University of Massachusetts Amherst

From the Selected Works of Christian Rojas

Spring 2012

The Role of Demand Information and Monitoring on Tacit Collusion

Christian Rojas, *University of Massachusetts - Amherst*

Available at: https://works.bepress.com/christian_rojas/1/

The role of demand information and monitoring in tacit collusion

Christian Rojas*

Abstract

Motivated by the Green and Porter (1984) and Rotemberg and Saloner (1986) models, we construct lab experiments to test the effects of two factors on collusion: information (regarding next period's demand state) and monitoring (of a rival's past action). Results indicate that information may facilitate collusion more than monitoring, especially as subjects gain experience. A robust finding is that subjects in the Rotemberg and Saloner treatment cooperate as predicted by this theory: collusion falls dramatically in anticipation of unusually large demand and returns to high levels otherwise. These results suggest that tacit and fairly elaborate collusion could arise in stochastic environments.

1. Introduction

■ Firms' coordination to obtain high profits has been a continuous concern for researchers and antitrust authorities. As a consequence, there exists a large body of theoretical work on the factors that determine the likelihood of collusion. But empirical analyses of collusion are difficult because the illegal status of cartels makes data scarcely available. Further, even if such data were available, the task of identifying and estimating the effect of different market conditions on collusion can be difficult due to many unobservable factors in field data. This article seeks to improve the understanding of the role of two factors that have been prominent in models of repeated interaction with stochastic demand: demand information (knowledge of next period's demand realization)¹ and monitoring (knowledge of rivals' past actions). More specifically, we analyze the effects of these factors on collusion by generating data from controlled experiments

*University of Massachusetts Amherst; rojas@isenberg.umass.edu.

I am extremely thankful to Catherine Eckel for her support and constructive comments. I am particularly indebted to Joe Harrington for all the time spent in helping me improve this article and for his encouragement. I am grateful to Matt Parrett and Zhongmin Wang for their detailed and thoughtful comments, which significantly improved the article. I also thank, Vince Crawford, Chetan Dave, Ernan Haruvy, Igal Hendel, Tom Hubbard, John Spraggon, John Stranlund, Bart Wilson, Alex Yuskavage, and several referees for their suggestions. Sheryl Ball provided grant administration support. Francesca Colantuoni and Chris Burns provided valuable research assistance. Funding was provided by the National Science Foundation (grant no. SES-0526229) and by the University of Massachusetts at Amherst (Department of Resource Economics and Healy grant no. 135573).

¹ We use the term "next period's demand" throughout even though we refer to the demand realization for which firms' current strategy choices (price or quantity) pertain. We choose to use this term because in practice (and when feasible) firms (and subjects in our experiment) obtain information about demand prior to choosing a strategy.

that have varying degrees of uncertainty along these two market conditions. In addition, we use the generated data to analyze whether some well-known theoretical predictions find support in the data.

The motivation for our inquiry comes from two influential models in the literature, Green and Porter (1984) and Rotemberg and Saloner (1986) (GP and RS henceforth). Both models assume a stochastic demand structure, but differ in their assumptions regarding what firms know about next period's demand and rivals' actions. First, RS assume that firms have perfect foresight about demand next period, whereas GP assume that firms are always uncertain about demand realizations (past and future). Second, GP assume that monitoring among cartel members is imperfect, in the form of a noisy public signal, whereas RS assume that monitoring is perfect. Our experimental design is guided by these differences in assumptions: in two of our treatments, monitoring and demand information differ in the same way as GP and RS differ. In a third, intermediate, treatment there is uncertainty about next period's demand realization (as in GP) but monitoring is perfect (as in RS); this treatment allows us to separate the imperfect monitoring effect from the imperfect demand information effect.

The differences in assumptions of the RS and GP models can generate different collusive equilibria. GP show that there can be a collusive equilibrium where price wars (possibly temporary ones) are part of the equilibrium path; the model predicts that such reversionary periods are triggered by unusually low demand realizations. Conversely, RS show that a cartel may experience price wars during periods of unusually large demand shocks. Thus, a common interpretation of these theories is that GP predicts more collusion during booms, whereas RS predicts more collusion during, or as a consequence of, recessions.²

Our article has two main objectives. First, we investigate whether and how collusion is affected by whether demand information and monitoring.³ Second, we study whether predictions of the RS and GP models find support in their respective treatments. In particular, we study whether collusion in the RS environment is inversely related to the level of demand to be faced next period, and whether a low demand realization in the previous period triggers less collusive behavior in later periods in the GP environment.⁴

Results indicate that for a sufficiently high discount rate, demand information seems to facilitate collusion more than monitoring, especially as subjects gain experience. This effect can range from being statistically insignificant to increasing the probability of collusion by 0.33. The effect of monitoring, on the other hand, is of smaller magnitude and less precise: it ranges from being statistically insignificant to increasing the probability of collusion by 0.15.

Although predictions of both the RS and GP models find some support in the data, a robust finding is that subjects' behavior in the RS treatment conforms closely to the countercyclical collusive prediction of the Rotemberg and Saloner theory. Relative to periods in which anticipated demand is low or medium, the probability of collusion in anticipation of a high demand period is lower by as much as 0.69 and 0.53, respectively. Remarkably, results indicate that a high demand realization does not collapse collusion entirely, as subjects are able to return to high levels of cooperation afterward.

From an industrial organization point of view, studying collusion in uncertain environments is important because no real-world market operates under complete certainty. There have been empirical efforts to study collusion under uncertainty in particular industry settings. Studies of

² Although GP and RS generate theoretical predictions that appear to be "at odds" (Ellison, 1994), there is an important caveat: in GP the duration of a price war is triggered by low demand and remains in place regardless of demand realizations during the "punishment" period, whereas in RS a price war is state dependent: it is only observed in periods when the demand shock is unusually large.

³ Strictly speaking, we measure the effect of "improved monitoring," as monitoring is not completely absent in the GP treatment (i.e., monitoring is noisy). We refer to "monitoring" throughout to simplify exposition.

⁴ Because the two motivating theories have different assumptions, we do not test the empirical validity of one theory against the other (i.e., each theory can or cannot be valid in its own domain). Instead, we analyze whether, in our experimental design, each theory's salient predictions emerge in their respective environments.

this type that have a connection to this article have primarily tested the predictions of the RS and GP models (jointly or individually). These papers include Porter (1983), Ellison (1994), Borenstein and Shepard (1996), and Chevalier, Kashyap, and Rossi (2003). In some cases the studies have found evidence in favor of the theory being evaluated, whereas in other cases there has been little or no support.

In contrast with the cited literature, our investigation exploits the control provided by the lab to ensure that the studied market environment satisfies the assumptions of the model of interest. This approach can be informative for two reasons. First, the large number of equilibria that are mathematically possible in infinitely repeated games has led some economists to criticize the usefulness of the theory. Second, there is relatively little evidence as to which market conditions make a particular collusive equilibrium more likely to emerge. Our experiments not only suggest that tacit collusion is more likely to arise in certain stochastic environments but that it can be fairly elaborate (as in our RS treatment).

We are not the first to study collusion under uncertainty in the lab, however.⁵ Several studies have introduced the notion of “noise” to the repeated prisoner’s dilemma in several ways. The basic feature of the main treatments in these studies is that although subjects face a “baseline” payoff matrix prior to choosing their strategy, the realized payoffs become “noisy” after the game is played. Specifically, a stochastic rule can change either the baseline payoffs or the strategy of each player. The studies by Feinberg and Snyder (2002), Bereby-Meyer and Roth (2006), and Aoyagi and Frechette (2008) fall into the first category, whereas the studies by Holcomb and Nelson (1997) and Fudenberg, Rand, and Dreber (forthcoming) fall into the latter. Except for the study by Bereby-Meyer and Roth, a common feature among these papers is that subjects cannot know with certainty their opponent’s choice; this feature makes these designs most similar to our GP treatment. Conversely, Bereby-Meyer and Roth allow for perfect monitoring after decisions have been made, making this design somewhat comparable to our intermediate treatment. The consensus in this literature is that noise plays an important role: relative to treatments where noise is absent or less important, collusion is lower (or slower to evolve).⁶

Our article differs from this experimental literature in two ways. First, we are interested in collusion in markets characterized by uncertainty rather than on “context-free” cooperation in the presence of some abstract notion of noise. In contrast, we frame our experiment as a game in which subjects act as sellers in a (hypothetical) market where noise takes the form of uncertain future demand.⁷ Second, one aim of prior work has been to compare uncertain environments with deterministic counterparts. Although informative, a deterministic benchmark might be less useful from a practical point of view. Instead, we preserve the stochastic nature of future demand and manipulate two market conditions within this uncertain baseline.

There is an industrial organization strand of the experimental literature that is related (in motivation and design) to our interest in the RS model, although in versions somewhat different from the original theory. Ruffle (2009) investigates whether large buyers pay lower prices when facing oligopolistic sellers in a market where buyers of different sizes appear randomly for an infinite numbers of periods. Consistent with RS’s predictions and our results, the author finds overall support for countercyclical pricing: lower prices are offered to high demand buyers. Abbink and Brandts (2009) study finitely repeated duopolies with a deterministic demand pattern

⁵ There have been several papers looking at the feasibility of cooperation in the two-player prisoner’s dilemma under no uncertainty. These lab investigations include Dal Bó (2005), Duffy and Ochs (2009), and Dal Bó and Frechette (2011). A common finding in this work is that collusion, when it is a feasible equilibrium, is more likely to occur when the discount factor is higher and when subjects gain more experience.

⁶ These studies differ in several dimensions. Monitoring is public in the Aoyagi and Frechette design but is private in the other studies. In addition to studying the effect of noise, the studies of Aoyagi and Frechette and Fudenberg et al. present a thorough analysis of subjects’ strategies. The last two studies provide useful accounts of this literature.

⁷ We could not verify whether the Feinberg and Snyder (2002) experiment used neutral language because instructions for this study were not available.

TABLE 1 Typical Payoff Table

		Seller 2	
		Low (L)	High (H)
Seller 1	Low (L)	Π_s^C, Π_s^C	Π_s^{ND}, Π_s^D
	High (H)	Π_s^D, Π_s^{ND}	Π_s^{NE}, Π_s^{NE}

Note: Collusion (C), deviation (D), nondeviation (ND), Nash equilibrium (NE); s is the demand state.

in which future demand either shrinks or grows. Although the authors find evidence that is at odds with RS predictions, it remains unclear whether this is attributable to a design that does not match the stochastic and infinitely repeated nature of the RS theory.

