inertia with respect to demand shocks is three times bigger than that with respect to cost shocks. Weiss also finds that although there is a positive correlation between concentration and price rigidity with respect to cost shocks, this relationship does not exist with respect to demand shocks.

[6] This is because, with independent cost shocks and inelastic demand, any future equilibrium punishment can be replaced by the corresponding price distortion to sustain the on-schedule constraint without affecting the equilibrium value. With correlated private information and elastic demand, this property does not hold in our setting, since future equilibrium punishments and price distortions now affect the equilibrium value differently.

has little economic interpretation and is difficult to carry out comparative static analysis.[7]

The rest of this article is organized as follows. Section 2 sets up the model. In Section 3, we derive the optimal pooling scheme and the optimal sorting scheme. Section 4 compares the two kinds of optimal schemes, and the equilibrium price war behavior is summarized. The case of impatient firms is also briefly studied. Section 5 extends the basic model. We offer discussions in Section 6 and conclusions in Section 7.

2. THE MODEL

Primitives. Consider two firms that play an infinitely repeated Bertrand game with homogenous products. In each period, each firm charges a price and the firm charging the lower price wins the whole market. If two firms charge the same price, they share the market equally. We assume that firms have the same marginal cost of production, which is constant and normalized to be 0.

In each period, the underlying demand can be either high or low. Let $S \in \{H, L\}$ denote the demand state and assume that each state arises equally likely, that is, $\Pr(H) = \Pr(L) = 0.5$. The demand function for state S is denoted by $D^S(p)$, and we assume the following conditions: $D^H(p) > D^L(p)$ for every p, and $D^H(p)$ and $D^L(p)$ are both downward-sloping, differentiable, and not too convex so that the industry profit functions $pD^S(p)$, $S \in \{H, L\}$, are strictly concave. Let $\pi^S(p) = pD^S(p)$, $S \in \{H, L\}$, denote the total profit of the industry when the (lower) price is p and the underlying demand is S. Assume that the industry profit-maximizing price for the high state is higher than that for the low state, that is, $\arg\max \pi^H(p) \equiv p^H > p^L \equiv \arg\max \pi^L(p)$. Suppose also that some $\bar{p} > 0$ exists such that $D^S(\bar{p}) \leq 0$ for $S \in \{H, L\}$. Then we can restrict the set of prices to the closed interval $[0, \bar{p}]$.

Demand is i.i.d. across periods. At the beginning of each period, firms do not observe the realized demand state of that period. Instead, each firm receives a private signal $s_i \in \{h, \ell\}$ about the underlying demand. The distribution of signals is conditionally independent and

$$\Pr(h \mid H) = \Pr(\ell \mid L) = \lambda; \quad \Pr(\ell \mid H) = \Pr(h \mid L) = 1 - \lambda,$$

where $\lambda \in (0.5, 1]$. By Bayes' rule, note that

$$\Pr(H \mid h) = \Pr(L \mid \ell) = \lambda; \quad \Pr(L \mid h) = \Pr(H \mid \ell) = 1 - \lambda.$$

This information structure captures the following business practice; nearly all important decisions made by firms require predictions about uncertain future demand, about which they usually have imperfect information. Moreover, firms might have different predictions about future demand.

[7] Another related paper is Stiglitz (1984). Focusing on a particular equilibrium, he argues that coordination costs due to informational asymmetry about demand may result in price rigidity. His approach is limited since the equilibrium he focused on is not an optimal SPPE.

The parameter λ captures the accuracy of the signals; a lower λ implies that future demand is more difficult to predict.[8] Since $\lambda \in (0.5, 1]$, the signals are informative. An increase in λ improves the informativeness of signals in the sense of Blackwell (1953). After each firm receives its private signal, it charges a price, which remains fixed during the period. At the end of the period, both firms observe the underlying demand and prices of that period. However, each firm does not observe the other firm's signal.

We first consider the stage game. A pricing schedule for firm i is a function $p_i(s_i)$, a mapping from the set of private signals, $\{h, \ell\}$, to the set of prices. When firm i ($-i$, respectively) charges p_i(p_{-i}, respectively) and the demand state is S, the profit for firm i is

$$\pi_i(p_i, p_{-i}; S) = \pi^S(\min\{p_i, p_{-i}\})m_i(p_i, p_{-i});\;\; m_i(p_i, p_{-i}) = \begin{cases} 0 & \text{if } p_i > p_{-i}, \\ 0.5 & \text{if } p_i = p_{-i}, \\ 1 & \text{if } p_i < p_{-i}. \end{cases}$$

Thus, the (ex ante) expected stage payoff for firm i is

$$u_i(p_i(\cdot), p_{-i}(\cdot)) = E[\pi_i\,(p_i\,(s_i)\,,\, p_{-i}(s_{-i}); S)].$$

Since the marginal costs are 0, as we know from the standard Bertrand game with homogenous products, the stage game has a unique Nash equilibrium: Both firms charge price 0 regardless of signals and earn zero profit in every period.[9] Note that the stage Nash equilibrium is independent of the demand states and private signals.

The Repeated Game. Firms repeat the stage game infinitely often. In any period, each firm chooses a pricing schedule depending on past history, which consists of a sequence of realized demand states, a sequence of prices (*public* history), and a sequence of signals the firm has received (*private* history). We focus on public strategies, that is, plans of stage-game pricing schedules that depend only on public histories. A profile of public strategies induces a distribution over the sequence of pricing schedules $\{p_1^t(\cdot), p_2^t(\cdot)\}_{t=0}^{\infty}$. Each firm maximizes the discounted sum of stage-game profits with a discount factor $\delta < 1$. Given a strategy profile, firm i's expected payoff (after normalization by multiplying $1-\delta$) is

$$E\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t u_i\left(p_i^t(\cdot), p_{-i}^t(\cdot)\right)\right].$$

[8] Empirically speaking, demand predictability is associated with a component of detrended demand fluctuations that cannot be predicted by an adaptive expectation process. Demand is said to be more difficult to predict if this component becomes larger.

[9] Alternatively, we can consider a different stage game in which firms compete in a market with differentiated products (e.g., Hotelling's linear city model) and common demand shocks. Unlike the homogeneous-product case, the stage Nash equilibrium for the differentiated-products case is sorting: Each firm charges different prices for different signals. Our main results apply qualitatively to this differentiated-products setting.

We adopt perfect public equilibria (PPE) as our solution concept. A PPE is a profile of public strategies for which no one-shot deviation is profitable for each firm at any point of time. We further restrict attention to *symmetric* PPE (SPPE), i.e., PPE for which firms' pricing schedules are symmetric after any public history. Intuitively, SPPE implies that, after any history, all firms together enjoy the same future rewards or suffer the same future punishments.

To investigate SPPE, we adopt the *APS approach* (Abreu et al., 1986, 1990). APS establish that, for each PPE, the equilibrium payoff, or the *value,* can be expressed as a recursive formula subject to the incentive compatibility (IC) constraints that ensure no (one-shot) deviation. Accordingly, we can decompose PPE values into the stage-game profits and the continuation values $\mathbf{v}(\cdot) = (v_1(\cdot), v_2(\cdot))$, where $\mathbf{v}$ map from each *public outcome*, $(p_1, p_2; S)$, to PPE payoffs. In addition, SPPE requires $p_1(\cdot) = p_2(\cdot) = p(\cdot)$ and $v_1(\cdot) = v_2(\cdot) = v(\cdot)$ with $v(p, p'; S) = v(p', p; S)$. Let $V_s^*(\lambda, \delta) \subset R$ denote the set of SPPE payoffs. It is nonempty since repetition of the stage-game Nash equilibrium is an SPPE, so $0 \in V_s^*(\lambda, \delta)$. We assume that firms have access to a public randomization device at the end of each period; thus $v(p_1, p_2; S)$ can be a probability distribution over $V_s^*(\lambda, \delta)$. Consequently, $V_s^*(\lambda, \delta)$ must be convex. By slightly modifying the argument of Abreu et al. (1986, 1990) for characterizing the PPE payoff set, we can show that $V_s^*(\lambda, \delta)$ is a closed interval of the form $[0, \bar{v}(\lambda, \delta)]$.[10]

Search for the Most Collusive SPPE. We focus on the most collusive SPPE (or an optimal SPPE), that is, an SPPE with the highest ex ante payoff $\bar{v}(\lambda, \delta)$. If firms have an opportunity to coordinate before playing the repeated game, it is natural for them to select this particular equilibrium among all SPPE. We are particularly interested in firms' pricing behavior that achieves $\bar{v}(\lambda, \delta)$.

The ex ante problem. Given the public randomization device, any SPPE payoff can be achieved by the following "trigger" strategies with the *bang-bang property* (see APS, 1990): Firms begin with playing a stage game strategy $p(\cdot)$, then after observing the public outcome $(p_i, p_{-i}; S)$, with probability $1 - \alpha(p_i, p_{-i}; S)$ they go back to the beginning (i.e., continue to play the same strategy $p(\cdot)$ in the next stage), and with probability $\alpha(p_i, p_{-i}; S)$ they go to perpetual Nash reversion. We can thus find an optimal SPPE payoff among such trigger strategy equilibria. The following problem then characterizes the optimal SPPE payoff:

$$\bar{v}(\lambda, \delta) = \max_{v, p(\cdot), \alpha(\cdot)} v$$

$$\begin{aligned} &\text{subject to} \\ &v = (1-\delta)u_i(p(\cdot), p(\cdot)) + \delta v E[1 - \alpha(p(s_i), p(s_{-i}); S)] \\ &v \geq (1-\delta)u_i(\tilde{p}_i(\cdot), p(\cdot)) + \delta v E[1 - \alpha(\tilde{p}_i(s_i), p(s_{-i}); S)], \\ &\quad \alpha(p, p'; S) = \alpha(p', p; S) \in [0, 1], \\ &\quad \forall \tilde{p}_i(\cdot),\ \forall (p, p'; S) \in [0, \bar{p}]^2 \times \{H, L\}. \end{aligned} \tag{1}$$

[10] The proof is available upon request.

The interim problem. The IC constraint (1) is formulated from the ex ante viewpoint, that is, before firms receive private signals. We reformulate this constraint as a system of incentive constraints from the interim viewpoint, that is, after firms receive private signals. These constraints require that each firm have an incentive to charge the price assigned for the received signal.

For the subsequent analysis, we classify these incentive constraints into two categories:[11] (i) *Off-schedule* constraints; each firm has no incentive to charge an off-schedule price, that is, a price assigned for neither signal. If an off-schedule price is observed, firms can immediately tell that deviation has occurred. (ii) *On-schedule* constraints; each firm receiving a specific signal must prefer charging the price assigned for that signal to charging the price assigned for the other signal. On-schedule deviations cannot be detected, since signals are privately observed.

Accordingly, we rewrite the problem from the interim viewpoint as

$$\text{(P)}\qquad \max_{v,p(\cdot),\alpha(\cdot)} \; v = \{v^h + v^\ell\}/2$$

subject to

$$v^k = (1-\delta)E[\pi_i(p(k), p(s_{-i}); S) \mid k] + \delta v E[1 - \alpha(p(k), p(s_{-i}); S) \mid k]$$

$$\text{(DIC)}\qquad v^h \geq (1-\delta)E[\pi_i(p(\ell), p(s_{-i}); S) \mid h] + \delta v E[1 - \alpha(p(\ell), p(s_{-i}); S) \mid h],$$

$$\text{(UIC)}\qquad v^\ell \geq (1-\delta)E[\pi_i(p(h), p(s_{-i}); S) \mid \ell] + \delta v E[1 - \alpha(p(h), p(s_{-i}); S) \mid \ell],$$

$$(\text{Off}_k)\qquad v^k \geq (1-\delta)E[\pi_i(\tilde{p}_i, p(s_{-i}); S) \mid k] + \delta v E[1 - \alpha(\tilde{p}_i, p(s_{-i}); S) \mid k],$$

$$\alpha(p, p'; S) = \alpha(p', p; S) \in [0, 1],$$

$$\forall \tilde{p}_i \neq p(k),\ \forall (p, p'; S) \in [0, \bar{p}]^2 \times \{H, L\},\ k = h, \ell.$$

The interim equilibrium payoff v^k, $k = h, \ell$, is defined in the first line of the constraints. Since the ex ante probability that each firm receives a particular signal is 1/2, the value of an SPPE is $\{v^h + v^\ell\}/2$. The two constraints following the definition of v^k are the on-schedule constraints: the *downward incentive constraint* (DIC), which deters a firm receiving a high signal from charging the price assigned

[11] These classification and terms follow ABS (2004).

for a low signal, and the *upward incentive constraint* (UIC), which is defined similarly. The constraints marked as (Off $_k$) are the off-schedule constraints: Charging any price $\tilde{p}_i$ other than $p(h)$ and $p(\ell)$, is not profitable.