The article is organized as follows. Section 2 describes the model. Section 3 provides details of the experimental design and Section 4 describes its implementation. Finally, Section 5 presents the results and Section 6 briefly discusses our main findings.

2. The model

■ Our experience, as well as prior experimental work (see Holt, 1995), suggests that games with a limited set of strategies and players give collusion a higher chance of occurrence. Therefore, we employ a prisoner's dilemma version of the model, specifically a two-firm Cournot game with symmetric firms and two quantity choices (low (L) and high (H)).⁸ Further, we assume three demand states: high (h), medium (m), and low (l). The game is infinitely repeated and the demand states occur with probabilities 0.60 for medium and 0.20 for both high and low. The payoff table for demand state $s \in (l, m, h)$ has the structure shown in Table 1.

Π_s^C denotes the collusive payoff, Π_s^D the payoff to a “deviating” firm, Π_s^{ND} the payoff to a “Nondeviating” firm, and Π_s^{NE} the Nash equilibrium payoff, with $\Pi_s^D > \Pi_s^C > \Pi_s^{NE} > \Pi_s^{ND}$. With one demand state, and if firms use a grim-trigger strategy, it is well known that the collusive outcome (L, L) can be supported in equilibrium if

$$\Pi^D - \Pi^C < \sum_{t=1}^{\infty} \delta^t (\Pi^C - \Pi^{NE}) = \frac{\delta}{1 - \delta} (\Pi^C - \Pi^{NE}), \quad (1)$$

where δ is the discount factor. Thus, collusion will be sustainable if firms' discount factor is larger than the critical value $\delta^* = [\Pi^D - \Pi^C] / [\Pi^D - \Pi^{NE}]$. With stochastic demand, collusive equilibrium paths with temporary reversions to the competitive outcome are possible. We turn to these theories in our setup.

□ **The RS model: perfect demand foresight and perfect monitoring.** In addition to perfect monitoring of rivals' actions, RS assume that firms know with certainty the demand realization in the next period, but that firms only know the distribution of demand realizations for the remaining future periods.⁹ In this setting, the future punishment of deviation is independent of next period's demand realization, whereas the incentive to deviate is only a function of next period's demand realization. RS show that for a sufficiently high demand realization next period, the incentive to deviate is greater than the future punishment and thus collusion can no longer be supported as an equilibrium. In the literature, this breakdown is interpreted as either countercyclical pricing

⁸ Both RS and GP can be cast as either price or quantity games with homogeneous products. Quantity competition seems more natural here, as matrices have nonzero payoffs at the Nash equilibrium.

⁹ RS consider IID demand shocks. Haltiwanger and Harrington (1991) and Kandori (1991) show that, under certain conditions, countercyclical pricing also holds with cyclical demand, but Bagwell and Staiger (1997) show that under other stochastic conditions the RS results are reversed.

or as price wars during booms. RS indicate that the model can yield either result, depending on the strategy space. If the strategy space is continuous, smooth countercyclical pricing will be observed in equilibrium. Conversely, if the strategy space is constrained to either compete or collude (as is the case here), then price wars (or reversion to Cournot) will be observed during periods of sufficiently high demand (RS). The corresponding RS version of (1) is thus

$$\Pi_s^D - \Pi_s^C < \frac{\delta}{1 - \delta} [E(\Pi_s^C - \Pi_s^{NE})], \quad (2)$$

where $E(\cdot)$ represents the expected punishment over the different demand states. To accommodate the main prediction of the RS theory, we construct payoff tables so that equation (2) can, in some cases, hold for the medium (m) and low (l) demand states but not for the high (h) demand state; that is, $[\Pi_h^D - \Pi_h^C] > [0.6(\Pi_m^C - \Pi_m^{NE}) + 0.2(\Pi_l^C - \Pi_l^{NE})] > [\Pi_m^D - \Pi_m^C] > [\Pi_l^D - \Pi_l^C]$. As is the case with infinitely repeated games, the RS prediction is one of several equilibria. In Appendix A.1 as well as later in the paper, we show when the RS prediction (and other outcomes) can be feasible equilibria in our experimental design.

□ **The GP model: no demand foresight and imperfect monitoring.** GP assume that in any period t , firms know only the distribution of demand realizations next period. In addition, firms can only imperfectly monitor their rivals' quantity choices through a noisy signal, that is, the market price. These two sources of uncertainty prevent firms from inferring their opponent's action even after the realization of demand. GP show that there can be a collusive equilibrium where finite price wars are part of the collusive path. A crucial feature of the model is that price wars are triggered by a low demand realization (revealed to firms in the form of a low market price) rather than by firms' deviation; the intuition for this result is that if any suspicious behavior is not castigated, the threat of punishment is not credible.

More formally, GP assume that firms use a trigger strategy in which all players revert to the Nash equilibrium (H, H) for N periods when the noisy signal (price) falls below threshold level k . The GP model involves a nontrivial optimization problem in which the length of the optimal punishment, the trigger price, and the level of collusion have to be adjusted. These calculations, however, are straightforward in our experimental design.

Denote firm i 's choice as $y_i \in \{L, H\}$. Also, define $f(p|y)$ as the density function that determines the probability of observing price level p given the outcome $y = (y_i, y_j)$, and $F(p|y)$ is its corresponding cumulative distribution function. The expected payoff in each reversionary period is given by $\lambda_i = 0.2\Pi_h^{NE} + 0.6\Pi_m^{NE} + 0.2\Pi_l^{NE}$, and the expected profit (for next period) when the opponent is sticking to the collusive quantity (L) is

$$\gamma_i(y_i) = \begin{cases} 0.2\Pi_h^C + 0.6\Pi_m^C + 0.2\Pi_l^C & \text{if } y_i = L \\ 0.2\Pi_h^D + 0.6\Pi_m^D + 0.2\Pi_l^D & \text{if } y_i = H. \end{cases}$$

Last, define V_i recursively as

$$V_i = \gamma_i(L) + (1 - F_k(L))\delta V_i + F_k(L) \left[\sum_{t=1}^N \delta^t \lambda_i + \delta^{N+1} V_i \right], \quad (3a)$$

where $F_k(y_i) = F(k|y_i; y_j = L)$. That is, for a given (price) threshold level k , firm i 's output choice can shift the probability with which the observed price level may fall below k (given that the other firm sticks to the collusive outcome). After some manipulation, (3a) can be written as

$$V_i = \frac{\gamma_i(y_i) + F_k(L) \frac{\delta(1 - \delta^N)}{1 - \delta} \lambda_i}{1 - (1 - F_k(L))\delta - F_k(L)\delta^{N+1}}. \quad (3b)$$

The resulting Bellman equation that ensures the existence of the GP equilibrium, for a given punishment length N and a threshold price k , is given by

$$V_i \geq \gamma_i(H) + (1 - F_k(H))\delta V_i + F_k(H) \left[\sum_{t=1}^N \delta^t \lambda_i + \delta^{N+1} V_i \right]. \quad (4)$$

There could be many values of N for which this condition holds (or none at all), including $N = \infty$. Similarly, there can be several threshold prices for which one can find a finite N such that (4) holds. Computation of all feasible equilibria for our experimental setup, including all feasible punishment lengths and threshold prices, can be found in Appendix A.

3. Experimental design

■ To simulate an infinitely repeated game with discount factor δ , a subject interacts with another seller with a continuation probability equal to δ (Fudenberg and Tirole, 1989). After a repeated game (a “supergame”) is terminated, subjects are rematched with another seller; this exercise continues until the allotted time for the session is over. This design is attractive for two reasons. First, it generates numerous infinitely repeated games. Second, it produces a few lengthy supergames where certain behavior (i.e., temporary reversals to competition, as in GP) is more likely to arise. We implement our baseline experiments with $\delta = 0.75$ and study additional experiments with $\delta = 0.60$ and $\delta = 0.90$.

The experiment is framed as a duopoly market in which participants act as sellers of a hypothetical good who decide between two output levels in each period. In each period, subjects are asked to simultaneously choose a quantity (high (H) or low (L)). Profits are determined by one of three tables: high demand (h), medium demand (m), or low demand (l). This is illustrated in Table 2.

Every period, each table is chosen with the probabilities indicated earlier (0.2, 0.6, and 0.2, respectively); probabilities are known by all subjects in all treatments. There are two sets of payoff matrices. The main difference between these two parameterizations is that deviation is less attractive in parameterization 2: the additional profits from deviation are smaller, and reversion to NE play is more costly. Using different parameterizations and discount factors allows us to investigate the robustness of the results to different collusive incentives.

We construct three treatments by varying the way in which demand information is presented to subjects. To keep treatments comparable, subjects in all treatments remain uninformed about their opponent’s choice; this means that when monitoring is allowed, the opponent’s choice can be inferred from the payoff tables.

- (i) Full information treatment (FI) (RS theory). Before a period starts, subjects are told which demand state they will face. Hence, uncertainty about next period’s demand state is removed, although demand states for future periods remain unknown. Similarly, monitoring is feasible because in our design it is possible to infer the rival’s strategy after a period’s profit is realized.
- (ii) Imperfect monitoring treatment (IM) (GP theory). Subjects only know the distribution for next period’s demand state: demand next period will most likely be “normal” (i.e., medium with probability 0.60), but there is also a chance of experiencing a demand shock (high or low, each with probability 0.20); further, our experimental software does not inform subjects of the demand state during the previous period. Note that with the constructed payoff tables, this setup allows for imperfect monitoring of the opponent’s choice of output even after profits have been realized for that period. To see this, suppose that player 1 chooses to

TABLE 2 Payoff Tables for the Three Demand States

Parameterization 1 High Demand (probability = 0.2)				Parameterization 2 High Demand (probability = 0.2)			
		Seller 2				Seller 2	
		Low	High			Low	High
Seller 1	Low	26.00, 26.00	7.50, 43.00	Seller 1	Low	31.00, 31.00	9.00, 43.00
	High	43.00, 7.50	12.50, 12.50		High	43.00, 9.00	12.50, 12.50
Medium Demand (probability = 0.6)				Medium Demand (probability = 0.6)			
		Seller 2				Seller 2	
		Low	High			Low	High
Seller 1	Low	7.50, 7.50	2.10, 12.50	Seller 1	Low	9.00, 9.00	2.50, 12.50
	High	12.50, 2.10	3.50, 3.50		High	12.50, 2.50	3.50, 3.50
Low Demand (probability = 0.2)				Low Demand (probability = 0.2)			
		Seller 2				Seller 2	
		Low	High			Low	High
Seller 1	Low	2.10, 2.10	0.60, 3.50	Seller 1	Low	2.50, 2.50	0.70, 3.50
	High	3.50, 0.60	1.00, 1.00		High	3.50, 0.70	1.00, 1.00

collude (L) and his profit turns out to be 2.10 (parameterization 1). If demand was low, it means that the opponent also chose the collusive outcome (L), but if demand was medium it means that the opponent deviated from it (H). This imperfect monitoring is possible whether a firm chooses the high or the low output.¹⁰ As assumed by GP, our design implies that imperfect monitoring is *public*. To see this, note that because subjects know their own profit as well as the chosen quantity, they can also (indirectly) perceive the corresponding price: this price is the noisy public signal as defined by GP (Appendix B illustrates this point in greater detail).^{11,12}

- (iii) Monitoring treatment (M) (intermediate treatment). Here, imperfect monitoring is removed from the IM treatment by informing subjects of the chosen demand state right after a period is over; this is equivalent to informing them of their opponent's choice.¹³ This treatment isolates one of the two factors that differentiate the FI treatment from the IM treatment. In the M treatment, collusion can be an equilibrium if the following condition (a modified version of (1)) is met¹⁴:

$$E(\Pi_s^D - \Pi_s^C) < \frac{\delta}{1 - \delta} [E(\Pi_s^C - \Pi_s^{NE})], \quad (5)$$

¹⁰ A system of equations allows us to maintain the imperfect monitoring structure (available upon request).