3. COLLUSION AMONG PATIENT FIRMS

In this section, we analyze the properties of solutions to (P). Note that, if the discount factor δ is sufficiently high, off-schedule deviations can be deterred by perpetual (off-path) Nash reversion, which causes no equilibrium value loss. For the time being, we consider such patient firms; the remaining effective incentive constraints are, therefore, the on-schedule ones. Impatient firm cases will be discussed in the next section.

3.1. *Two Classes of Pricing Schemes.* Firms can possibly adopt two classes of pricing schemes: pooling schemes or sorting schemes. Under a pooling scheme, firms charge the same price regardless of the signals, that is, $p(h) = p(\ell)$. Under a sorting scheme, firms charge different prices when they receive different signals, namely, $p(h) \neq p(\ell)$. The key difference between these schemes is that, on-schedule constraints are redundant for pooling schemes, but are relevant for sorting schemes. This is because there is only one price on the equilibrium path if firms adopt a pooling scheme, so all deviations are off-schedule. In contrast, there are two prices on the equilibrium path if firms adopt a sorting scheme, so on-schedule deviations are possible.

Using technical terms in repeated games, we can concisely rephrase the above key difference between the two schemes. Namely, *whether monitoring is perfect depends on which type of pricing scheme is adopted*. When firms adopt a pooling scheme, monitoring is perfect because any deviation can be detected for sure. However, when firms adopt a sorting scheme, monitoring becomes imperfect, because each firm cannot observe its rival's signals and pricing schedule.

We are mainly interested in under what conditions an optimal solution to (P) exhibits sorting or pooling pricing. To proceed, we consider pooling and sorting in turn.[12]

3.2. *Pooling Pricing Scheme.* Under pooling, the on-schedule constraints are redundant. This implies that there is no need to impose future punishments on the optimal equilibrium path. The optimal pooling price must maximize the ex ante stage game profits, namely,

$$p^{r*} = \arg\max_p \{\pi^H(p) + \pi^L(p)\}/2. \tag{2}$$

[12] Treating both schemes at the same time is in fact inconvenient. This is because the deviation gain exhibits discontinuity at points where the price for a high signal equals that for a low signal. This discontinuity in the programming problem creates analytical inconvenience.

Since each firm receives a half of the industry-wide profits every period, the equilibrium value is

$$v^r = \{\pi^H(p^{r*}) + \pi^L(p^{r*})\}/4. \tag{3}$$

Note that p^{r*} and v^r are independent of the accuracy of private signals, λ.

3.3. *Sorting Pricing Scheme*

Effective Constraints. Define $(p^{hm}, p^{\ell m})$ as the prices that maximize the ex ante stage payoff $(\pi^h + \pi^\ell)/2$, or those that capture the full informational gain, namely,

$$p^{hm} = \arg\max_{p^h \geq 0} \pi^h; \quad p^{\ell m} = \arg\max_{p^\ell \geq 0} \pi^\ell.$$

Naturally, for $\lambda \in (0.5, 1)$, we have $p^H > p^{hm} > p^{\ell m} > p^L$. We can show the following properties of optimal sorting schemes.

First, a sorting SPPE with $\tilde{p}(h) < \tilde{p}(\ell)$ is suboptimal. To see this, consider the following sorting profile $\{\tilde{p}(h), \tilde{p}(\ell)\}$ with $p^{hm} > \tilde{p}(\ell) > \tilde{p}(h) > p^{\ell m}$. By the concavity of the profit functions, the stage-game profit can be increased by shifting $\tilde{p}(h)$ and $\tilde{p}(\ell)$ toward p^{hm} and $p^{\ell m}$, respectively. Then there exists a pooling SPPE with $p \in (\tilde{p}(h), \tilde{p}(\ell))$ that dominates the initial sorting SPPE, since the pooling SPPE can generate a higher current profit without any on-path future punishments.

Second, the UIC must be slack under an optimal scheme. If a firm receiving a signal ℓ charges $p(h)$ $(>p(\ell))$, its expected market share and its stage payoff decrease. A firm would have no incentive to make such a deviation unless the future punishments that the firm could avoid would be large. Such on-path future punishments are excessive for an optimal scheme.[13]

Third, the cases in which the DIC is slack are trivial. This is because, if the DIC is also slack at an optimal solution, sorting entails no loss from the on-schedule constraint, but earns the full informational gain. To make our problem interesting, we impose the following condition that guarantees the DIC always binds for *all* λ in optimal sorting schemes.

ASSUMPTION 1. $\pi^H(p^L) \geq \pi^H(p^H)/2$.

Recall that $\arg\max \pi^H(p) \equiv p^H > p^L \equiv \arg\max \pi^L(p)$. Assumption 1 says that p^H and p^L are fairly close to each other, which implies that p^{hm} and $p^{\ell m}$ are also fairly close to each other for any λ. Under the sorting scheme $\{p^{hm}, p^{\ell m}\}$, which captures the full informational gain, a firm receiving a signal h is thus tempted to charge $p^{\ell m}$. Thus under Assumption 1 the DIC is effective.[14]

[13] If the UIC binds for a sorting SPPE with $p(h) > p(\ell)$, it can be shown straightforwardly that a pooling SPPE with $p^r = p(h)$ and no on-path future punishment generates a higher equilibrium value.

[14] Proofs of the above three claims are available upon request.

By the above discussions, we can, therefore, drop the UIC, and impose $p(h) > p(\ell)$ and a binding DIC. The problem for an optimal sorting SPPE becomes:[15]

$$
\begin{aligned}
&\text{(P–S)} \qquad \max_{v,p(\cdot),\alpha(\cdot)} v = \{v^h + v^\ell\}/2 \\
&\qquad \text{subject to} \\
&\qquad v^k = (1-\delta)E[\pi_i(p(k), p(s_{-i}); S) \mid k] \\
&\qquad\qquad + \delta v E[1 - \alpha(p(k), p(s_{-i}); S) \mid k], \quad k = h, \ell, \\
&\text{(DIC)} \qquad v^h = (1-\delta)E[\pi_i(p(\ell), p(s_{-i}); S) \mid h] \\
&\qquad\qquad + \delta v E[1 - \alpha(p(\ell), p(s_{-i}); S) \mid h], \\
&\qquad \alpha(p, p'; S) = \alpha(p', p; S) \in [0, 1], \\
&\qquad \forall (p, p'; S) \in \{p(h), p(\ell)\}^2 \times \{H, L\}, \\
&\qquad \bar{p} \geq p(h) > p(\ell) \geq 0.
\end{aligned}
$$

Notations. We hereafter use the following simplified notations for convenience:

$$
\begin{aligned}
\pi^k &= E[\pi_i(p(k), p(s_{-i}); S) \mid k], \quad p^k = p(k), \quad k = h, \ell, \\
\pi_d^h &= E[\pi_i(p(\ell), p(s_{-i}); S) \mid h] \\
\alpha_{ij}^S &= \alpha(p^i, p^j; S), (i, j, S) \in \{h, \ell\}^2 \times \{H, L\}.
\end{aligned}
$$

Given p^h and p^ℓ, we can calculate the relevant stage payoffs as follows:

$$
\begin{aligned}
\pi^h &= \frac{\lambda^2}{2}\pi^H(p^h) + \frac{(1-\lambda)^2}{2}\pi^L(p^h); \\
\pi^\ell &= \frac{(1-\lambda)(1+\lambda)}{2}\pi^H(p^\ell) + \frac{\lambda(2-\lambda)}{2}\pi^L(p^\ell); \\
\pi_d^h &= \frac{\lambda(1+\lambda)}{2}\pi^H(p^\ell) + \frac{(1-\lambda)(2-\lambda)}{2}\pi^L(p^\ell).
\end{aligned}
$$

These prices and profits certainly depend on the level of signal accuracy, λ. For notational simplicity, we drop λ from them unless otherwise noted.

Optimal Future Punishments and Price Distortions. Firms have two instruments to satisfy the DIC: price distortions and future equilibrium punishments. By distorting p^ℓ downward from $p^{\ell m}$, the stage game deviation gain can be lowered.[16] By increasing the probabilities of triggering perpetual Nash reversion, the continuation value can be lowered for a deviating firm. Although these instruments can

[15] This problem may not have a solution, since the constraint set is not compact. This is not a problem, however. In the case where no solution exists to (P–S), the original problem (P) must have a pooling pricing scheme as its optimal solution.

[16] Distorting p^h from p^{hm} is harmful for the DIC.

decrease the temptations for cheating, they cause equilibrium value loss at the same time. Our goal is to identify the optimal pricing and punishment scheme.

Consider a feasible solution $(p^h, p^\ell, v, (\alpha_{ij}^S)_{S\in\{H,L\}, ij\in\{h,\ell\}})$ to (P–S). With our new notations, the value of this profile is

$$v = \frac{1-\delta}{2}\{\pi^h + \pi^\ell\} + \delta v \sum_{\substack{i,j\in\{h,\ell\}\\ S\in\{H,L\}}} \Pr(ij;S)\big(1-\alpha_{ij}^S\big),$$

where $\Pr(ij;S) = \Pr(S)\Pr(i \mid S)\Pr(j \mid S)$, the probability that firm 1 receives signal i, firm 2 receives signal j, and the underlying demand is S. The DIC holds with equality, that is,

$$(1-\delta)\big\{\pi_d^h - \pi^h\big\} = \delta v \sum_{S,j} \Pr(ij;S \mid i=h)\big(\alpha_{\ell j}^S - \alpha_{hj}^S\big).$$

Assume, for the time being, $\alpha_{ij}^S > 0$ for some (i, j, S). Define

$$LR = \frac{\sum_{S,j}\Pr(hj;S)\alpha_{\ell j}^S + \sum_{S,j}\Pr(\ell j;S)\alpha_{\ell j}^S}{\sum_{S,j}\Pr(hj;S)\alpha_{hj}^S + \sum_{S,j}\Pr(\ell j;S)\alpha_{\ell j}^S}.$$

The denominator is the ex ante probability of triggering perpetual Nash reversion on the equilibrium path. This probability measures the equilibrium value loss. The numerator is the ex ante probability of triggering perpetual Nash reversion if a firm deviates to charging p^ℓ after receiving a signal h. This probability thus represents future punishments off the equilibrium path. The likelihood ratio LR measures how effective the future punishments are in deterring deviation per unit of value loss on the equilibrium path. A large LR implies that the trigger probabilities impose heavy punishments on the deviator per unit of equilibrium value loss.