¹¹ Alternatively, subjects could be informed of the noisy price. A possible downside of this approach is that instructions increase in complexity and subjects are not informed of earnings until the end of the experiment (see Aoyagi and Frechette, 2008 for an example of such a procedure).

¹² In the analysis below, we focus on the threshold given by the price corresponding to the collusive profit in the low demand state, Π_s^L . Appendix A provides an explanation for this choice.

¹³ A "perfect demand foresight" with "imperfect monitoring" treatment is unfeasible, as this case would allow subjects to infer their opponent's strategy.

¹⁴ Although we do not have a strong theoretical prediction about whether collusion would be easier (or more difficult) in the M treatment with respect to the FI treatment, it is likely that the M treatment is a more collusive-friendly environment because this treatment is equivalent to the case of deterministic demand (at least when firms are risk neutral) and, by the logic of RS, we know that collusion is more difficult when demand fluctuates.

TABLE 3 Range of CRRA Parameter (r) for Which Collusive Equilibrium Is Theoretically Feasible in Each Experimental Design Cell

Collusive Equilibrium Considered	FI		M Collusion in Every Period	IM Finite Price War of $N = 2$ Periods (GP Equilibrium) ^b
	Countercyclical Collusion (RS Equilibrium) ^a	Collusion in All Demand States		
$\delta = 0.60\text{-}P1$	[+0.63, +0.89]	[+0.70, + ∞)	[-0.29, + ∞)	[+1.60, + ∞)
$\delta = 0.75\text{-}P1$	[-1.68, +0.48]	[+0.07, + ∞)	[-1.49, + ∞)	[+1.20, + ∞)
$\delta = 0.75\text{-}P2$	($-\infty$, 0.10]	($-\infty$, + ∞)	[-3.18, + ∞)	[+0.06, + ∞)
$\delta = 0.90\text{-}P2$	($-\infty$, -1.20]	($-\infty$, + ∞)	[-6.03, + ∞)	[-0.30, + ∞)

Note: δ is the discount factor (continuation probability); FI = full information treatment, M = monitoring treatment, IM = imperfect monitoring treatment; P1 = parameterization 1, P2 = parameterization 2. The collusive equilibria considered in the FI and M treatments are based on a grim-trigger strategy. Appendix A discusses additional feasible equilibria in both FI and IM treatments r is the coefficient of relative risk aversion from the CRRA utility specification; $r < 0$ denotes risk-loving behavior, $r = 0$ is risk neutrality, and $r > 0$ is risk aversion.

^aCollusion is feasible during medium and low demand states but not during the high demand state.

^b N represents the length of a price war for which the GP equilibrium exists; if a GP equilibrium for N exists, then a GP equilibrium for $N + i$ ($i = 1, \dots$) also exists. The equilibrium considered here assumes that the threshold is the price that corresponds to the collusive profit in the low demand state (Π_f); see Appendixes A and B. $N = 2$ is arbitrarily chosen in this table for illustration purposes; the range of the risk parameter for which the GP equilibrium exists grows with N . See Appendix A for a computation in all possible threshold prices and for the limiting $N = 1$ and $N = \infty$ cases.

One important feature of the experimental design is that the payoff matrices are constructed so that the percentage difference between payoffs and across entries remains invariant across demand states. For example, in parameterization 1, the payoff in the collusive outcome is about 100% higher (with rounding error) than the payoff in the Nash equilibrium. The reason for constructing payoff matrices in this fashion is that individuals seem to care about relative variation in payoffs rather than the absolute variation (Weber, Shafir, and Blais, 2004). Thus, potential confounding effects of variation in relative payoffs across demand states are reduced.

In sum, the experiment considers three treatments and four δ -parameterization pairs (δ -P): $\delta = 0.60\text{-}P1$, $\delta = 0.75\text{-}P1$, $\delta = 0.75\text{-}P2$, $\delta = 0.90\text{-}P2$. Because different δ -P pairs offer different incentives for collusion, it is important to discuss the feasibility of collusive equilibria in our design. In the FI treatment, we consider two possible collusive equilibria, the equilibrium devised by RS (collusion in medium and low demand states but not in high demand states) as well as the “fully” collusive equilibrium (collusion in all demand states). For the M treatment we consider collusion in every period, whereas for the IM treatment we consider the GP equilibrium (temporary price wars triggered by a low demand state).¹⁵

Theories of infinitely repeated games assume risk neutrality by subjects; thus, one possibility is to report the feasibility of each collusive equilibrium under this assumption. However, the experimental economics literature has documented varying degrees of risk tolerance by subjects in the lab, with risk aversion being the most common finding (Harrison and Ruström, 2008). We adopt the convention in the literature of assuming a constant relative risk aversion (CRRA) utility function and compute the range of the risk aversion parameter (r) for which each collusive equilibrium is feasible.¹⁶ Table 3 reports this range for each cell in our experimental design.

As stated earlier, the experiments with discount factor $\delta = 0.75$ can be considered as our baseline (because $\delta = 0.75$ is used in both parameterizations). The other two discount factors either expand or restrict the ranges of the CRRA coefficient over which collusive equilibria are feasible. As one moves from the first to the last δ -P pair in the table, the CRRA range becomes larger, indicating that the environment becomes more “collusive friendly.” This is expected, as

¹⁵ The collusive equilibria we contemplate in the FI and M treatments consider a grim-trigger strategy.

¹⁶ The CRRA utility function is given by $u(c) = c^{1-r}/(1-r)$, where r is the coefficient of relative risk is aversion and c is the monetary payoff; $r < 0$, denotes risk-loving behavior, $r = 0$ is risk neutrality, and $r > 0$ is risk aversion. In a comparison of different elicitation mechanisms, Harrison and Ruström report CRRA coefficients between 0.51 and 0.86.

TABLE 4 Number of Participants and Average Earnings per Session

Session number	Parameterization 1					
	Full Information		Monitoring		Imperfect Monitoring	
	I	II	III	IV	V	VI
δ	<i>0.75</i>	0.60	<i>0.75</i>	0.60	<i>0.75</i>	0.60
Number of participants	22	20	18	18	24	20
Average earnings (\$)	<i>31.30</i>	30.06	<i>32.20</i>	28.72	<i>29.70</i>	28.46

Session number	Parameterization 2					
	Full Information		Monitoring		Imperfect Monitoring	
	VII	VIII	IX	X	XI	XII
δ	<i>0.75</i>	0.90	<i>0.75</i>	0.90	<i>0.75</i>	0.90
Number of participants	24	20	18	18	22	20
Average earnings (\$)	<i>38.34</i>	38.60	<i>37.47</i>	34.06	<i>34.25</i>	34.28

Note: Excludes show-up fee (\$7). Italic numbers denote baseline treatments.

theory indicates that more patient firms (i.e., a larger discount factor) and smaller net gains from deviation (i.e., moving from parameterization 1 to parameterization 2) should result in stronger incentives to collude.¹⁷

4. Implementation

■ Twelve sessions with a total of 244 subjects were run. Six sessions (two per treatment) were run with each parameterization. Three sessions were run with parameterization 1 using $\delta = 0.60$, three sessions were run with parameterization 2 using $\delta = 0.90$, and the remaining six sessions (three with each parameterization) were run using the baseline discount factor, $\delta = 0.75$. Subjects were recruited from a variety of courses at the University of Massachusetts at Amherst throughout the spring of 2010 and were informed that the experiment would last approximately two hours. Demographic composition was not unusual for laboratory experiments with college students: 35% were females, 66% were white, and the combined number of juniors and seniors was 63% (with the remaining 37% distributed relatively evenly among freshmen, sophomores, and graduate students). Subjects received a \$7 show-up fee and earned additional money from their decisions; earnings from decisions were in experimental dollars (\$1 = 30 experimental \$). The number of participants and average earnings are presented in Table 4.

All experiments were computerized and programmed in z-Tree (Fischbacher, 2007). Web Appendix C contains instructions for one of the treatments. Students were assigned a computer terminal and advised that they would be randomly paired with multiple “other sellers” in the room for as long as time permitted and that communication with other participants was forbidden. The continuation probability (0.60, 0.75, or 0.90) was implemented from period 1 and subjects were informed that the matching mechanism was determined by a random number: if after a period the random number chosen by the computer was greater than δ , they would be rematched with another participant in the room; otherwise they would remain interacting with the current “other seller.”

Special efforts were made to ensure subjects’ comprehension and familiarity with the experiment before the start of the game. Extensive instructions were coupled with three practice

¹⁷ An outcome where collusion is feasible only in the low demand state would, technically, also qualify as an “RS” equilibrium. We do not study this case, as its feasibility is rare in our design: it is unfeasible in the $\delta = 0.90$ treatment and it is feasible in the other treatments only for unlikely high levels of risk aversion $r > 3$.

TABLE 5 Summary of Supergame Lengths

	$\delta = 0.60$	$\delta = 0.75$	$\delta = 0.90$
Number of supergames faced by a seller ^a	37	23	12
Average number periods per supergame	3.16	5.21	10
Maximum supergame length (number of periods)	8	26	29
Minimum supergame length	1	1	1

^aThis is the number of times a participant was matched with a different seller.

periods, each with one demand state, and a quiz. If a subject did not respond correctly to a question in the quiz, the participant was approached by the experimenter for explanation.

A variety of colors was used for the different demand states and to distinguish own profits from the other seller's profits. Also, the three possible profit tables were permanently displayed on the left-hand side of the screen (see Web Appendix C). All subjects in all treatments were informed about the probability of appearance of each profit table. To determine the demand state (high, medium, or low), random draws from a uniform distribution were taken once and the same set of demand states implied by these draws was used in all treatments to preserve comparability. Similarly, a predetermined set of draws was used to establish the length of supergames across all sessions. Table 5 summarizes the length of supergames for the three discount factors (the complete sequence of supergame lengths appears in Appendix D). Finally, all sessions ended at period 120 to make each of them last approximately two hours, including instructions.¹⁸

A period consisted of subjects making simultaneous decisions between low output (*L*) and high output (*H*); afterward, subjects were informed of profits and the period ended. In the FI treatment, the decision screen presented subjects with the demand state they would face. Conversely, in the IM and M treatments, the decision screen only reminded subjects of the probability with which each demand state would be chosen for play.

In the M treatment, the profit screen revealed the chosen demand state. Also, this screen highlighted the cell in the chosen profit table that determined the subject's profit. Because the FI and M treatments imply perfect monitoring *after* the period is over, the profit screen for both of these treatments was the same. In the IM treatment, the profit screen presented subjects with the *possible* outcomes that might have occurred, effectively implementing the desired imperfect monitoring. The figures in Appendix E display the screen shots of decision and profit screens for the three treatments.

The program also contains a history table where subjects can see their cumulative earnings. In addition to the experiment, subjects completed a small survey that contained questions on demographics, subjects' assessment of the clarity of the experiment,¹⁹ and an open-ended question regarding the strategy used. We use these responses later in the analysis. A total of 29,280 observations were collected (120 periods $\times$ 244 participants). At the end of the experiment, subjects were individually called in private and were paid their cumulative earnings in cash.

5. Results

■ Our strategy for reporting results was to group findings into (i) the effect of information and monitoring on collusion and (ii) theoretical predictions of the RS and GP models supported by the

¹⁸ One potential concern with the use of a predetermined number of total rounds is that subjects could exchange this information outside the lab, thereby generating an undesirable end-of-game effect. As explained later (and in Appendix G), this effect seems unimportant in this experiment.

¹⁹ Answers to the statement "The instructions for the experiment were clear and easy to follow" were: strongly agree (70%), agree (23%), neither agree nor disagree (3%), disagree (2%), and strongly disagree (2%).

FIGURE 1

FRACTION OF TIMES THAT SUBJECTS CHOSE THE COLLUSIVE OUTPUT ACROSS ALL 120 PERIODS (BY SESSION)

Note: FI is the full information treatment, M is the monitoring (intermediate) treatment, and IM is the imperfect monitoring treatment. P1 and P2 are the parameterizations considered; δ is the discount factor.

data. The following section presents a descriptive analysis of the data, and the section after that presents the econometric analysis. In what follows, the analysis focuses, primarily, on a subject's decision to collude.

□ **Descriptive results.** In this section, we slice the data in subsets that are meaningful for the two objectives of the article. We report graphically, for each subset, the fraction of subjects who choose to collude. We then discuss the most salient patterns of the data.

Effect of demand information and monitoring on collusion and other patterns. Figure 1 presents the fraction of times, across all 120 periods, that subjects chose the collusive outcome (y axis); this information is displayed for each of the 12 experimental sessions (x axis). The 12 sessions are sorted from left to right first by parameterization (P1, P2), then by discount factor ($\delta = 0.60, 0.75, 0.90$), and then by treatment (FI, M, and then IM). Each line on the graph corresponds to a different δ -P pair, and it connects the frequencies of the three treatments in that pair.

There are three aspects of the figure worth highlighting. First, collusion rates appear to increase as the theoretical incentives to collude become stronger: for a given treatment, the frequency of collusion (7 out of 9 cases) increases as one moves across δ -P pairs from left to right. Second, a drop in the frequency of collusion from the FI to the M treatment in all δ -P pairs suggests that lack of knowledge of next period's demand state hinders collusion. Further, the importance of this effect appears to increase as the environment becomes more collusive friendly. Third, the effect of monitoring appears to be mixed, with two δ -P pairs showing an increase in the frequency of collusion from the M to the IM treatment and the other two experiencing decreases (albeit larger in magnitude than the increases).

Learning and end-of-game effects often arise in repeated-interaction experiments. In our setting, we are primarily concerned with learning that expresses itself as growing collusion as time progresses. Conversely, an end-of-game effect is associated with a sharp decline in collusion late in the experiment as subjects "sense" the last period of play is approaching. We explore the existence of these effects by looking at the frequency of collusion in each supergame.²⁰

²⁰ There are sufficient supergames toward the end of the experiment for this exercise to be a reliable check of the end-of-game effect. In the last 13 periods of the experiment there are, respectively, six, five, and three supergames in the three discount factor environments considered ($\delta = 0.60, 0.75, 0.90$). The last supergame initiates in period 120 for

FIGURE 2

FRACTION OF TIMES THAT SUBJECTS CHOSE THE COLLUSIVE OUTPUT, BY SUPERGAME

Note: The smooth lines correspond to the fit of a quadratic trend.

The four panels in Figure 2 report these frequencies, together with the fit given by a quadratic trend.

In terms of the overall evolution of collusion, the FI environment shows the steepest trend of all three treatments: it is clearly positive in all but the least collusive-friendly setting ($\delta = 0.60$). Conversely, the M and IM treatments show trends that are either flat or slightly downward sloping, except in the $\delta = 0.75$ -P2 pair, where there appears to be an upward trend in both the M and IM treatments. We interpret these patterns as suggestive of stronger learning effects in the RS treatment. The figure suggests that an end-of-game effect is unlikely.²¹

Finally, the data exhibit two other features, a “restart” effect and a “reciprocate” effect. The former refers to an increase in collusion when a new supergame starts (i.e., when subjects are rematched, with another seller), whereas the latter refers to a greater likelihood of collusion in the current period when the other seller decided to collude in the previous period. Figures F1 and F2 illustrate these two effects.

Theory-related patterns. In this section, we investigate whether collusion frequencies are consistent in ways predicted by RS and GP theories. We first focus on how the frequency of collusion varies by type of demand under two conditions: (i) demand that is to be faced next period, and (ii) demand faced in the previous period. The reason for looking at collusion under these two conditions is that RS and GP provide distinct predictions in each situation. In RS, collusion is expected to decrease when firms anticipate a high demand state, but it should be

$\delta = 0.60$ and $\delta = 0.75$ and in period 114 for $\delta = 0.90$. Looking at the frequency of collusion by supergame (rather than by period) helps to better visualize the evolution of play.

²¹ This effect was formally tested in the next regressions similar to those presented in the next section. Appendix G displays the results.

FIGURE 3

FRACTION OF TIMES THAT SUBJECTS CHOSE THE COLLUSIVE OUTPUT, BY LEVEL OF DEMAND TO BE FACED NEXT PERIOD (BY SESSION)

Note: See notes in Figure 1 and text for acronym details.

independent of the prior demand realization (i.e., collusion should return to “normal” levels after a high demand period). Conversely, next period’s demand level should be unimportant in GP, whereas collusion should become less likely after a low demand realization.

Figure 3 reports the frequency of collusion by the type of demand to be faced the next period. Unsurprisingly, in the two treatments that do not provide information on the demand realization (M and IM), next period’s demand realization does not appear to have an impact on collusive behavior. Conversely, and consistent with RS’s prediction, collusion in the FI treatment is significantly lower when subjects anticipate a high demand state, with the frequency of collusion being between 2.3 and 5.7 times smaller than the frequency of collusion when facing a low demand realization next period (these two numbers correspond, respectively, to the 0.90-P2 and 0.60-P1 pairs). Further, the figure shows that collusion in the FI treatment appears to be lower in the medium demand state relative to the high demand state, although this difference is much smaller in parameterization 2. This last observation suggests that the RS theory is a more accurate predictor of behavior in parameterization 2. This is consistent with the RS equilibrium being feasible for a wider range of risk attitudes in parameterization 2 than in parameterization 1 (see Table 3).

Figure 4 reports the frequency of collusion by the type of demand faced in the previous period. It can be seen that a small difference in the frequency of collusion across the three demand realizations is more clearly present in the IM treatment, especially when high and low demand are contrasted. Although the existence of GP-style patterns in the data (i.e., finite price wars) is more difficult to visualize and capture descriptively, the observed drop in the frequency of collusion following a low demand period is consistent with the trigger mechanism suggested by GP.

Notably, the frequency of collusion in the FI treatment is slightly higher after a high demand state than after the other two demand realizations. In other words, the high levels of defection observed in anticipation of a high demand period (Figure 3) do not appear to trigger long-lasting price wars, which is in line with the countercyclical path of the RS model. This return to “normal” levels of collusion after a high demand realization can be observed in more detail in Figures F4 and F5.

In GP, once a price war starts, its length is independent of the demand realizations during the punishment phase. Because some of the low demand realizations may well arise during a price war, a fairer exercise might be to compare frequencies between (i) all periods leading up to the first low demand realization in a supgame and (ii) the period following such realization.

FIGURE 4

FRACTION OF TIMES THAT SUBJECTS CHOSE THE COLLUSIVE OUTPUT, BY LEVEL OF DEMAND FACED IN THE PREVIOUS PERIOD (BY SESSION)

Note: See notes in Figure 1 and text for acronym details.

Figure F3 reports these results and confirms the moderate support for the type of trigger suggested by GP in the IM treatment.

Because many supergames in our design are too short for finite price wars to emerge (as predicted by GP), we visually inspected the collusive frequencies for longest supergame in the IM treatment (not shown). In line with the evidence reported earlier, collusive behavior appears to drop after a low demand realization, especially in parameterization 2. This occurs in five out of six cases when $\delta = 0.75$ and in five out of seven cases when $\delta = 0.90$, versus four out of six cases in parameterization 1 ($\delta = 0.75$).²²

□ **Econometric model and results.** A concern with experimental data is that observations within a session are unlikely to be independent of each other (Houser, 2008). First, behavior might depend on past actions or experience (the so-called dynamic session effects, as discussed in Frechette, forthcoming). Further, individual-specific unobservables make each subject's decisions correlated across time. We rely on standard panel data techniques to address these issues and to carry out hypothesis tests of interest.