Using the likelihood ratio LR, we reformulate the DIC as

$$\begin{aligned}
(1-\delta)\big\{\pi_d^h - \pi^h\big\} &= \frac{\delta v}{\Pr(i=h)} \sum_{S,j} \Pr(hj;S)\big(\alpha_{\ell j}^S - \alpha_{hj}^S\big) \\
&= \frac{\delta v}{\Pr(i=h)} \times (LR-1)\left\{\sum_{S,j}\Pr(hj;S)\alpha_{hj}^S + \sum_{S,j}\Pr(\ell j;S)\alpha_{\ell j}^S\right\} \\
&= 2\delta v(LR-1)\sum_{S,i,j}\Pr(ij;S)\alpha_{ij}^S.
\end{aligned}$$

Suppose $LR \neq 1$. Substituting this formula into the value function, we obtain

$$\text{(AMP)} \qquad v = \frac{1}{2}\left\{\pi^h - \frac{\pi_d^h - \pi^h}{LR-1}\right\} + \frac{1}{2}\pi^\ell.$$

This formula incorporates the binding DIC into the value function. The term of $\{\pi^h + \pi^\ell\}/2$ is the expected profit from the cooperative phase. The term of

$\{\pi_d^h - \pi^h\}/\{LR - 1\}$ is the expected loss resulting from on-path future punishments. This is indeed a modified version of the formula developed by Abreu et al. (1991, henceforth AMP).[17,18]

To justify the modified AMP formula at optimum, we show the following lemma.

LEMMA 1. *Suppose that Assumption 1 holds*, $\lambda > 0.5$, *and that a solution to (P–S) exists. For each solution, we must have* $\alpha_{ij}^S > 0$ *for some* (i, j, S), *and* $LR > 1$.

PROOF. See the Appendix.

This lemma shows that using price distortion alone to satisfy the DIC is suboptimal; in any optimal sorting SPPE, $\pi_d^h > \pi^h$ holds, and future equilibrium punishments are necessary. These punishments cause equilibrium value loss, which is one component of the coordination costs.

According to the modified AMP formula, an optimal future punishment scheme maximizes LR subject to the binding DIC. This is equivalent to finding the maximum likelihood test to detect deviation. To attain a higher power of test, we need to find a public outcome that is more likely to occur off the equilibrium path relative to its likelihood on the equilibrium path. For each public outcome (i, j, S), the power of test, LR_{ij}^S, is defined by the likelihood ratio when future punishments are triggered only after either (i, j, S) or (j, i, S) is observed.[19] For instance,

$$LR_{\ell\ell}^H = \frac{\Pr(h\ell; H)\alpha_{\ell\ell}^H + \Pr(\ell\ell; H)\alpha_{\ell\ell}^H}{\Pr(\ell\ell; H)\alpha_{\ell\ell}^H} = \frac{\lambda(1-\lambda) + (1-\lambda)^2}{(1-\lambda)^2} = \frac{1}{1-\lambda}.$$

Similarly,

$$LR_{\ell\ell}^L = \frac{1}{\lambda}, \quad LR_{h\ell}^L = \frac{1}{2\lambda}, \quad LR_{h\ell}^H = \frac{1}{2(1-\lambda)}.$$

The public outcome (ℓ, ℓ, H) has the highest power of test.[20] Note that the likelihood ratio LR is a convex combination of the powers of test, LR_{ij}^S . This

[17] AMP consider repetition of the following partnership game with imperfect public monitoring: If both players cooperate, each player receives π; if only one player cooperates, the deviator receives $\pi + g$; if both players defect, the stage payoff is 0 for each player. AMP show that, if the incentive constraint holds with equality, the value of an SPPE, v, can be written as $v = \pi - g/(l - 1)$, where l is the likelihood ratio of the future punishment scheme.

[18] AMP's model is simpler than ours: It has no in-stage private information, and each player's action space is binary. In AMP, the payoff from the cooperative phase and the stage-game deviation gain are exogenously given. In contrast, due to the continuous action space, those payoffs in our model are endogenously determined. Due to private information, the monitoring structure is also endogenous in our model, as we have discussed earlier, whereas the monitoring structure in AMP is exogenous.

[19] Note that, by the symmetry requirement, $\alpha_{ij}^S = \alpha_{ji}^S$ is used in calculating LR_{ij}^S.

[20] This result depends on our restriction to symmetric equilibria: After any public outcome, both firms' continuation values must be the same. Because of this restriction, after observing the public outcome (h, ℓ, H), the firm charging p^h cannot be rewarded. This is why (ℓ, ℓ, H) has a higher power of test than (ℓ, h, H) does.

implies the following two important facts. First, *to optimize a future punishment scheme, (i) first only set* $\alpha^H_{\ell\ell} > 0$ *and (ii) then set* $\alpha^S_{ij} > 0$ *for* $(i, j, s) \neq (\ell, \ell, H)$ *only if setting* $\alpha^H_{\ell\ell} = 1$ *is insufficient to deter deviation.* Second, *LR attains the least upper bound of* $1/(1-\lambda)$ *iff* $\alpha^S_{ij} > 0$ *only for* $(i, j, S) = (\ell, \ell, H)$.

Distorting p^ℓ downward from $p^{\ell m}$ can also be used to satisfy the DIC. To determine an optimal $p^{\ell *}$ to (P–S), we need to evaluate the effectiveness of price distortion relative to the likelihood ratios of future punishments. By the modified AMP formula,

$$\frac{\partial v}{\partial p^\ell} = \frac{1}{2}\left[\partial \pi^\ell / \partial p^\ell - \frac{\partial \pi^h_d / \partial p^\ell}{LR - 1}\right].$$

By manipulation, we have

$$\frac{\partial v}{\partial p^\ell} < 0 \iff LR < \frac{\partial \pi^h_d / \partial p^\ell + \partial \pi^\ell / \partial p^\ell}{\partial \pi^\ell / \partial p^\ell} \equiv R(p^\ell). \tag{4}$$

Therefore, if $LR < R(p^\ell)$, then lowering p^ℓ and reducing future punishments (keeping the DIC binding) can increase the equilibrium value. Now $R(p^\ell)$ becomes an effectiveness measure comparable to the powers of test. Intuitively, $R(p^\ell)$ measures how much deviation loss is created per unit of the equilibrium value loss by lowering p^ℓ. The following lemma evaluates $R(p^\ell)$.

LEMMA 2. *For any sorting scheme such that* $p^\ell \leq p^{\ell m}$, $R(p^\ell) > 1/\lambda$. *Moreover,* $R(p^\ell) \geq 1/(1-\lambda)$ if $p^{\ell m} \geq p^\ell \geq p^L$.

PROOF. See the Appendix.

Now, we can determine the optimal punishment scheme in an optimal SPPE. Combined with (4), Lemma 2 implies that, if $p^\ell > p^L$, p^ℓ should be lowered before using any future punishment, since $R(p^\ell)$ is greater than all the powers of test. Lemma 2 also implies that future punishments should be used only after (ℓ, ℓ, H), (h, ℓ, H), and (ℓ, h, H) are observed, since the other public outcomes have smaller powers of test than $1/\lambda$.[21]

An Upper Bound for the Optimal Sorting SPPE Value. We would like to solve (P–S) for all λ, but it is indeed hard to derive an explicit solution for λ close to 1. Given the previous results, we can instead find an upper bound by setting $LR = 1/(1-\lambda)$.

LEMMA 3. *Suppose that Assumption 1 holds,* $\lambda > 0.5$, *and that a solution to (P–S) exists. An upper bound for the value of an optimal solution to (P–S) is given by*

$$\bar{v}^s(\lambda) = \frac{1}{2\lambda}\pi^h(p^{h*}) + \frac{(2\lambda - 1)(2-\lambda)}{4\lambda}\pi^L(p^{\ell *}), \tag{5}$$

[21] The implications about price war behavior are summarized in Proposition 3 (Section 4).

where

$$p^{h*} = p^{hm}; \quad p^{\ell*} = p^L. \tag{6}$$

PROOF. From the above analysis, the value of an optimal solution to (P–S) cannot exceed

$$\frac{1}{2}\left\{\pi^h - \frac{\pi_d^h - \pi^h}{1/(1-\lambda)-1}\right\} + \frac{1}{2}\pi^\ell = \frac{1}{2}\left\{\frac{\pi^h}{\lambda} + \pi^\ell - \frac{1-\lambda}{\lambda}\pi_d^h\right\}. \tag{7}$$

Maximizing (7) with respect to p^h and p^ℓ provides an upper bound of the value. Thus, set

$$p^{h*} = \arg\max_{p^h} \pi^h = p^{hm}; \quad p^{\ell*} = \arg\max_{p^\ell}\left\{\pi^\ell - \frac{1-\lambda}{\lambda}\pi_d^h\right\}.$$

Note that

$$\begin{aligned}\pi^\ell - \frac{1-\lambda}{\lambda}\pi_d^h &= \frac{(1-\lambda)(1+\lambda)}{2}\pi^H(p^\ell) + \frac{\lambda(2-\lambda)}{2}\pi^L(p^\ell)\\ &\quad - \frac{1-\lambda}{\lambda}\left\{\frac{\lambda(1+\lambda)}{2}\pi^H(p^\ell) + \frac{(1-\lambda)(2-\lambda)}{2}\pi^L(p^\ell)\right\}\\ &= \frac{(2\lambda-1)(2-\lambda)}{2\lambda}\pi^L(p^\ell).\end{aligned}$$

This shows that, since $\lambda > 1/2$,

$$p^{\ell*} = p^L = \arg\max \pi^L(p). \tag{8}$$

Substituting the prices into (7), we obtain the upper bound as desired. ■

This upper bound is attainable if the DIC binds with a future punishment only after (ℓ, ℓ, H) is observed. However, such future punishment is not enough if λ is close enough to 1, because then the public outcome (ℓ, ℓ, H) arises with a very small probability when deviation occurs. In such a case, other future punishments are necessary. As a result, the attainable equilibrium value is strictly less than the upper bound.

4. OPTIMAL COLLUSION

4.1. *Patient Firms.* Based on the analysis in Section 3, we derive the implications of optimal collusion for patient firms. First of all, if the upper bound of sorting SPPE $\bar{v}^s(\lambda)$ is smaller than the value of the optimal pooling scheme v^r, then the optimal SPPE must exhibit price rigidity.

PROPOSITION 1. *Suppose that Assumption 1 holds and the discount factor is high enough such that all the off-schedule constraints can be ignored. Then there exists a* $\hat{\lambda} \in (0.5, 1)$ *such that for* $\lambda \in [0.5, \hat{\lambda}]$, *the pricing scheme for an optimal solution to (P) must be pooling (price rigidity).*

PROOF. From now on, let us include the accuracy level λ in the value and profit functions, for instance, $\pi^h = \pi^h(\lambda)$. Recall (3) and (5):

$$\bar{v}^s(\lambda) = \frac{1}{2\lambda}\pi^h(\lambda) + \frac{(2\lambda-1)(2-\lambda)}{4\lambda}\pi^L(p^{\ell*}),$$
$$v^r = \frac{\pi^H(p^{r*}) + \pi^L(p^{r*})}{4}.$$

Thus

$$\bar{v}^s(\lambda) \le v^r$$
$$\iff \frac{1}{2\lambda}\pi^h(\lambda) + \frac{(2\lambda-1)(2-\lambda)}{4\lambda}\pi^L(p^{\ell*}) \le \frac{\pi^H(p^{r*}) + \pi^L(p^{r*})}{4}. \tag{9}$$

Note that the right-hand side (RHS) of (9) is constant. Now we will show (i) that the left-hand side (LHS) of (9) becomes smaller than the RHS as $\lambda \to 0.5$, (ii) that the LHS becomes larger than the RHS as $\lambda \to 1$, and (iii) that the LHS is increasing in λ. Let $LHS(\lambda)$ denote the LHS in (9).

(i) First, note that $p_\lambda^{h*} \to p^{r*}$ as $\lambda \to 0.5$ (see (2) and (6)). Then

$$\lim_{\lambda\downarrow 0.5} LHS(\lambda) = \pi^h(0.5) = \frac{\pi^H(p_{0.5}^{h*}) + \pi^L(p_{0.5}^{h*})}{8}$$
$$= \frac{\pi^H(p^{r*}) + \pi^L(p^{r*})}{8} < \frac{\pi^H(p^{r*}) + \pi^L(p^{r*})}{4}.$$

(ii) Similarly,

$$\lim_{\lambda\uparrow 1} LHS(\lambda) = \frac{1}{2}\pi^h(1) + \frac{1}{4}\pi^L(p^{\ell*})$$
$$= \frac{\pi^H(p_1^{h*}) + \pi^L(p^{\ell*})}{4} > \frac{\pi^H(p^{r*}) + \pi^L(p^{r*})}{4}.$$

The inequality follows from the fact that $p_1^{h*} = \arg\max \pi^H(p)$, and $p^{\ell*} = \arg\max \pi^L(p)$.