Specifically, we model individual i 's decision to cooperate in period t , and write it as

$$y_{it} = 1[\gamma z_{it} + c_i + e_{it} > 0], \quad (6)$$

where y_{it} denotes cooperation (1) or defection (0), $1[\cdot]$ is an indicator function, z_{it} is a vector of explanatory variables, $c_i|z_{it} \sim N(\alpha + \psi\zeta_i, \sigma_c^2)$ is an unobserved and random heterogeneity term, and e_{it} is an idiosyncratic error term. We assume a nonzero mean for c_i to deal with the initial-conditions problem (Wooldridge, 2002) and set ζ_i equal to y_{i0} .²³ The vector of parameters γ as well as α , ψ and σ_c^2 are estimated from the data. The vector z_{it} contains indicator variables for the experimental controls of interest (treatment dummies), variables controlling for time aspects of the session (the overall round in the experiment and an indicator for the first round of a supergame), variables that account for experience (own and rival behavior in prior periods),

²² We did not analyze data from the $\delta = 0.60$ -P1 session as the longest supergame in this pair contains only eight rounds.

²³ Here, y_{i0} refers to the choice in the first overall period in the experiment. Other specifications that set ζ_i equal to the mean or the median of y_{it} across all first rounds (in all supergames) produced similar results (available upon request). A similar specification is adopted by Aoyagi and Frechette (2008).

and indicators for prior and/or future demand conditions. Importantly, besides controlling for time-and history-dependent behavior, this specification accounts for correlation across decisions within the same individual through σ_c^2 .²⁴

Effect of demand information and monitoring on collusion. Table 6 displays results of probit regressions of (6) that test whether collusive intensity varies across the three treatments. The table contains three specifications for each of the four δ -parameterization combinations. The three specifications control for the evolution of collusion over time, but they differ in the flexibility with which such evolution is accommodated. Specification 1 allows for a linear trend (the variable “Period”), whereas specifications 2 and 3 allow this trend to vary by treatment. Specification 3 is more flexible, as it allows for a quadratic trend by treatment. Finally, all regressions include an indicator for whether the other seller colluded in the prior period as well as an indicator for the first round of a supergame.²⁵

Before discussing the main results from the regressions, we note a few consistent findings across specifications. First, the strong statistical significance of $\rho = \sigma_c^2 / \sigma_e^2 + 1$ provides support for the random-effects specification. Second, the estimate of ψ indicates that an individual’s propensity to collude at the start of the experiment (ζ_i) appears to be, in all but one case, an important determinant of collusion. Third, collusion is significantly more likely during the first round of a supergame as well as when the other seller colluded in the prior period.

Using specification 1 as a benchmark, it can be seen that although collusion appears to be higher (lower) in the FI (IM) treatment relative to the M treatment, the only statistically significant result is that of the FI treatment in the $\delta = 0.90$ -P2 pair. Also, in three out of the four δ -parameterization pairs, overall collusive behavior appears to depend on the overall number of periods of play; this relationship is negative in one case ($\delta = 0.60$ -P1) and positive in the other two ($\delta = 0.75$ -P2 and $\delta = 0.90$ -P2). When the trend is allowed to vary by treatment (spec. 2), it can be seen that there are substantial differences across treatments as to how collusion evolves. Although collusion decays in all treatments in the $\delta = 0.60$ -P1 pair, the fastest decay occurs in the GP treatment and the slowest in the RS treatment. Results in specification 2 for the three remaining δ -parameterization pairs indicate that collusion (i) increases with time in all cases in the FI treatment, (ii) increases with time in one case in the M treatment ($\delta = 0.75$ -P2), and (iii) decreases with time in one case in the IM treatment ($\delta = 0.75$ -P1). Specification 3 confirms the uneven evolution of collusion across treatments.

Using the estimates from the most flexible specification (spec. 3), we quantified the difference in the predicted probability of collusion between treatments (FI-M, M-IM, and FI-IM) at three different stages of the experiment: start (1st period), middle (61st period) and end (120th period). Figure 5 displays the results. The figure indicates (with dark gray markers) whether the differences are statistically different from zero at the 10% level (or less). The figure shows a general tendency for the computed differences to increase as one moves from less to more collusive-friendly δ -P pairs; this is particularly true for the FI-M comparison (i.e., the demand information effect). Also, the largest differences in the graph occur in period 120 ($\delta = 0.75$ -P1 and $\delta = 0.90$ -P2)²⁷; in these cases, learning seems to be important, but it appears to be present only in the FI treatment. Finally, the joint effect of monitoring and demand information is statistically significant in the three most collusive-friendly environments; however, this difference does not emerge until later in the experiment.

²⁴ Alternatively, one can be agnostic about the precise nature of dependence of observations and allow the variance-covariance matrix of estimates to accommodate for correlation across error terms. Probit and OLS regressions of this type yield similar results (see Appendix G).

²⁵ We opt to report estimated coefficients rather than their marginal effects because interpretation of marginal effects in the presence of interactions (specifications 2 and 3) is not straightforward. After the discussion of Table 6, we describe a way of looking at the marginal effects of monitoring and demand information on collusion in the presence of interactions.

²⁶ Because the remaining variables are 0/1 indicators (First, Rival Colluded, and ψ), their averages are fractions between zero and one. The results of Figure 6 are similar if, instead, the calculation sets these variables at zero (or one).

²⁷ These differences are also the most statistically significant.

TABLE 6 Random Effects Probit Estimates of the Effect of Information and Monitoring on Collusion, 120 Periods

Regressor	Parameterization 1						Parameterization 2					
	$\delta = 0.60$			$\delta = 0.75$			$\delta = 0.75$			$\delta = 0.90$		
	Spec. 1	Spec. 2	Spec. 3	Spec. 1	Spec. 2	Spec. 3	Spec. 1	Spec. 2	Spec. 3	Spec. 1	Spec. 2	Spec. 3
α (constant)	-1.367*** (0.195)	-1.373*** (0.202)	-1.611*** (0.221)	-0.911*** (0.154)	-0.847*** (0.162)	-1.089*** (0.178)	-1.023*** (0.177)	-1.069*** (0.183)	-1.078** (0.197)	-1.090*** (0.166)	-0.932*** (0.173)	-0.896*** (0.188)
FI (RS)	0.066 (0.268)	0.050 (0.280)	0.269 (0.303)	0.273 (0.182)	-0.0172 (0.197)	0.284 (0.220)	0.285 (0.208)	0.306 (0.220)	0.411* (0.241)	0.611*** (0.206)	0.211 (0.220)	0.137 (0.241)
IM (GP)	-0.102 (0.267)	-0.0652 (0.280)	0.4483 (0.303)	-0.178 (0.180)	-0.0463 (0.196)	0.291 (0.218)	-0.317 (0.210)	-0.204 (0.223)	-0.023 (0.244)	0.054 (0.205)	0.034 (0.220)	-0.222 (0.243)
FI $\times$ Period		-0.0024*** (0.0009)	-0.0013 (0.0035)		0.0040*** (0.0008)	0.0012 (0.0032)		0.0025*** (0.0008)	-0.0021 (0.0031)		0.0068*** (0.0008)	0.0149*** (0.0034)
M $\times$ Period		-0.0026*** (0.0011)	0.0089** (0.0043)		-0.0009 (0.0009)	0.0110*** (0.0037)		0.0028*** (0.0009)	0.0033 (0.0037)		-0.0001 (0.0009)	-0.0018 (0.0038)
IM $\times$ Period		-0.0033*** (0.0009)	-0.0169*** (0.0040)		-0.0032*** (0.0008)	-0.0079** (0.0033)		0.0009 (0.0008)	-0.0075** (0.0034)		0.0003 (0.0008)	0.0109*** (0.0035)
FI $\times$ Period ²			-0.000009 (0.00002)			0.00002 (0.00003)			0.00004 (0.00003)			-0.00007** (0.00003)
M $\times$ Period ²			-0.0001*** (0.00003)			-0.0001*** (0.00003)			-0.000004 (0.00003)			0.00001 (0.00003)
IM $\times$ Period ²			0.0001*** (0.00003)			0.00004 (0.00003)			0.00007** (0.00003)			-0.00009*** (0.00003)
Period	-0.003*** (0.001)			0.0001 (0.0005)			0.002*** (0.0005)			0.0025*** (0.0005)		
Rival Colluded	0.693*** (0.041)	0.693*** (0.041)	0.682*** (0.041)	0.702*** (0.034)	0.687*** (0.034)	0.682*** (0.034)	0.860*** (0.034)	0.858*** (0.034)	0.856*** (0.034)	0.769*** (0.035)	0.756*** (0.036)	0.751*** (0.036)
Previous Period	0.408*** (0.041)	0.408*** (0.041)	0.409*** (0.041)	-0.041 (0.042)	-0.042 (0.043)	-0.040 (0.043)	0.265*** (0.043)	0.265*** (0.043)	0.254*** (0.043)	0.346*** (0.059)	0.351*** (0.060)	0.360*** (0.060)
First Round	1.038*** (0.238)	1.038*** (0.238)	1.042*** (0.238)	1.042*** (0.238)	1.042*** (0.238)	1.042*** (0.238)	0.519*** (0.167)	0.519*** (0.167)	0.520*** (0.168)	0.189 (0.178)	0.190 (0.179)	0.189 (0.180)
ψ												
ρ	0.375*** (0.238)	0.375*** (0.238)	0.377*** (0.238)	0.237*** (0.150)	0.240*** (0.151)	0.240*** (0.151)	0.293*** (0.167)	0.294*** (0.167)	0.295*** (0.168)	0.269*** (0.178)	0.273*** (0.179)	0.274*** (0.180)
Log likelihood	-2952	-2952	2942	-4025	-4004	-3997	-3901	-3900	-3895	-3659	-3639	-3631
Number of observations		6902			7616			7616			6902	
Mean dependent variable		0.26			0.33			0.45			0.42	

Note: Dependent variable is equal to 1 if subject chose “Low Output” (i.e., collude) in period t , and zero otherwise. FI, and IM are dichotomous variables that are equal to 1 if the observation corresponds to either of the two treatments, and zero otherwise (i.e., the left-out treatment is M). Period indicates the period number in the session (1–120). Rival Colluded Previous Period is equal to 1 if the other seller colluded in the previous period and zero otherwise. First Round is a dichotomous variable indicating the first round in each supergame. ψ is equal to 1 if subject colluded in the first period of the experiment and zero otherwise. $\rho = \sigma_c^2/(\sigma_c^2 + 1)$. Standard errors are in parentheses. Coefficient significant at ***1%, **5%, and *10%.

FIGURE 5

DIFFERENCES IN THE PREDICTED PROBABILITY OF COLLUSION BETWEEN TREATMENTS AT DIFFERENT STAGES OF THE EXPERIMENT

Note: The estimates in the figure correspond to the difference in probability $P(T1, period|\bar{x};\hat{\theta}) - P(T2, period|\bar{x};\hat{\theta})$, where $T1$ and $T2$ correspond to the treatments being compared, $period$ denotes the period of interest (1, 61, or 120), $\bar{x}$ denotes the average value of the remaining variables (except the two other treatment dummies, which are held at zero), and $\hat{\theta}$ denotes the estimated parameters.²⁶ Standard errors are calculated using the delta method; a dark gray marker indicates that the corresponding p value is smaller than 0.1. Appendix G displays this figure for alternative specifications.

Overall, the results in Figure 5 suggest that monitoring may play a more important role as a determinant of collusion; however, the magnitude and existence of this effect appear to depend both on the specifics of the market environment as well as on experience. In terms of magnitude, the effect of demand information on collusion ranges from being statistically insignificant ($\delta = 0.60$ -P1) to increasing the probability of collusion by 0.33 ($\delta = 0.90$ -P2), whereas the effect of monitoring ranges from being statistically insignificant ($\delta = 0.60$ -P1 and $\delta = 0.90$ -P2) to increasing the probability of collusion by 0.15 ($\delta = 0.75$ -P2).

One last point worth noting is that reciprocating behavior appear to be a strong predictor of collusive behavior: the coefficient on “Rival Colluded Previous Period” has a large magnitude and is always statistically significant. Using the results in specification 3 to compute the corresponding marginal effect reveals that this variable increases the probability of collusion between 0.199 and 0.328 (depending on the δ -P pair).

Theory-related results. In this section, we report probit results for each session. As stated earlier, RS and GP predictions can be linked to demand realizations in different ways. Thus, we choose a specification that tests the effect of future and prior demand states on collusion. Specifically, our probit regression includes the indicator variables Next High and Next Medium (equal to 1 if next period’s demand realization is high or medium, respectively) and Last Low and Last Medium (equal to 1 if the previous period’s demand realization was low or medium, respectively).

If the collusive pattern predicted by RS is present in our data, we would expect collusion to be the lowest when demand is anticipated to be high (coefficient on Next High should be negative). Moreover, collusion should return to higher levels after a high demand period (i.e., Last Medium or Last Low should be negative, or at least nonpositive). Conversely, a prediction of the GP model is that collusion should decline after a negative demand shock (coefficient on Last Low should be negative) but should be unaffected by next period’s demand realization (coefficients on Next High or Next Medium should be unimportant). The variables we just discussed do not capture

TABLE 7 Marginal Effects from Random Effects Probit Estimates, Theoretical Patterns, 120 Periods

Regressor	Parameterization 1				Parameterization 2			
	$\delta = 0.60$				$\delta = 0.75$			
	FI	M	IM	FI	M	IM	FI	M
Next High	-0.521*** (0.059)	-0.023 (0.028)	-0.038 (0.025)	-0.693*** (0.039)	-0.004 (0.035)	-0.043* (0.025)	-0.656*** (0.027)	-0.004 (0.042)
Next Medium	-0.340*** (0.036)	-0.042 (0.026)	-0.003 (0.021)	-0.293*** (0.032)	0.0147 (0.031)	-0.010 (0.023)	-0.127*** (0.028)	-0.012 (0.038)
Last Low	-0.180*** (0.038)	-0.039 (0.028)	-0.094*** (0.032)	-0.205*** (0.040)	-0.001 (0.036)	-0.103*** (0.028)	-0.174*** (0.028)	-0.022 (0.043)
Last Medium	-0.086*** (0.027)	-0.049** (0.023)	-0.074*** (0.025)	-0.139*** (0.030)	-0.035 (0.026)	-0.084*** (0.022)	-0.141*** (0.028)	0.006 (0.032)
Period	-0.001 (0.001)	0.001 (0.0009)	-0.004*** (0.0013)	0.00003 (0.0014)	0.0027** (0.0013)	-0.002** (0.0009)	-0.00077 (0.0013)	0.0009 (0.0015)
Period ²	0.00005 (0.00001)	-0.00001* (0.00001)	0.00003*** (0.0001)	0.0001 (0.0001)	-0.00002** (0.00001)	0.0001 (0.00001)	0.00002*** (0.00001)	-0.00001 (0.00001)
Rival Colluded	0.309*** (0.038)	0.288*** (0.053)	0.114*** (0.029)	0.431*** (0.027)	0.356*** (0.028)	0.109*** (0.023)	0.426*** (0.026)	0.525*** (0.024)
Previous Period	0.128*** (0.027)	0.146*** (0.036)	0.069*** (0.024)	0.139*** (0.031)	-0.137*** (0.026)	0.029 (0.022)	0.160*** (0.028)	0.066** (0.026)
First Round	0.216** (0.104)	0.345 (0.230)	0.410*** (0.138)	0.163 (0.115)	0.089 (0.080)	0.043 (0.083)	0.096 (0.119)	0.305** (0.122)
ψ	0.331** (0.107)	0.396*** (0.142)	0.436*** (0.138)	0.316*** (0.121)	0.186*** (0.108)	0.274*** (0.135)	0.340*** (0.125)	0.319*** (0.122)
Log likelihood	-1027 (2380)	-763 (2142)	-923 (2380)	-1210 (2618)	-1087 (2142)	-1354 (2856)	-1255 (2856)	-971 (2142)
Number of observation	0.30	0.21	0.26	0.43	0.32	0.24	0.59	0.45
Mean dependent Variable								

Note: Dependent variable is equal to 1 if subject chose "Low Output" (i.e., collude) in period t , and zero otherwise. Next High is equal to 1 if next period's demand is high, and zero otherwise. Next Medium is equal to 1 if next period's demand is medium, and zero otherwise. Last Low is equal to 1 if the previous period's demand was low, and zero otherwise. Last Medium is equal to 1 if the previous period's demand was medium, and zero otherwise. First is a dichotomous variable indicating the first round in each supgame. Period indicates the period number in the session (1–120). Rival Colluded Previous Period is equal to 1 if the other seller colluded in the previous period and zero otherwise. ψ is equal to 1 if subject colluded in the first period of the experiment and zero otherwise. All specifications include a constant (not reported). Standard errors are in parentheses. Bold numbers indicate whether the marginal effect is significantly different (at the 5% level) from the other included category (i.e., Next Medium vs. Next High and Last Medium vs. Last Low).

Significant at *** 1%, ** 5%, and * 10%. See Appendix G for additional specifications.
^a $\rho = \sigma_c^2 / (\sigma_c^2 + 1)$; this value corresponds to the estimate from the original regression (i.e., not a marginal effect).

FIGURE 6

EFFECT OF NEXT AND PAST DEMAND REALIZATIONS ON PROBABILITY OF COLLUSION, FI AND IM TREATMENTS

Note: Vertical bars denote 95% confidence intervals. Next Is Low is the reference category for Next Is High and Next Is Medium. Last Was High is the reference category for Last Was Low and Last Was Medium.

all features of the theoretical predictions, for example, the existence of price wars in the GP treatment. We discuss additional evidence that addresses this issue in Web Appendix H.

Table 7 reports the estimated marginal effects of our probit specification for each treatment in every δ -P pair. The M treatment is included only for comparison purposes. Appendix G contains results of alternative specifications. Results on explanatory variables from the fifth regressor onward are consistent with those obtained in Table 6. Among these variables, the other seller's behavior appears to be the main determinant of collusive behavior; its magnitude ranges from 0.109 to 0.525 and it is always statistically significant.

Figure 6 helps visualize the importance of variables of theoretical interest (regressors 1–4) by graphically contrasting their marginal effects in the FI and M treatments.²⁸ It can be seen that a high demand state in the FI treatment plays a key role: the probability of collusion is lower before its occurrence (between -0.181 and -0.529 with respect to medium demand and between -0.521 and -0.693 with respect to low demand) but higher afterward (between 0.086 and 0.141 with respect to medium demand and between 0.174 and 0.205 with respect to low demand).²⁹ In other words, collusion significantly breaks down in anticipation of a high demand period, but this does not seem to interfere with subjects' ability to later maintain high collusion levels in medium and low demand states. This countercyclical pattern appears to also hold between medium and low demand periods, although it becomes less important as one moves to more collusive-friendly δ -P pairs.

Conversely, the most important determinant of collusion in the IM treatment (among the four regressors in the figure) is past demand: collusion is lower after a low demand period (between -0.064 and -0.185 with respect to high demand and between -0.008 and -0.051 with respect to medium) although this difference is only statistically significant with respect to a high demand realization.

Although we interpret the results above as consistent with the predictions of both the RS and GP theories, we believe that subjects' ability to tacitly collude in the way predicted by RS is the most salient feature in our experiments. This level of complexity in collusive strategies can be seen in subjects' responses to the open-ended question mentioned earlier, which are shown in Web Appendix H.

²⁸ Consistent with expectations, results in Table 7 indicate that these four regressors are only important in the FI and IM treatments.

²⁹ The difference in probability of collusion between high and medium demand periods (not reported in the figure) is calculated by subtracting the marginal effect on Next Medium from the marginal effect on Next High.

6. Concluding remarks

■ In this article, we focus on two factors that have been prominent in the theory of infinitely repeated games with stochastic demand: monitoring and demand information. Guided by theory, we construct experiments to study the effect of these two factors on collusion. In addition, we explore the extent to which theoretical predictions from two prominent models find support in our data.

The results of our investigation can be grouped into two findings. First, demand information might facilitate collusion more than monitoring. We believe that this result needs further investigation, as results appear to depend both on the parameterization of the experiment as well as on subjects' experience. For example, our experimental design has the feature that there is a large difference between payoffs across demand states, which may make demand information relatively more important than monitoring in this setting. On the other hand, as payoffs in demand states approach one another, monitoring becomes unnecessary. A hypothesis worth exploring is whether and how demand information continues to be as important as the distribution of demand states is tighter.