(iii) By taking the derivative of $LHS(\lambda)$,

$$\frac{dLHS(\lambda)}{d\lambda} = \frac{d}{d\lambda}\left\{\frac{1}{2\lambda}\pi^h(\lambda) + \frac{(2\lambda-1)(2-\lambda)}{4\lambda}\pi^L(p^{\ell*})\right\}$$
$$= \frac{\lambda^2\pi^H(p_\lambda^{h*}) - (1-\lambda^2)\pi^L(p_\lambda^{h*})}{4\lambda^2} + \frac{2(1-\lambda^2)\pi^L(p^{\ell*})}{4\lambda^2}$$
$$\ge \frac{\lambda^2\pi^H(p_\lambda^{h*}) + (1-\lambda^2)\pi^L(p^{\ell*})}{4\lambda^2} > 0.$$

The second equality follows by the envelope theorem. The first inequality follows by $\pi^L(p^{\ell *}) \geq \pi^L(p_\lambda^{h*})$ since $p^{\ell *}$ is the maximizer for $\pi^L(p)$.

Clearly $LHS(\lambda)$ is continuous and thus by (i)–(iii), there must be a $\hat{\lambda} \in (0.5, 1)$ such that for $\lambda \in [0.5, \hat{\lambda})$, $\bar{v}^s(\lambda) < v_\lambda^r$ and for $\lambda \in (\hat{\lambda}, 1]$, $\bar{v}^s(\lambda) > v_\lambda^r$. This completes the proof. ■

Proposition 1 can be understood in the following intuitive way. Compared to a sorting scheme, a pooling scheme has both an advantage and a disadvantage. The advantage is that firms can enforce it without the fear of cheating, since there is no on-schedule deviation in this scheme. The disadvantage is that a pooling scheme cannot enjoy informational gain, because it does not utilize the information contained in private signals. In contrast, a sorting scheme entails future punishments and price distortions on the equilibrium path even though firms behave honestly in equilibrium. These are regarded as coordination costs due to incentive compatibility. A pooling scheme avoids the coordination costs but forgoes the informational gain at the same time.

Which scheme is optimal thus depends on the magnitude of the informational gain relative to that of the coordination costs. When the accuracy of signals is sufficiently low (λ is close to 0.5), the informational gain is close to 0. But the coordination costs are strictly positive.[22] As a result, the optimal pooling scheme strictly dominates any sorting scheme.

We conjecture that, not only its upper bound, $\bar{v}^s(\lambda)$, but also the value of an optimal solution to (P–S) is increasing in λ, thereby leading to a stronger result such that "the value of an optimal sorting SPPE is greater than v^r *if and only if* $\lambda \geq \hat{\lambda}$." However, we have not yet been able to conclude whether such a result holds, because it is hard to characterize an optimal sorting scheme for λ close to 1. Nevertheless, we establish a weaker result that optimal collusion exhibits sorting pricing if λ is close enough to 1 (see Proposition 2).[23]

Price Rigidity. An immediate implication of Proposition 1 is that, if the accuracy of signals is low enough ($\lambda < \hat{\lambda}$), it is optimal for firms to adopt a pooling scheme where they charge the same price, p^{r*}, over time regardless of their private signals on the equilibrium path. To give some sense of how big this $\hat{\lambda}$ is, we provide a concrete example. Suppose that the demand functions are $D^H(p) = 1.2 - p$ and $D^L(p) = 1 - p$, and the discount factor is $\delta = 0.9$. Then $p^{r*} = 0.55$ and $v^r = 1.21/8$. The cutoff value is $\hat{\lambda} = 0.986$. In this example, the range of λ for which a price rigidity scheme is optimal is very big.

[22] When λ is close to 0.5, p^{hm} and $p^{\ell m}$ are very close to p^{r*}. For an optimal sorting SPPE, however, Lemma 3 says that $p^{\ell *} = p^L$; thus the coordination cost due to price distortion is positive. Moreover, the deviation gain $\pi_d^h - \pi^h \approx 3[\pi^H(p^L) + \pi^L(p^L)]/8 - [\pi^H(p^{hm}) + \pi^L(p^{hm})]/8$ is also strictly positive by Assumption 1. Thus, the coordination cost due to future punishments is also positive.

[23] In a repeated game with imperfect public monitoring, Kandori (1992) shows that the sequential equilibrium payoff set expands as the quality of monitoring technology improves in Blackwell's (1953) sense. We cannot directly apply his argument since the monitoring technology in Kandori's model is exogenous, whereas in our model it is endogenous due to private information. This is why the monotonicity result is not immediate in our model.

Weiss (1993) empirically tests the relationship between price rigidity and the predictability of demand using the data of Austrian manufacturing industries. He uses coefficient of variation (CV) of time series data as a measure of demand predictability; the higher the CV, the less predictable is the demand. His test result is consistent with our prediction: As demand becomes more difficult to predict, prices tend to be more rigid. Of course, this is just one empirical test, and is not specifically designed for collusive industries. More empirical work needs to be done to directly test our prediction. Possibly, we can pick a collusive industry, and compare the pricing behavior of this industry across different countries with different variations of demand.

It is important to note that the demand predictability in a particular industry is typically affected by the predictability of economy-wide variables. For example, as the aggregate demand becomes more difficult to predict, the demand in a collusive industry might become more difficult to predict as well. As a result, the prices of collusive industries tend to be more rigid. Thus, our model predicts that prices tend to be more rigid when the macroeconomic environment becomes more unstable. This empirical implication can be tested by cross-country studies (as Lucas, 1973, does in testing the neutrality of money).

Optimality of Sorting Schemes. Since $\bar{v}^s(\lambda)$ is just an upper bound for the optimal sorting SPPE value, we have not established whether a sorting scheme is optimal for some $\lambda > \hat{\lambda}$. Although it is hard to specify the entire parameter region of λ for which a sorting scheme is optimal, we can show the optimality of sorting schemes for highly accurate signals. To verify that, we construct a feasible sorting profile that generates an equilibrium value greater than that of the optimal pooling profile.

PROPOSITION 2. *Suppose that Assumption 1 holds and firms are sufficiently patient. There exists a* $\bar{\lambda} < 1$ *such that* $\lambda > \bar{\lambda}$ *implies that optimal collusion exhibits sorting pricing.*

PROOF. Consider the following profile:

$$
\begin{aligned}
p_\lambda^h &= p_\lambda^{hm},\ p_\lambda^\ell = p^L \\
\alpha_{h\ell,\lambda}^H &= \alpha_{\ell h,\lambda}^H = k_\lambda \\
\alpha_{ij,\lambda}^S &= 0, \quad \text{for } (i, j, S) \neq (h, \ell, H), (\ell, h, H),
\end{aligned}
$$

where k_λ, and the corresponding value v_λ, are defined by the binding DIC and the modified AMP formula, namely,

$$
\text{(DIC)} \qquad \pi_d^h(p_\lambda^\ell) - \pi^h(p_\lambda^h) - \frac{\delta}{1-\delta} v_\lambda(\lambda^2 - \lambda(1-\lambda))k_\lambda = 0
$$

$$
\text{(AMP)} \qquad v_\lambda - \frac{1}{2}\left\{\pi^h(p_\lambda^h) + \pi^\ell(p_\lambda^\ell) - \frac{\pi_d^h(p_\lambda^\ell) - \pi^h(p_\lambda^h)}{LR-1}\right\} = 0.
$$

Two remarks are worth mentioning. First, this is a feasible profile as long as $k_\lambda \leq 1$. But it is not optimal since it imposes no future punishment after (ℓ, ℓ, H), which should be used prior to imposing any punishment after (h, ℓ, H). Second, the existence of k_λ and v_λ is immediate; since $LR = 1/\{2(1-\lambda)\}$ regardless of k_λ, (AMP) uniquely defines v_λ, and then (DIC) pins down k_λ.

As $\lambda \to 1$, $p_\lambda^h \to p^H = \arg\max_p \pi^H(p)$, $k_\lambda \to k_{\lambda=1} < 1$, and $1/(LR-1) \to 0$. The prices at $\lambda = 1$ are the monopoly ones under perfect information, and by continuity, the value of this profile thus converges to (a half of) the monopoly profit. Let $v^*_{\lambda=1}$ denote the optimal collusive value under perfect information.

Recall $v^*_{\lambda=1} > v^r$. This implies that, for λ close to 1, there exists a feasible sorting scheme that generates a higher value than the optimal pooling scheme does. This shows that the optimal pricing scheme must be sorting if the accuracy of signals is high enough. ■

Price Wars. If a pooling scheme is not optimal to (P), an optimal SPPE must entail a sorting pricing scheme with downward price distortion and future punishments, which we call *current* and *future price wars*, respectively. Based on the results of Section 3.3, we summarize the price war behavior as follows.

PROPOSITION 3. *Suppose Assumption 1 holds and firms are sufficiently patient. If rigid-pricing is not optimal among SPPE, then an optimal SPPE exhibits (i) a current price war in which only the price for the low signal is distorted downward, and (ii) future price wars that are triggered either only after* (ℓ, ℓ, H) *or only after* (ℓ, ℓ, H), (h, ℓ, H), *and* (ℓ, h, H).

If a sorting scheme is optimal, our model generates two implications about price wars that are different from those in the existing literature.[24] First, in an optimal sorting SPPE, only the price for the low signal is distorted downward. This result is in contrast to Rotemberg and Saloner (1986), who predicts that only the price in high demand states needs to be distorted downward, and it can be smaller than the price in low demand states. In their model on-schedule deviation is absent because there is no private information. Instead, they focus on off-schedule deviations when the discount factor is low. Smaller discounting generates stronger temptations for undercutting in high demand states. So far this effect has been absent in our model since we assume firms are sufficiently patient. Our model instead focuses on private information about demand, and shows that the on-schedule DIC is the only relevant incentive constraint, which can be relaxed only if the price assigned for the low signal is distorted downward. The general message here is that the relationship between prices and demand states in collusive industries sensitively depends on the information structure and the discount factor.

Second, in contrast to Green and Porter (1984), in which future price wars are only triggered in low demand states, in our model future price wars are only triggered in high demand states, specifically after observing (ℓ, ℓ, H), (h, ℓ, H),

[24] These price war implications are derived from an optimal SPPE that has two demand states. When the underlying demand has more than two states, these implications might not follow.

and (ℓ, h, H).[25] We do not claim that our result is more empirically relevant. Again, the general conclusion is that at which demand state future price wars are more likely to be triggered crucially depends on the information structure. Actually, empirical evidence shows no pattern about whether collusion is more likely to break down in an economic downturn or upturn. According to a survey by Levenstein and Suslow (2002, table 14), cartels in the Beer and Steel industries break up during economic downturn, whereas cartels in the Rayon industry break up during economic upturn.[26]

4.2. *Impatient Firms.* We have so far restricted attention to patient firms, and have dismissed all the off-schedule incentive constraints. Here we consider how firms' impatience affects the optimal SPPE.