Our second result is that we find strong support for the countercyclical collusive pattern predicted by the RS theory. The large and statistically significant results across all variants of our experiment as well as across all descriptive and econometric tests allow us to attach more confidence to this finding. This result is interesting for two reasons. First, a large portion of subjects are able to achieve collusion in a stochastic market environment even in the absence of communication. Second, this type of collusion is fairly sophisticated as behavior follows a state-dependent strategy.

In general, we interpret the empirical evidence as highlighting the importance of uncertainty on collusion: collusion is highest in the most certain environment (RS) and lowest in the least certain one (GP). The certainty granted by the precise anticipation of next period's demand is particularly important. One can interpret this feature of the RS environment as an effective mechanism through which firms can coordinate on a collusive equilibrium. More specifically, one can imagine that the RS theory has an advantage over other theories in that tomorrow's demand state can act as a public correlation device.

The RS and GP theories are used to motivate our research and to design our experiments. Thus, this article should not be viewed as a test of one theory versus another, nor should we interpret the weaker evidence for GP-type predictions as absolute absence of empirical support for this influential model. It is possible that the predictions of the GP model are valid in other market environments or experimental designs beyond the scope of this article. Although our results are informative as to how complex tacit collusion can be in practice, we see a pressing need for more research looking at this issue.

Appendix A. Feasible equilibria

□ **The full information treatment (RS theory).** Besides the two "collusive" equilibria described in the text (Table 3), playing the one-shot NE, (H, H) , in every period is, as usual, an equilibrium. In addition, as shown by Stahl (1991), the strategy pair (L, H) can also be sustained as an equilibrium outcome. Adapting Stahl's approach to our stochastic demand structure, the strategy pair (L, H) can be an equilibrium outcome in demand state s if there exists at least one payoff Π^* for player 1, such that the following condition is met:^{i,ii}

$$(1 - \delta)\Pi_s^{ND} + \delta E(\Pi^*) \geq (1 - \delta)\Pi_s^{NE} + \delta E(\Pi^*),$$

where $E(\Pi^*) = 0.2\Pi_h^* + 0.6\Pi_m^* + 0.2\Pi_l^*$. The left-hand side consists of the payoff of playing L today (when the opponent is playing H) plus the expected discounted profits of obtaining profits Π^* for each of the remaining rounds of the game; the right-hand side consists of the payoff of switching to H today (when the opponent is playing H) plus the expected discounted payoffs of playing a strategy that yields a payoff of Π^* . By setting Π^* equal to Π^{NE} (i.e., the grim strategy),

ⁱ Stahl shows that only player 1's condition is the relevant one.

ⁱⁱ Stahl's paper is known as a study of collusion with correlated equilibria. The analysis here uses Stahl's results that are relevant for our case (i.e., in the absence of a public randomization signal).

TABLE A1 Range of CRRA Coefficients for Which the $(H,L)/(L,H)$ Outcome Is a Feasible Equilibrium in the FI Treatment, by Demand State and δ -P Pair

Demand State	Parameterization 1		Parameterization 2	
	$\delta = 0.60$	$\delta = 0.75$	$\delta = 0.75$	$\delta = 0.90$
High (h)	$(-\infty, +\infty)$	$(-\infty, +\infty)$	$(-\infty, +\infty)$	$(-\infty, +\infty)$
Medium (m)	$(-\infty, +\infty)$	$(-\infty, +\infty)$	$(-\infty, +\infty)$	$(-\infty, +\infty)$
Low (l)	$(-\infty, +1.4]$	$(-\infty, +1.9]$	$(-\infty, +2.4]$	$(-\infty, +3.8]$

TABLE A2 Range of CRRA Coefficient for Which the GP Equilibria in the IM Treatment Are Feasible; Different Threshold Levels: $[N = 1]$, $[N = \infty]$

Threshold	Parameterization 1		Parameterization 2	
	$\delta = 0.60$	$\delta = 0.75$	$\delta = 0.75$	$\delta = 0.90$
p_0	Never	Never	Never	Never
p_1	$[3.7, +\infty)$; $[2.2, +\infty)$	$[3.4, +\infty)$; $[1.2, +\infty)$	$[2.4, +\infty)$; $[0.2, +\infty)$	$[2.1, +\infty)$; $[-2, +\infty)$
p_2	$[2.2, +\infty)$; $[0.97, +\infty)$	$[1.9, +\infty)$; $[0.2, +\infty)$	$[0.85, +\infty)$; $[-1, +\infty)$	$[0.6, +\infty)$; $[-2.1, +\infty)$
p_3	$[4.3, +\infty)$; $[3.5, +\infty)$	$[4.1, +\infty)$; $[3.2, +\infty)$	$[3, +\infty)$; $[2.1, +\infty)$	$[2.9, +\infty)$; $[1.8, +\infty)$
p_4	Never	Never	Never	Never

one makes the right-hand side as small as possible, thus making the following condition a sufficient one for (L,H) to be an equilibrium:

$$E(\Pi^x) \geq \frac{(1 - \delta)(\Pi_s^{NE} - \Pi_s^{ND}) + \delta E(\Pi^{NE})}{\delta}.$$

In our case, it suffices to check whether, for each demand state, the above condition holds when we set $\Pi^x = \Pi^C$. If it does, then, by symmetry, the equilibrium (H,L) is also feasible; it is important to note, again, that these equilibria are *state dependent*. In our design, the (H,L) equilibrium is feasible under the grim-trigger strategy for almost all values of the CRRA parameter (see Table A1).

□ **The imperfect monitoring treatment (GP theory).** As with the FI treatment, it is easy to show that playing the one-shot NE (H,H) in every period is an equilibrium of this game. As opposed to the FI treatment, the collusive equilibria that we consider here refer to the “collusive path” proposed by GP: reversion to the one-shot NE occurs for some period of time (finite or infinite).ⁱⁱⁱ In this type of equilibrium, two variables need to be calibrated: the punishment length (N) and the threshold level (*price*). As shown by Abreu, Pearce, and Stacchetti (1990), the set of equilibrium payoffs can be obtained by considering a grim-trigger strategy (or less severe punishments) and then finding the smallest public signal for which collusion is incentive compatible. Using Abreu et al.’s insight, it is straightforward to compute the feasible set of “collusive” equilibria and the corresponding range of incentive-compatible punishment lengths. We proceed as follows. For each of the five possible threshold levels (i.e., each of the five prices implied by our prisoner’s dilemma structure; see Appendix B), we compute the range of the CRRA coefficients for which the $N = 1$ and $N = \infty$ GP outcomes (the two extreme GP equilibria) satisfy condition (4). Table A2 presents the results of this calculation. Note that except for the two extreme threshold levels (p_0 and p_4), the GP equilibria are feasible. The CRRA feasibility range grows as one moves to more collusive-friendly δ -P pairs as well as with punishment length. The analysis in the article (Table 3 and Section 6) considers p_2 as it accommodates the GP equilibrium for the broadest set of risk preferences (including, in all but one case, risk aversion, $r = 0$).

Appendix B

□ **Imperfect public monitoring.** Consider parameterization 1. Without loss of generality, let us assume that the collusive quantity (L) is equal to 1; given our design, $L = 1$ automatically defines H (the high quantity), as well as the prices (the public signal) in all the cells of the three payoff tables (high, medium, and low demand), in Table B1.

Note that the noisy price signal (indirectly) received by both sellers is the same, and hence it is public. For example, suppose that seller 1 chooses to play the collusive outcome (L) and that seller 2 chooses to defect (H). Further, suppose

ⁱⁱⁱ Because monitoring is imperfect, Stahl’s results regarding the feasibility of $(H,L)/(L,H)$ do not apply to this case.

TABLE B1 Implied High Quantity (H) and Prices (p_i) in Parameterization 1 when $L = 1$

		High Demand Seller 2		Medium Demand Seller 2		Low Demand Seller 2	
		$L = 1$	$H = 5.73$	$L = 1$	$H = 5.73$	$L = 1$	$H = 5.73$
Seller 1	$L = 1$	$p_4 = 26$	$p_3 = 7.5$	$p_3 = 7.5$	$p_2 \approx 2.1$	$p_2 = 2.1$	$p_1 \approx 0.6$
	$H = 5.73$	$p_3 = 7.5$	$p_2 \approx 2.1$	$p_2 \approx 2.1$	$p_1 \approx 0.6$	$p_1 \approx 0.6$	$p_0 = 0.17$

Notes: $H = 5.73$ is obtained by dividing 43 (player 2's profit when player 1 chooses L , player 2 chooses H , and high demand occurs) by 7.5 (the price that must hold in the L,L cell of the medium-demand state and in the L,H cell of the high-demand state). A slightly different H (because of rounding of profit numbers) can be obtained by dividing 12.5 (player 2's profit when player 1 chooses L , player 2 chooses H , and medium demand occurs) by 2.1 (the price that must hold in the L,L cell of the low-demand state and in the L,H cell of the medium-demand state). Because of rounded profit (which is easier for subjects to understand), the implied prices in some instances have also been approximated (the ones with a " $\approx$ ").

that medium demand occurred. Then, seller 1 receives the noisy price signal of p_2 (upper right cell of the medium-demand matrix or upper left cell of the low-demand matrix) and so does seller 2 (upper right cell of the medium-demand matrix or lower right cell of the high-demand matrix). A similar exercise can demonstrate the public nature of monitoring for parameterization 2. As indicated in the second section of Appendix A, the analysis in the article focuses on the noisy signal p_2 .

Importantly, the public nature of the noisy signal is achieved for any number one wishes to choose for L . More generally, one can think of this design as a symmetric two-player Cournot game with two strategies and stochastic demand. This game maps choices $y = (y_1, y_2)$ onto price through a stochastic (demand) function, generating a conditional distribution $f(p|y)$; in our case f has discrete support, that is, $p \in (p_0, \dots, p_4)$.