Pooling Schemes. Under any pooling scheme, given that the price p^r is independent of signals, firms have a stronger incentive to engage in off-schedule deviation when they receive the h signal. Thus, to deter off-schedule deviations the following condition is necessary and sufficient:

$$(1-\delta)(2\pi^h(p^r) - \pi^h(p^r)) \leq \frac{1}{2}\delta[\pi^h(p^r) + \pi^\ell(p^r)]$$
$$\Leftrightarrow \delta \geq \frac{2\lambda\pi^H(p^r) + (2-2\lambda)\pi^L(p^r)}{(2\lambda+1)\pi^H(p^r) + (3-2\lambda)\pi^L(p^r)} \equiv \hat{\delta}(p^r, \lambda).$$

Note that $\hat{\delta}(p^r, \lambda) \in (0.5, 1)$ is increasing in λ, which means that more patience is required to prevent off-schedule deviations as the accuracy of signals increases. For the optimal pooling scheme p^{r*}, we can define $\hat{\delta}(p^{r*}, \lambda)$ accordingly. If $\delta \in [\hat{\delta}(p^{r*}, \lambda), 1]$, then the optimal pooling scheme p^{r*} is sustainable. Combining with the previous results, we can draw the following conclusion: If $\lambda \leq \hat{\lambda}$ and $\delta \in [\hat{\delta}(p^{r*}, \lambda), 1]$, an optimal SPPE is pooling with price p^{r*}.[27]

If $\delta < \hat{\delta}(p^{r*}, \lambda)$, then p^{r*} cannot be sustained as an equilibrium pooling price. To reduce firms' incentive to deviate after a signal h, the pooling price p^r needs to be distorted downward from p^{r*}.[28] Given that the expected stage payoff is concave in price, the optimal pooling price p^r is the highest price that satisfies the following condition: $\delta = \hat{\delta}(p^r, \lambda)$. Thus impatience would naturally decrease the optimal pooling SPPE payoff.

[25] A low profit is the only good indicator of deviation in Green and Porter's setting, and therefore punishments are triggered in low demand states on the equilibrium path. In our model, observing both a high demand and low prices indicates a high likelihood of deviation.

[26] Our model has another price war implication distinct from these papers: A price war can be triggered when just one of several firms charges a lower price. This can be interpreted as one firm deviating first and the other firms retaliating later. This pattern is also delivered in ABS, though, in their extended model with public iid demand shocks. We thank a referee for pointing this out.

[27] This is because in deriving an upper bound of the optimal sorting SPPE value for patient firms we ignored the off-schedule constraints; incorporating these constraints would decrease the upper bound.

[28] Distorting p^r upward from p^{r*} would increase deviation payoff after the h signal.

Sorting Schemes. For a sorting scheme with $p^h > p^\ell$, it is after a signal h that firms have stronger incentives to engage in off-schedule deviations. Now, there are two candidates for the most profitable off-schedule deviation. The first possibility is to undercut p^h and receive a deviation gain of $\pi^h(p^h)$. The second possibility is to undercut p^ℓ and receive a deviation gain of $\lambda\pi^H(p^\ell) + (1-\lambda)\pi^L(p^\ell) - \pi^h(p^h)$. When firms are impatient, both types of off-schedule deviations should be prevented. To prevent the former possibility, p^h might be distorted downward from p^{hm}; and to prevent the latter possibility, p^ℓ might be distorted downward from p^L. Note that these price distortions will also affect the on-schedule constraints.[29] Generally speaking, without imposing more concrete structure on the demand functions, it is a complex task to pin down the optimal sorting scheme and to evaluate the optimal SPPE value for impatient firms.

Since an optimal SPPE for patient firms exhibits sorting only when λ is high, here we discuss how impatience affects the optimal sorting scheme when λ is close to 1. When $\lambda = 1$, our model coincides with that of Rotemberg and Saloner. Thus optimal collusion might involve countercyclical pricing ($p^h_{\lambda=1} < p^\ell_{\lambda=1}$) if firms are sufficiently impatient. If this is the case, in our working paper version we show that for λ sufficiently close to 1 the optimal sorting scheme also exhibits countercyclical pricing ($p^h < p^\ell$). Interestingly, now the relevant on-schedule constraint is the UIC instead of the DIC. Moreover, future equilibrium price wars are triggered after low demand states; thus the price war implications in Proposition 3 are reversed. Again, the general conclusion is that the pattern of price war behavior is sensitive to the information structure and the firms' patience level, which means that empirical tests have to be careful about those factors.

It would be desirable to study how firms' impatience affects the form of the optimal SPPE for each λ (whether it exhibits rigid pricing or sorting). However, no general result has yet been obtained.[30] We leave this for future work.

5. SOME EXTENSIONS

Our model is highly stylized. There are only two firms in the industry and only two underlying demand states. In this section, we discuss how our results can be generalized. We start with the case of more than two firms. Note that we come back to the setting of patient firms.

5.1. *More Firms.*

For an industry in which there are more than two firms, the price rigidity result still carries through qualitatively. In the Appendix, we prove the following proposition.

[29] Basically, firms can use price distortions of p^h and p^ℓ to satisfy the off-schedule constraints, and to satisfy the on-schedule constraints firms can use both price distortions and future equilibrium punishments.

[30] ABS derive a result that it is more difficult to satisfy off-schedule constraints with a sorting scheme using future equilibrium punishments than with a payoff-equivalent pooling scheme. Given the common value property of demand and the correlated signal structure, this result does not hold in our model.

PROPOSITION 4. *Suppose there are $n \geq 3$ firms in an industry, and δ is high enough such that all the off-schedule constraints can be ignored. Then there exists a* $\hat{\lambda}(n) \in (0.5, 1)$ *such that for* $\lambda \in [0.5, \hat{\lambda}(n)]$, *the optimal SPPE exhibits price rigidity.*

Proposition 4 shows that the result of Proposition 1 extends to the n-firm case: A price rigidity scheme is optimal when λ is small. Again, the main reason is that the coordination costs of adopting a sorting scheme are positive as long as private signals are not perfectly accurate, whereas the informational gain vanishes if λ goes to 0.5.

Price Rigidity and Concentration. How does the price rigidity region (in terms of λ) change when the number of firms in a collusive industry increases? Indeed, it is difficult to derive general results on this matter. As a first step, we examine this question by adding one more firm to the two firm case.

PROPOSITION 5. *Fix* $\lambda > 0.5$. *Suppose a sorting scheme is optimal for* $n = 3$, *with associated* $p^{h*}(3)$, $p^{\ell*}(3)$ *and* $\alpha^*(3)$. *If*

$$[\pi^H(p^{h*}(3)) - \pi^H(p^{\ell*}(3))] \geq [\pi^L(p^{\ell*}(3)) - \pi^L(p^{h*}(3))], \tag{10}$$

then there is a sorting scheme that strictly dominates the optimal pooling scheme for $n = 2$.

PROOF. See the Appendix.

Assumption (10) plays an important role in proving Proposition 5. This holds, for example, in the following case in which demand fluctuation is driven by a parameter:

$$D^S(p) = a^S - bp, \quad S = H, L \quad \text{and} \quad a^H > a^L.$$

Now (10) is equivalent to $a^H + a^L \geq 2b(p^{h*} + p^{\ell*})$, the validity of which can be shown straightforwardly.[31]

Intuitively, coordination becomes more difficult as the number of colluding firms increases. With more firms, the probability that at least one firm charges the low price on the equilibrium path becomes higher. As a result, harsher current or future price wars are necessary to sustain the DIC. This leads to higher coordination costs. In addition, the informational gain of sorting schemes can be decreasing in the number of firms (this is so if (10) holds). As a result, the price rigidity region expands as the number of firms increases. Since the same intuition applies to a general n-firm setting, we conjecture that Proposition 5 extends to that general setting. That is, as the number of firms increases in a collusive industry, a price rigidity scheme is more likely to be optimal.

[31] By the fact that p^{h*} and $p^{\ell*}$ are optimal sorting prices for $n = 3$, $p^{h*} \leq a^H/2b$ (no distortion in p^{h*}, and it equals $a^H/2b$ only if $\lambda = 1$), and $p^{\ell*} \leq a^L/2b$ (because it is distorted downward).

Empirical Implications and Evidence. A common wisdom in the literature of industrial organization (e.g., Carlton, 1986) tells us that there is a positive correlation between concentration and price rigidity. However, our model suggests that the relationship between concentration and price rigidity may be *nonmonotonic*. In fact, in highly concentrated industries our model predicts a negative correlation between concentration and price rigidity: Prices in monopolies are more flexible, since there is no need to coordinate, than those in duopolies; and as the number of firms increases, prices are more likely to be rigid as long as the firms still manage to collude. As the number of firms increases further, however, collusion cannot be sustained any more because satisfying the off-schedule incentive constraints becomes more difficult.[32] As a result, prices become flexible.[33] Therefore, we should observe that monopolistic and fairly competitive industries have more flexible prices, whereas in oligopolistic industries prices are more rigid.

This nonmonotonicity may shed light on why the existing empirical studies about the relationship between price rigidity and concentration have generated conflicting results. For example, whereas Dixon (1983) and Carlton (1986) support a positive correlation between concentration and price rigidity, Chappell and Addison (1983) and Weiss (1993) find little support for this positive correlation with respect to demand shocks. None of these studies properly deals with a nonmonotonic relationship. Typically, they have only regressed the degree of price rigidity on the concentration ratio. Our model thus calls for more careful empirical studies to test the potential nonmonotonic relationship.

Several existing empirical results are consistent with our prediction. First, Fisher and Konieczny (1995) (studying the Canadian newspaper industry) find that oligopolies change prices less often than monopolies. Second, Qualls (1979) shows that price–cost margins are more flexible in highly concentrated oligopolies than in less concentrated oligopolies.[34] Finally, both Posner (1970) and Dick (1996) find that cartel duration increases with the number of firms, which is consistent with our prediction.[35]

5.2. *More Underlying States*. Our results remain qualitatively valid if the number of underlying states increases. As long as private signals are not perfectly informative, the incentive compatibility conditions for any sorting scheme cause positive coordination costs. On the other hand, the informational gain clearly disappears as the informativeness of signals vanishes.

If there are more than two underlying demand states, firms may find it optimal to adopt a partial sorting scheme. An optimal nonrigid scheme is harder to

[32] This is because, as the number of firms n increases, each firm's equilibrium payoff decreases at the rate of $1/n$, whereas its most profitable off-schedule deviation payoff roughly remains the same.

[33] This argument for flexible pricing is based on a modified version of our model with product differentiation (see footnote 9).

[34] By normalizing the marginal cost to a positive number, our model predicts a negative relationship between concentration and the rigidity of price–cost margin.

[35] In our model, cartels using sorting schemes have a shorter expected duration (due to future punishments on the equilibrium path) than those using price rigidity schemes. Since our model suggests that cartels with more firms is more likely to adopt a price rigidity scheme, these cartels have a longer expected duration than those with fewer firms.

characterize, but as discussed above, a price rigidity scheme must be optimal if private signals are of sufficiently low accuracy.

6. DISCUSSION

Asymmetric PPE. We have restricted our attention to symmetric PPE, and this symmetry assumption is important for our results. Asymmetric PPE can in principle attain higher profits than an optimal SPPE, and implications from optimal collusion with asymmetric equilibrium can be different. Specifically, firms may receive different continuation values on the equilibrium path, with a firm that is more suspicious of deviating being punished whereas the other firm is rewarded, for example.[36]

Although the symmetry assumption inevitably entails some loss of industry-wide profits,[37] we argue that APPE might not be so plausible for the following reasons. First, to play an (optimal) APPE requires a high level of sophistication because it can be highly nonstationary, and this might be overwhelming to the colluding firms. In contrast, an optimal SPPE can be played by simple strategies with two states. Second, the concern for off-schedule deviations becomes more serious, since the firm being at a disadvantage in a period has a stronger incentive for undercutting. The firms' patience level, therefore, needs to be sufficiently high to support such an APPE. Third, simple APPE may be easily detected as evidence of collusion by the antitrust agencies.

To elaborate on the second and the third points, consider the following simple APPE that attains higher ex ante profits for all firms than an optimal SPPE does: Colluding firms take turns to serve the entire market one by one, and if one firm breaks this rule, all firms go to perpetual Nash reversion. This APPE with rotation solves the coordination problem (there is only one firm active in each period), yet retains the informational gain (the active firm can utilize the information contained in its private signal).

Arguably, the feature of serving the market alternately can be easily detected by the antitrust agencies. In other APPE that require some firm to be nonactive in some period (say, by posting a very high price), the equilibrium play would also arouse the suspicion of the antitrust agencies.

To see the second point, consider the APPE with rotation and an optimal SPPE with n firms and λ close to 0.5. Since the signals are very uninformative, the incentives to engage in off-schedule deviation are almost independent of signals. To satisfy the off-schedule constraints for the APPE with rotation,

$$\pi^m \leq \frac{\delta^{n-1}\pi^m}{1-\delta^n} \Leftrightarrow \delta^{n-1}(1+\delta) \geq 1.$$

[36] The set of APPE payoffs is indeed hard to characterize, because the state space for an APPE can be daunting.

[37] This can be seen in models of collusion with private cost shocks, analyzed by ABS (2004) for SPPE and Athey and Bagwell (2001) for APPE. Athey and Bagwell show that if firms are sufficiently patient, an optimal APPE attains efficiency even though private information is present (Proposition 8), whereas an optimal SPPE is not efficient.

To satisfy the off-schedule constraints for the optimal SPPE (price rigidity scheme),

$$(n-1)\pi^r \leq \frac{\delta\pi^r}{1-\delta} \Leftrightarrow \delta \geq \frac{n-1}{n}.$$

Comparing the above two inequalities, we can see that the APPE with rotation requires a much higher discount factor to satisfy the off-schedule constraints, especially when n is large.[38] The main reason is that in the APPE there are big stage payoff differences among firms, which makes low payoff firms very tempted to cheat. Since APPE generally share this feature of having big differences in stage payoffs, we suspect that APPE require a higher discount factor to deter off-schedule deviations than an optimal SPPE does.

Communication. We have not addressed the issue of communications among firms. Communication can in principle avoid the coordination problem and thus improves the efficiency of sorting schemes. However, the plausibility of communications among firms is questionable because of the following reasons. First, if the enforcement of antitrust law is stringent, firms may optimally avoid communications, which could be evidence of collusion per se.[39] Second, as the number of firms increases, communications among firms become more tedious. Third, if each firm has to incur some individual cost to get the signal about demand, the very presence of communication might dampen firms' incentives to acquire information, since each firm can rely on other firms' information. For these reasons, we have restricted our attention to the case without communication.

7. CONCLUSION

This article studies pricing behavior of a collusive industry when demand is fluctuating and individual firms receive private information about demand in each period. Our model provides a theoretical microfoundation for price rigidity with respect to demand shocks. In particular, we find that the predictability of underlying demand, modeled as the accuracy of private information, plays a crucial role in determining an industry's optimal pricing scheme. If demand is poorly predictable from private information, then it is optimal for firms to adopt a rigid-pricing scheme.

Our model also finds new price war implications different from those in the previous literature. In an optimal sorting SPPE with patient firms, prices tend to be higher in high demand states than in low demand states, and future price wars are

[38] Intuitively, as n increases, after a firm makes a sale it would then need to "sit out" for a greater number of periods, and this in turn raises the incentive for this firm to cheat.

[39] In the antitrust legal context (see Hay, 1982), there was a per se approach to any kind of information exchange, that is, any information exchange among firms in the same industry is illegal per se. Recently, a modified per se approach was suggested: Under certain structural conditions, information exchange is per se illegal.

triggered only in high demand states. Moreover, our model suggests that the relationship between price rigidity and concentration is not monotonic: monopolistic and competitive industries have more flexible prices than oligopolistic industries, and in oligopolistic industries, prices tend to be more rigid as the number of firms increases.

As seen in Section 5, our model can be extended in less stylized ways without qualitatively changing the main results. One restrictive assumption that we have maintained throughout the article is that the demand shocks are i.i.d.. It seems more plausible to incorporate business cycles explicitly in the model.[40] This is left for future research.

Our model can be applied to other situations where (i) economic agents' payoffs are affected by common shocks, (ii) information about common shocks is imperfect and private, and (iii) coordination with respect to common shocks is important. If information about common shocks is poor, tailoring actions to the changing environment entails high coordination costs. It is then optimal for players to adopt a "rigid rule," in which actions are not responsive to the changing environment.

APPENDIX: PROOFS

PROOF OF LEMMA 1. We show that neither $\alpha_{ij}^S = 0$ for all (i, j, S) nor $LR \leq 1$ is optimal. Suppose $\alpha_{ij}^S = 0$ for all (i, j, S) at optimum. The binding DIC then requires $\pi_d^h = \pi^h$. As we argued before, $p^h = p^{hm}$ at a solution to (P–S) (no distortion for p^h). In the first step we show that $p^\ell < p^L$. Define

$$p_d^{hm} \equiv \arg\max_p \pi_d^h(p) = \arg\max_p \frac{\lambda(1+\lambda)}{2}\pi^H(p) + \frac{(1-\lambda)(2-\lambda)}{2}\pi^L(p).$$

It is obvious that $p^{hm} > p_d^{hm}$, since $\arg\max \pi^H(p) = p^H > p^L = \arg\max \pi^L(p)$ and the relative weight on π^H is larger in π^h than in π_d^h. Now suppose $p_d^{hm} \leq p^\ell < p^{hm}$; then

$$\begin{aligned}\pi_d^h &= \frac{\lambda(1+\lambda)}{2}\pi^H(p^\ell) + \frac{(1-\lambda)(2-\lambda)}{2}\pi^L(p^\ell)\\ &> \frac{\lambda(1+\lambda)}{2}\pi^H(p^{hm}) + \frac{(1-\lambda)(2-\lambda)}{2}\pi^L(p^{hm})\\ &> \frac{\lambda^2}{2}\pi^H(p^{hm}) + \frac{(1-\lambda)^2}{2}\pi^L(p^{hm}) = \pi^h,\end{aligned}$$

a contradiction (the first inequality follows since the RHS is a single-peaked function in p). Thus, we must have $p^\ell < p_d^{hm}$. Now suppose $p_d^{hm} > p^\ell \geq p^L$; then

[40] There are articles that study collusion under serially correlated demand with no private information. Bagwell and Staiger (1997) consider persistent demand shocks, and Haltiwanger and Harrington (1991) incorporate deterministic business cycles.

$$
\begin{aligned}
\pi_d^h &= \frac{\lambda(1+\lambda)}{2}\pi^H(p^\ell) + \frac{(1-\lambda)(2-\lambda)}{2}\pi^L(p^\ell) \\
&\geq \frac{\lambda(1+\lambda)}{2}\pi^H(p^L) + \frac{(1-\lambda)(2-\lambda)}{2}\pi^L(p^L) \\
&> \frac{\lambda^2}{2}\pi^H(p^H) + (1-\lambda)^2\frac{\pi^L(p^L)}{2} \\
&> \frac{\lambda^2}{2}\pi^H(p^{hm}) + (1-\lambda)^2\frac{\pi^L(p^{hm})}{2} = \pi^h,
\end{aligned}
$$

a contradiction (the first inequality follows from the single-peakedness of π_d^h and $p^\ell < p_d^{hm}$, and the second inequality is based on Assumption 1). Therefore, we must have $p^\ell < p^L$.

So far, we have $\alpha_{ij}^S = 0$ and $p^h = p^{hm} > p^L > p^\ell$ for a solution to (P–S). In the next step, we show that this profile is dominated by the following profile: $p^h = p^{hm}$, $\tilde{p}^\ell = p^\ell + \varepsilon \in (p^\ell, p^L)$, $\alpha_{ij}^S = 0, \forall(i, j, S) \neq (\ell, \ell, H)$, and $1 > \alpha_{\ell\ell}^H(\varepsilon) > 0$ such that

$$(1-\delta)\big(\tilde{\pi}_d^h(\varepsilon) - \pi^h\big) = \delta\tilde{v}(\varepsilon)\Pr(h\ell; H)\alpha_{\ell\ell}^H(\varepsilon),$$

where $\tilde{\pi}_d^h(\varepsilon)$ and $\tilde{v}(\varepsilon)$ are defined according to this new profile. First, note that $\tilde{v}(\varepsilon)$ and $\alpha_{\ell\ell}^H(\varepsilon)$ are continuous, and $\tilde{v}(0) = v$ and $\alpha_{\ell\ell}^H(0) = 0$.[41] It should also be noted that $LR = \{\lambda(1-\lambda) + (1-\lambda)^2\}/(1-\lambda)^2 = 1/(1-\lambda)$. For a sufficiently small ε, we show that $\alpha_{\ell\ell}^H(\varepsilon) \leq 1$, and $\tilde{v}(\varepsilon) > v$, which implies this profile is feasible and dominates the original profile. Indeed,

$$
\begin{aligned}
\tilde{v}(\varepsilon) - v &= \frac{1}{2}\left\{\pi^h - \frac{\tilde{\pi}_d^h(\varepsilon) - \pi^h}{LR-1}\right\} + \frac{1}{2}\pi^\ell(\varepsilon) - \frac{1}{2}\{\pi^h + \pi^\ell\} \\
&= \frac{1}{2}\left\{\pi^\ell(\varepsilon) - \pi^\ell - \frac{1-\lambda}{\lambda}\left(\tilde{\pi}_d^h(\varepsilon) - \pi^h\right)\right\},
\end{aligned}
$$

where $\pi^\ell(\varepsilon) = \pi^\ell(p^\ell + \varepsilon)$. Recall $\tilde{v}(0) = v$. In order to have $\tilde{v}(\varepsilon) - v > 0$ for some small ε, it suffices to show that

$$\frac{d}{d\varepsilon}\left\{\pi^\ell(\varepsilon) - \frac{1-\lambda}{\lambda}\tilde{\pi}_d^h(\varepsilon)\right\}\bigg|_{\varepsilon=0} > 0.$$

[41] The value $\tilde{v}(\varepsilon)$ is defined by the modified AMP formula (page 493). Note that the likelihood ratio LR equals $\{\Pr(h\ell; H) + \Pr(\ell\ell; H)\}/\Pr(\ell\ell; H)$, independent of the level of $\alpha_{\ell\ell}^H(\varepsilon) > 0$. Thus

$$\tilde{v}(\varepsilon) = \frac{1}{2}\left\{\pi^h - \frac{\tilde{\pi}_d^h(\varepsilon) - \pi^h}{LR-1}\right\} + \frac{1}{2}\pi^\ell(\varepsilon)$$

is continuous in ε, and $\tilde{v}(0) = v$. This also implies that

$$\alpha_{\ell\ell}^H(\varepsilon) = \frac{(1-\delta)\left(\tilde{\pi}_d^h(\varepsilon) - \pi^h\right)}{\delta\tilde{v}(\varepsilon)\Pr(h\ell; H)}$$

is continuous (at least around $\varepsilon = 0$), and $\alpha_{\ell\ell}^H(0) = 0$.

In fact,

$$\frac{d}{d\varepsilon}\pi^{\ell}(\varepsilon)\,|_{\,\varepsilon=0} = \frac{(1-\lambda)(1+\lambda)}{2}\pi^{H\prime}(p^{\ell}) + \frac{\lambda(2-\lambda)}{2}\pi^{L\prime}(p^{\ell})$$

$$\frac{1-\lambda}{\lambda}\frac{d}{d\varepsilon}\tilde{\pi}^{h}_{d}(\varepsilon)\,|_{\,\varepsilon=0} = \frac{(1-\lambda)(1+\lambda)}{2}\pi^{H\prime}(p^{\ell}) + \frac{(1-\lambda)^2(2-\lambda)}{2\lambda}\pi^{L\prime}(p^{\ell}),$$

where $\pi^{S\prime}(p) = d\pi^{S}/dp$, $S = H, L$. Thus

$$\frac{d}{d\varepsilon}\left\{\pi^{\ell}(\varepsilon) - \frac{1-\lambda}{\lambda}\pi^{h}_{d}(\varepsilon)\right\}\bigg|_{\varepsilon=0} = \left\{\lambda - \frac{(1-\lambda)^2}{\lambda}\right\} \times \frac{2-\lambda}{2}\pi^{L\prime}(p^{\ell}) > 0,$$

since, for $\lambda > 1/2$, $\lambda > (1-\lambda)^2/\lambda$, and $\pi^{L\prime}(p^{\ell}) > 0$ by single-peakedness. This is the desired fact, which leads to a contradiction.

Finally, if $LR \leq 1$, then the profile must be dominated by another one with no future punishments, that is, all α's being zero. This shows that a profile with $LR \leq 1$ cannot be optimal to (P–S). ■

PROOF OF LEMMA 2. By manipulation,

$$\frac{\partial \pi^{h}_{d}}{\partial p^{\ell}} = \frac{\lambda(1+\lambda)}{2}\pi^{H\prime}(p^{\ell}) + \frac{(1-\lambda)(2-\lambda)}{2}\pi^{L\prime}(p^{\ell}),$$

$$\frac{\partial \pi^{\ell}}{\partial p^{\ell}} = \frac{(1-\lambda)(1+\lambda)}{2}\pi^{H\prime}(p^{\ell}) + \frac{\lambda(2-\lambda)}{2}\pi^{L\prime}(p^{\ell}) \geq 0.$$

The last inequality results from $p^{\ell} \leq p^{\ell m}$, and equality holds only if $p^{\ell} = p^{\ell m}$. Now suppose that $p^{\ell} \geq p^{L} = \arg\max \pi^{L}(p)$, and therefore $\pi^{L\prime}(p^{\ell}) \leq 0$. Note that $\pi^{H\prime}(p^{\ell}) > 0$. Then

$$\begin{aligned}\frac{\partial \pi^{h}_{d}/\partial p^{\ell} + \partial \pi^{\ell}/\partial p^{\ell}}{\partial \pi^{\ell}/\partial p^{\ell}} &= \frac{\lambda(1+\lambda)\pi^{H\prime}(p^{\ell}) + (1-\lambda)(2-\lambda)\pi^{L\prime}(p^{\ell})}{(1-\lambda)(1+\lambda)\pi^{H\prime}(p^{\ell}) + \lambda(2-\lambda)\pi^{L\prime}(p^{\ell})} + 1 \\ &\geq \frac{\lambda(1+\lambda)\pi^{H\prime}(p^{\ell}) + \lambda(2-\lambda)\pi^{L\prime}(p^{\ell})}{(1-\lambda)(1+\lambda)\pi^{H\prime}(p^{\ell}) + \lambda(2-\lambda)\pi^{L\prime}(p^{\ell})} + 1 \\ &\geq \frac{\lambda}{1-\lambda} + 1 = \frac{1}{1-\lambda}.\end{aligned}$$

Suppose instead that $p^{\ell} < p^{L}$, and therefore $\pi^{L\prime}(p^{\ell}) > 0$. In this case the ratio is

$$\begin{aligned}\frac{\partial \pi^{h}_{d}/\partial p^{\ell} + \partial \pi^{\ell}/\partial p^{\ell}}{\partial \pi^{\ell}/\partial p^{\ell}} &= \frac{\lambda(1+\lambda)\pi^{H\prime}(p^{\ell}) + (1-\lambda)(2-\lambda)\pi^{L\prime}(p^{\ell})}{(1-\lambda)(1+\lambda)\pi^{H\prime}(p^{\ell}) + \lambda(2-\lambda)\pi^{L\prime}(p^{\ell})} + 1 \\ &> \frac{(1-\lambda)(1+\lambda)\pi^{H\prime}(p^{\ell}) + (1-\lambda)(2-\lambda)\pi^{L\prime}(p^{\ell})}{(1-\lambda)(1+\lambda)\pi^{H\prime}(p^{\ell}) + \lambda(2-\lambda)\pi^{L\prime}(p^{\ell})} + 1 \\ &> \frac{1-\lambda}{\lambda} + 1 = \frac{1}{\lambda}.\end{aligned}$$

This completes the proof. ■

PROOF OF PROPOSITION 4. In the optimal price rigidity scheme,

$$p^{r*}(n) = \arg\max_{p^b} \left\{ \frac{\pi^H(p) + \pi^L(p)}{2} \right\}$$
$$v^r(n) = \frac{\pi^H(p^{r*}) + \pi^L(p^{r*})}{2n}.$$

Thus, the optimal rigid price $p^{r*}(n)$ is independent of n, and so is the industry profit. Moreover, each firm's collusive profit $v^r(n)$ is $1/n$ of the industry profit.

To solve for the optimal sorting scheme, we first need to identify the optimal punishment scheme on the equilibrium path. Due to the symmetry requirement, firms' decision about future punishments only depends on the observed demand $S \in \{H, L\}$ and the observed number of firms charging price p^ℓ, $z \in \{0, 1, \ldots, n\}$. More specifically, given an outcome (S, z), firms go to perpetual Nash reversion with probability α_z^S. Let $v^{FS}(n)$ denote the value of an optimal sorting scheme with n firms.[42] Now the value function becomes

$$v^{FS}(n) = \frac{1-\delta}{2}\{\pi^h(n) + \pi^\ell(n)\} + \delta v^{FS}(n) \sum_{\substack{z\in\{0,1,\ldots,n\}\\ S\in\{H,L\}}} \Pr(z, S)\big(1 - \alpha_z^S\big)$$

and the binding DIC becomes

$$(1-\delta)\big(\pi_d^h(n) - \pi^h(n)\big) = \delta v^{FS}(n) \sum_{\substack{z\in\{0,1,\ldots,n-1\}\\ S\in\{H,L\}}} \Pr(z, S \mid h)\big(\alpha_{z+1}^S - \alpha_z^S\big),$$

where

$$\pi^h(n) = \lambda^n \frac{\pi^H(p^h)}{n} + (1-\lambda)^n \frac{\pi^L(p^h)}{n}$$

$$\pi_d^h(n) = \lambda \sum_{k=0}^{n-1} C_{n-1}^k \lambda^{n-1-k}(1-\lambda)^k \frac{\pi^H(p^\ell)}{k+1} + (1-\lambda) \sum_{k=0}^{n-1} C_{n-1}^k \lambda^k (1-\lambda)^{n-1-k} \frac{\pi^L(p^\ell)}{k+1}$$

$$\pi^\ell(n) = \lambda \sum_{k=0}^{n-1} C_{n-1}^k \lambda^k (1-\lambda)^{n-1-k} \frac{\pi^L(p^\ell)}{k+1} + (1-\lambda) \sum_{k=0}^{n-1} C_{n-1}^k \lambda^{n-1-k}(1-\lambda)^k \frac{\pi^H(p^\ell)}{k+1}.$$

[42] As before, we drop λ in v's, p's, and α's for notational simplicity.

Again, each public outcome has a distinct likelihood ratio. By simple calculation, the LR of each (S, z) is

$$LR(H, z) = \frac{\lambda C_{n-1}^{z-1}\lambda^{n-z}(1-\lambda)^{z-1} + (1-\lambda)C_{n-1}^{z-1}\lambda^{n-z}(1-\lambda)^{z-1}}{\lambda C_{n-1}^{z}\lambda^{n-z-1}(1-\lambda)^{z} + (1-\lambda)C_{n-1}^{z-1}\lambda^{n-z}(1-\lambda)^{z-1}} = \frac{1}{1-\lambda}\frac{z}{n}.$$

Similarly,

$$LR(L, z) = \frac{\lambda C_{n-1}^{z-1}\lambda^{z-1}(1-\lambda)^{n-z} + (1-\lambda)C_{n-1}^{z-1}\lambda^{z-1}(1-\lambda)^{n-z}}{\lambda C_{n-1}^{z-1}\lambda^{z-1}(1-\lambda)^{n-z} + (1-\lambda)C_{n-1}^{z}\lambda^{z}(1-\lambda)^{n-z-1}} = \frac{1}{\lambda}\frac{z}{n}.$$

According to these formulas, all the public outcomes can be ranked in terms of punishment efficiency. The outcome (H, n) obviously has the highest LR, thus is the most efficient punishment. For (H, z) the LR is increasing in z, and so is the LR for (L, z). In the optimal punishment scheme, a particular $\alpha_z^S(n) > 0$ only if $\alpha_k^S(n) = 1$ for any (S, k) with $k > z$.

Using the modified AMP formula, we can specify an upper bound of $v^{FS}(n)$. From the formulas of LRs, the actual LR cannot exceed $1/(1-\lambda)$ for each n. Now by the modified AMP formula, the following is an upper bound of $v^{FS}(n)$:

$$\text{(A.1)} \qquad \begin{aligned} \bar{v}^s(n) &= \frac{1}{2}\left\{\pi^h(n) + \pi^\ell(n) - \frac{\pi_d^h(n) - \pi^h(n)}{1/(1-\lambda) - 1}\right\} \\ &= \frac{1}{2}\left\{\frac{\pi^h(n)}{\lambda} + \pi^\ell(n) - \frac{1-\lambda}{\lambda}\pi_d^h(n)\right\}. \end{aligned}$$

Maximizing (A.1) with respect to $p^h(n)$ and $p^\ell(n)$ yields

$$\begin{aligned} p^{h*}(n) &= \arg\max_{p^h}\left\{\lambda^n\frac{\pi^H(p^h)}{n} + (1-\lambda)^n\frac{\pi^L(p^h)}{n}\right\} \\ p^{\ell*}(n) &= p^{L*} = \arg\max_{p}\pi^L(p). \end{aligned}$$

In deriving $p^{\ell*}(n)$, we use the fact that

$$\pi^\ell(n) - \frac{1-\lambda}{\lambda}\pi_d^h(n) = \left[\lambda - \frac{(1-\lambda)^2}{\lambda}\right]\left[\sum_{k=0}^{n-1} C_{n-1}^k\lambda^k(1-\lambda)^{n-1-k}\frac{1}{k+1}\right]\pi^L(p^\ell).$$

As $\lambda \to 0.5$, $p^{h*}(n) \to p^{r*}(n)$. Then,

$$\begin{aligned} \lim_{\lambda\to 0.5}\bar{v}^s(n) = \pi^h(n) &= \frac{(1/2)^n}{n}\left[\pi^H(p^{r*}) + \pi^L(p^{r*})\right] \\ &< \frac{1}{2n}\left[\pi^H(p^{r*}) + \pi^L(p^{r*})\right] = v^r(n). \end{aligned}$$

Therefore, when λ is small, a price rigidity scheme is optimal. ■

PROOF OF PROPOSITION 5. The proof is by construction. Specifically, we construct an equilibrium sorting scheme for $n = 2$ such that its value is higher than $v^r(2)$.

Fix the prices for the sorting scheme such that $p^h = p^{h*}(3)$ and $p^\ell = p^{\ell *}(3)$. We proceed in several steps.

Step 1. We want to show

$$2[\pi^h(2) + \pi^\ell(2)] \geq 3[\pi^h(3) + \pi^\ell(3)]. \tag{A.2}$$

On the equilibrium path, a firm's ex ante stage profit is just $1/n$ of the whole industry's ex ante stage profit. Hence, it is sufficient to prove that the industry's stage profit is decreasing in n, that is,

$$\begin{aligned} G(n) \equiv \frac{1}{2}[&\lambda^n \pi^H(p^h) + (1-\lambda^n)\pi^H(p^\ell) \\ &+ (1-\lambda)^n \pi^L(p^h) + (1-(1-\lambda)^n)\pi^L(p^\ell)] \end{aligned}$$

is decreasing in n.

$$\begin{aligned} 2[G(n) - G(n+1)] = {}& \lambda^n(1-\lambda)[\pi^H(p^h) - \pi^H(p^\ell)] \\ &+ (1-\lambda)^n \lambda[\pi^L(p^h) - \pi^L(p^\ell)] \geq 0. \end{aligned}$$

The inequality comes from condition (10).

Step 2. We want to show $\pi_d^h(3) - \pi^h(3) \geq \pi_d^h(2) - \pi^h(2)$. By calculation, we have

$$\begin{aligned} \pi_d^h(3) &= \lambda\left[\lambda^2 + 2\lambda(1-\lambda) \times \frac{1}{2} + (1-\lambda)^2 \times \frac{1}{3}\right]\pi^H(p^\ell) \\ &\quad + (1-\lambda)\left[\lambda^2 \times \frac{1}{3} + 2\lambda(1-\lambda) \times \frac{1}{2} + (1-\lambda)^2\right]\pi^L(p^\ell) \\ \pi_d^h(2) &= \lambda\left[\lambda + (1-\lambda) \times \frac{1}{2}\right]\pi^H(p^\ell) + (1-\lambda)\left[\lambda \times \frac{1}{2} + (1-\lambda)\right]\pi^L(p^\ell) \\ \pi^h(3) &= \frac{\lambda^3}{3}\pi^H(p^h) + \frac{(1-\lambda)^3}{3}\pi^L(p^h) \\ \pi^h(2) &= \frac{\lambda^2}{2}\pi^H(p^h) + \frac{(1-\lambda)^2}{2}\pi^L(p^h). \end{aligned}$$

Thus,

$$\begin{aligned} &\left[\pi_d^h(3) - \pi^h(3)\right] - \left[\pi_d^h(2) - \pi^h(2)\right] \\ &\quad = \left[\pi_d^h(3) - \pi_d^h(2)\right] + \left[\pi^h(2) - \pi^h(3)\right] \\ &\quad = \lambda\frac{2\lambda - 1}{6}\pi^H(p^h) + \lambda\left[\frac{1-\lambda}{2} - \frac{(1-\lambda)^2}{3}\right][\pi^H(p^h) - \pi^H(p^\ell)] \\ &\qquad - (1-\lambda)\frac{2\lambda-1}{6}\pi^L(p^\ell) - (1-\lambda)\left[\frac{1-\lambda}{2} - \frac{(1-\lambda)^2}{3}\right][\pi^L(p^\ell) - \pi^L(p^h)] > 0. \end{aligned}$$

The last inequality comes from the fact $\pi^H(p^h) > \pi^L(p^\ell)$ and condition (10).

Step 3. We want to show $v^{FS}(2) > 3v^{FS}(3)/2$. To abuse notation, we use $v(n)$ and $v^{FS}(n)$ interchangeably. We first need to identify the optimal punishment scheme for $n = 3$, that is, $\alpha^*(3)$. Given the ranking of the LR of each public outcome, $\alpha_z^{S*} = 0$ if there is a (S', z') such that $\alpha_{z'}^{S'*} < 1$ and $LR(S', z') > LR(S, z)$. We need to consider several cases. Here we just provide the proof for two cases. The proof of the other cases is similar.

CASE 1. Suppose $\alpha^*(3)$ sets $\alpha_3^{H*} \le 1$ and all other $\alpha_z^{S*} = 0$. For $n = 2$, we construct $\alpha_2^H = (1-\lambda)\alpha_3^{H*}$ and all other $\alpha_z^S = 0$. Note that this construction is feasible since the constructed $\alpha_2^H \in [0, 1)$. Then by the binding DIC for $n = 3$,

$$\pi_d^h(3) - \pi^h(3) = \frac{\delta}{1-\delta} v(3)\lambda(1-\lambda)^2\alpha_3^{H*}.$$

By step 2,

$$\begin{aligned}\pi_d^h(2) - \pi^h(2) \le \pi_d^h(3) - \pi^h(3) &= \frac{\delta}{1-\delta} v(3)\lambda(1-\lambda)^2\alpha_3^{H*} \\ &= \frac{\delta}{1-\delta} v(3)\lambda(1-\lambda)\alpha_2^H < \frac{\delta}{1-\delta} v(2)\lambda(1-\lambda)\alpha_2^H. \qquad \text{(A.3)}\end{aligned}$$

Thus, the DIC for $n = 2$ holds if inequality (A.3) holds. But inequality (A.3) holds if $v(2) > v(3)$. Now we only need to verify that from the value function.

$$v(3) = \frac{(1-\delta)[\pi^h(3) + \pi^\ell(3)]/2}{1-\delta+\delta(1-\lambda)^3\alpha_3^{H*}/2} < \frac{2}{3}\frac{(1-\delta)[\pi^h(2) + \pi^\ell(2)]/2}{1-\delta+\delta(1-\lambda)^2\alpha_2^{H}/2} = \frac{2}{3}v(2).$$

Note that we use the result of step 1 in deriving the above inequality.

CASE 2. Now suppose $\alpha^*(3)$ sets $\alpha_3^{H*} = 1$, $\alpha_2^{H*} \in (0, 1]$ and all other $\alpha_z^S = 0$. Then by the binding DIC for $n = 3$,

$$\pi_d^h(3) - \pi^h(3) = \frac{\delta}{1-\delta} v(3)\big[\lambda(1-\lambda)^2 + \lambda(1-\lambda)(3\lambda-1)\alpha_2^{H*}\big].$$

For $n = 2$, we construct α_2^H and α_1^H as follows:

$$\begin{aligned}&\alpha_2^H = (1-\lambda) + (3\lambda-1)\alpha_2^{H*} \quad \text{and} \quad \alpha_1^H = 0 \quad \text{if } (1-\lambda) + (3\lambda-1)\alpha_2^{H*} \le 1,\\ &\alpha_2^H = 1 \quad \text{and} \quad \alpha_1^H = \frac{1-\lambda}{2\lambda-1}\big[(3\lambda-1)\alpha_2^{H*} - \lambda\big] \quad \text{if } (1-\lambda) + (3\lambda-1)\alpha_2^{H*} > 1.\end{aligned}$$

Note that α_1^H is well defined since $\alpha_2^{H*} \le 1$. By the construction, in both subcases

$$\begin{aligned}\pi_d^h(2) - \pi^h(2) \le \pi_d^h(3) - \pi^h(3) &= \frac{\delta}{1-\delta} v(3)\big[\lambda(1-\lambda)^2 + \lambda(1-\lambda)(3\lambda-1)\alpha_2^{H*}\big] \\ &\le \frac{\delta}{1-\delta} v(2)\big[\lambda(1-\lambda)\alpha_2^H + \lambda(2\lambda-1)\alpha_1^H\big]. \qquad \text{(A.4)}\end{aligned}$$

Thus the DIC for $n = 2$ holds if inequality (A.4) is valid. But inequality (A.4) holds as long as $v(2) \geq v(3)$. Now we only need to verify it in both subcases. We start with the first subcase. From the value function for $n = 3$,

$$v(3) = \frac{(1-\delta)[\pi^h(3) + \pi^\ell(3)]/2}{1 - \delta + \delta\big[(1-\lambda)^3/2 + 3\lambda(1-\lambda)^2\alpha_2^{H*}/2\big]} < \frac{2}{3}\frac{(1-\delta)[\pi^h(2) + \pi^\ell(2)]/2}{1 - \delta + \delta\big[(1-\lambda)^2\alpha_2^{H}/2\big]}$$
$$= \frac{2}{3}v(2).$$

The last inequality comes from the fact that (the denominator of $v(3)$ is bigger than that of $v(2)$)

$$\frac{1}{2}(1-\lambda)^2\big[1 - \lambda + 3\lambda\alpha_2^{H*} - \alpha_2^{H}\big] = \frac{1}{2}(1-\lambda)^2\alpha_2^{H*} \geq 0.$$

In the second subcase,

$$\begin{aligned} v(3) &= \frac{(1-\delta)[\pi^h(3) + \pi^\ell(3)]/2}{1 - \delta + \delta[(1-\lambda)^3/2 + 3\lambda(1-\lambda)^2\alpha_2^{H*}/2]} \\ &< \frac{2}{3}\frac{(1-\delta)[\pi^h(2) + \pi^\ell(2)]/2}{1 - \delta + \delta[(1-\lambda)^2/2 + \lambda(1-\lambda)\alpha_1^{H}]} \\ &= \frac{2}{3}v(2). \end{aligned}$$

The last inequality comes from the fact that

$$\begin{aligned} &\frac{1}{2}(1-\lambda)^2\left[1 - \lambda + 3\lambda\alpha_2^{H*} - 1 - \frac{\lambda}{2\lambda - 1}\big[(3\lambda - 1)\alpha_2^{H*} - \lambda\big]\right] \\ &= \frac{1}{2}\frac{\lambda}{2\lambda - 1}(1-\lambda)^2\big(1 - \alpha_2^{H*}\big) \geq 0. \end{aligned}$$

Step 4. $v^{FS}(2) > v^r(2)$. In the previous three steps, we show that there is a sorting scheme for $n = 2$ such that $v^{FS}(2) > 3v^{FS}(3)/2$. By the assumption that $v^{FS}(3) \geq v^r(3)$ and the fact that $v^r(2) = 3v^r(3)/2$, we get our desired result that $v^{FS}(2) > v^r(2)$. ■

REFERENCES

ABREU, D., P. MILGROM, AND D. PEARCE, "Information and Timing in Repeated Partnerships," *Econometrica* 59 (1991), 1713–33.

———, D. PEARCE, AND E. STACCHETTI, "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory* 39 (1986), 251–69.

———, ———, AND ———, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* 58 (1990), 1041–63.

ATHEY, S., AND K. BAGWELL, "Optimal Collusion with Private Information," *Rand Journal of Economics* 32 (2001), 428–65.

———, ———, AND C. SANCHIRICO, "Collusion and Price Rigidity," *Review of Economic Studies* 71 (2004), 317–49.

BAGWELL, K., AND R. W. STAIGER, "Collusion over the Business Cycle," *Rand Journal of Economics* 28 (1997), 82–106.

BLACKWELL, D., "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics* 24 (1953), 265–72.

CARLTON, D., "The Rigidity of Prices," *American Economic Review* 76 (1986), 637–58.

CHAPPELL, H., AND J. ADDISON, "Relative Prices, Concentration, and Money Growth," *American Economic Review* 73 (1983), 1122–26.

DICK, A., "When Are Cartels Stable Contracts?" *Journal of Law and Economics* 39 (1996), 241–83.

DIXON, R., "Industry Structure and the Speed of Price-Adjustment," *Journal of Industrial Economics* 32 (1983), 25–37.

FISHER, T., AND J. KONIECZNY, "The Relative Rigidity of Oligopoly Pricing," *Economics Letters* 48 (1995), 33–38.

GEROSKI, P. A., "Price Dynamics in UK Manufacturing: A Microeconomics View," *Economica* 59 (1992), 403–19.

GREEN, E. J., AND R. H. PORTER, "Noncooperative Collusion under Imperfect Information," *Econometrica* 52 (1984), 87–100.

HALTIWANGER, J., AND J. E. HARRINGTON, JR., "The Impact of Cyclical Demand Movements on Collusive Behavior," *Rand Journal of Economics* 22 (1991), 89–106.

HAY, G., "Oligopoly, Shared Monopoly, and Antitrust Law," *Cornell Law Review* 67 (1982), 439–81.

KANDORI, M., "The Use of Information in Repeated Games with Imperfect Monitoring," *Review of Economic Studies* 59 (1992), 581–93.

LEVENSTEIN, M., AND V. SUSLOW, "What Determines Cartel Success?" Mimeo, University of Michigan Business School, 2002.

LUCAS, R., "Some International Evidence on Output–Inflation Tradeoffs," *American Economic Review* 63 (1973), 326–34.

POSNER, R., "A Statistical Study of Antitrust Enforcement," *Journal of Law and Economics* 13 (1970), 365–419.

QUALLS, D., "Market Structure and the Cyclical Flexibility of Price–Cost Margins," *Journal of Business* 2 (1979), 305–25.

ROTEMBERG, J., AND G. SALONER, "A Supergame-Theoretic Model of Price Wars during Booms," *American Economic Review* 76 (1986), 390–407.

STIGLITZ, J. E., "Price Rigidities and Market Structure," *American Economic Review* 74 (1984), 350–55.

WEISS, C., "Price Inertia and Market Structure: Empirical Evidence from Austrian Manufacturing," *Applied Economics* 25 (1993), 1175–86.