Web Appendix C (To view web appendix, please visit <http://onlinelibrary.wiley.com/doi/10.1111/rand.2012.43.issue-1/issuetoc>)

Appendix D

TABLE D1 Sequence of Supergame Lengths

Supergame Number	Number of Rounds		
	$\delta = 0.60$	$\delta = 0.75$	$\delta = 0.90$
1	7	7	7
2	4	4	28
3	2	5	10
4	3	10	5
5	2	9	6
6	3	6	29
7	5	1	11
8	7	3	2
9	2	5	9
10	6	6	5
11	1	26	1
12	3	3	7
13	5	8	
14	1	3	
15	4	2	
16	1	1	
17	4	3	
18	4	5	

TABLE D1 continued

Supergame Number	Number of Rounds		
	$\delta = 0.60$	$\delta = 0.75$	$\delta = 0.90$
19	8	2	
20	1	3	
21	4	1	
22	2	6	
23	3	1	
24	3		
25	1		
26	7		
27	3		
28	2		
29	1		
30	3		
31	4		
32	1		
33	1		
34	1		
35	3		
36	1		
37	6		
38	1		

Appendix E

Decision and profit screens for the three treatments (parameterization 1)

FIGURE E1

DECISION SCREEN IN THE FULL INFORMATION (FI) TREATMENT

FIGURE E2
DECISION SCREEN IN THE MONITORING (M) AND IMPERFECT MONITORING (IM) TREATMENTS

FIGURE E3
PROFIT SCREEN IN THE FULL INFORMATION (FI) AND MONITORING (M) TREATMENTS

FIGURE E4
PROFIT SCREEN IN THE IMPERFECT MONITORING (IM) TREATMENT

Appendix F

Additional descriptive results

FIGURE F1
FRACTION OF TIMES THAT SUBJECTS CHOSE THE COLLUSIVE OUTPUT, FIRST ROUND OF A SUPERGAME VERSUS ALL OTHER PERIODS (BY SESSION)

See notes in Figure 1 and text for acronym details.

FIGURE F2

FRACTION OF TIMES THAT SUBJECTS CHOSE THE COLLUSIVE OUTPUT, PERIODS FOLLOWING OTHER SELLER'S DECISION TO COLLUDE VERSUS PERIODS FOLLOWING OTHER SELLER'S DECISION TO DEVIATE (BY SESSION)

FIGURE F3

FRACTION OF TIMES THAT SUBJECTS CHOSE THE COLLUSIVE OUTPUT, PERIOD AFTER FIRST LOW-DEMAND REALIZATION IN A SUPERGAME VERSUS PERIODS PRIOR TO FIRST LOW-DEMAND REALIZATION IN A SUPERGAME

FIGURE F4
EVOLUTION OF SUBJECTS CHOOSING THE COLLUSIVE OUTPUT, 120 PERIODS, FI TREATMENT (RS THEORY, PARAMETERIZATION 1)

FIGURE F5
EVOLUTION OF SUBJECTS CHOOSING THE COLLUSIVE OUTPUT, 120 PERIODS, FI TREATMENT (RS THEORY, PARAMETERIZATION 2)

Note: Squares denote the first round of a supergame; circles denote a high-demand period.

TABLE G1 Sign and Statistical Significance of “Late in The Experiment” Variable When Included in Regressions Reported in Table 6

Cutoff period	Parameterization 1						Parameterization 2					
	$\delta = 0.60$			$\delta = 0.75$			$\delta = 0.75$			$\delta = 0.90$		
	Spec.1	Spec. 2	Spec.3	Spec.1	Spec. 2	Spec.3	Spec.1	Spec. 2	Spec.3	Spec.1	Spec. 2	Spec.3
100	—	—	—	—	—	— [†]	+	+	—	—	—	+
110	+ [†]	+ [†]	+ [†]	—	—	—	+	+	—	—	—	+

[†]Significant at the 5% level or lower.

TABLE G2 Sign and Statistical Significance of “Late in the Experiment” Variable When Included in Regressions Reported in Table 7

Cutoff period	Parameterization 1						Parameterization 2					
	$\delta = 0.60$			$\delta = 0.75$			$\delta = 0.75$			$\delta = 0.90$		
	FI	M	IM	FI	M	IM	FI	M	IM	FI	M	IM
100	—	—	—	—	—	—	—	+	—	—	+	+ [†]
110	+	+ [†]	+ [†]	—	+	—	—	—	+	—	+	+ [†]

[†]Significant at the 5% level or lower.

Appendix G

Additional specifications

- **End of game effect.** In this section, we test for the existence of an end-of-game effect by rerunning regressions in the first and second subsections of Section 5 with an additional dichotomous variable, which takes a value of 1 for “late” periods in the experiment and zero otherwise. If an end-of-game effect is present in our data, the estimated coefficient on this variable should be negative and statistically significant. We carry out this exercise with two cutoff points for the “late in the experiment” dichotomous variable (100 periods and 110 periods). For the sake of brevity, Tables G1 and G2 display the sign of the estimated coefficient and information about its statistical significance. As can be seen, the dichotomous variable is negative and statistically significant in only one of the 48 regressions.
- **Robustness checks for Section 5.** Here we report results from probit specification and a linear probability model (OLS), without random effects; standard errors are clustered at the subject level (clustering at the session level yields similar results). We also report results of a probit specification with random effects that restricts the data to the two longest supergames in each δ -P pair. For the sake of brevity, we report the results graphically as in Figures 6 and 7 in the paper.
- **Robustness check for Figure 6**

FIGURE G1

PROBIT RESULTS, CLUSTERED STANDARD ERRORS, ALL SUPERGAMES

FIGURE G2

OLS RESULTS, CLUSTERED STANDARD ERRORS, ALL SUPERGAMES

□ Robustness Check for Figure 6

FIGURE G3

PROBIT RESULTS, CLUSTERED STANDARD ERRORS, ALL SUPERGAMES

FIGURE G4

OLS RESULTS, CLUSTERED STANDARD ERRORS, ALL SUPERGAMES

FIGURE G5

PROBIT RESULTS, RANDOM EFFECTS, TWO LONGEST SUPERGAMES

FIGURE G6

OLS RESULTS, CLUSTERED STANDARD ERRORS, TWO LONGEST SUPERGAMES

Web Appendix H (To view web appendix, please visit

<http://onlinelibrary.wiley.com/doi/10.1111/rand.2012.43.issue-1/issuetoc>

References

- ABBINK, K. AND BRANDTS, J. "Collusion in Growing and Shrinking Markets: Empirical Evidence from Experimental Duopolies." In J. HINLOOPEN AND H.-T. NORMANN, eds., *Experiments and Antitrust Policy*. Cambridge, UK: Cambridge University Press, 2009.
- ABREU, D., PEARCE, D., AND STACCHETTI, E. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica*, Vol. 5 (1990), pp. 1041–1063.
- AOYAGI, M. AND FRECHETTE, G. "Collusion as Public Monitoring Becomes Noisy: Experimental Evidence." *Journal of Economic Theory*, Vol. 144 (2008), pp. 1135–1165.
- BAGWELL, K. AND STAIGER, R.W. "Collusion over the Business Cycle." *RAND Journal of Economics*, Vol. 28 (1997), pp. 82–106.
- BEREBY-MEYER, Y. AND ROTH, A. "The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation." *American Economic Review*, Vol. 96 (2006), pp. 1029–1042.
- BORENSTEIN, S. AND SHEPARD, A. "Dynamic Pricing in Retail Gasoline Markets." *RAND Journal of Economics*, Vol. 27 (1996), pp. 429–451.
- CHEVALIER, J., KASHYAP, A., AND ROSSI, P. "Why Don't Prices Rise during Periods of Peak Demand? Evidence from Scanner Data." *American Economic Review*, Vol. 93 (2003), pp. 15–37.

- DAL BÓ, P. "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games." *American Economic Review*, Vol. 95 (2005), pp. 1591–1604.
- AND FRECHETTE, G. "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence." *American Economic Review*, Vol. 101 (2011), pp. 411–429.
- DUFFY, J. AND OCHS, J. "Cooperative Behavior and the Frequency of Social Interaction." *Games and Economic Behavior*, Vol. 66 (2009), pp. 785–812.
- ELLISON, G. "Theories of Cartel Stability and the Joint Executive Committee." *RAND Journal of Economics*, Vol. 25 (1994), pp. 37–57.
- FEINBERG, R. AND SNYDER, C. "Collusion with Secret Price Cuts: An Experimental Investigation." *Economics Bulletin*, Vol. 3 (2002), pp. 1–11.
- FISCHBACHER, U. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments—Experimenter's Manual." *Experimental Economics*, Vol. 10 (2007), pp. 171–178.
- FRECHETTE, G. "Session-Effects in the Laboratory." *Experimental Economics* (forthcoming).
- FUDENBERG, D. AND TIROLE, J. "Non-Cooperative Game Theory for Industrial Organization: An Introduction and Overview." In R. Schmalensee and R. Willig, eds., *Handbook of Industrial Organization*, Vol. 1. Amsterdam: Elsevier Science, 1989.
- , RAND, D., AND DREBER, A. "Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World." *American Economic Review* (forthcoming).
- GREEN, E. AND PORTER, R. "Noncooperative Collusion under Imperfect Price Competition." *Econometrica*, Vol. 52 (1984), pp. 87–100.
- HALTIWANGER, J. AND HARRINGTON, J. "The Impact of Cyclical Demand Movements on Collusive Behavior." *RAND Journal of Economics*, Vol. 22 (1991), pp. 89–106.
- HARRISON, G. AND RUSTRÖM, E. "Risk Aversion in the Laboratory." In J.C. Cox AND G.W. Harrison, eds., *Risk Aversion in Experiments*, Vol. 12. Bingley, UK: Emerald, 2008.
- HOLCOMB, J. AND NELSON, P. "The Role of Monitoring in Duopoly Market Outcomes." *Journal of Socio-Economics*, Vol. 26 (1997), pp. 79–93.
- HOLT, C. "Industrial Organization: A Survey of Laboratory Research." In J.H. Kagel AND A.E. Roth, eds., *Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press, 1995.
- HOUSER, D. "Experiments and Econometrics." In S. DURLAUF AND L. BLUME, eds., *New Palgrave Dictionary of Economics*, 2d ed. New York: Palgrave-MacMillan, 2008.
- KANDORI, M. "Correlated Demand Shocks and Price Wars during Booms." *Review of Economic Studies*, Vol. 58 (1991), pp. 171–180.
- PORTER, R. "A Study of Cartel Stability: The Joint Executive Committee, 1880–1886." *Bell Journal of Economics*, Vol. 14 (1983), pp. 301–314.
- ROTEMBERG, J.J. AND SALONER, G. "A Supergame-Theoretic Model of Price Wars during Booms." *American Economic Review*, Vol. 76 (1986), pp. 390–407.
- RUFFLE, B. "When Do Large Buyers Pay Less? Experimental Evidence." Unpublished manuscript, 2009. http://www.bgu.ac.il/~bradley/#Papers_under_Review.
- STAHL, D. "The Graph of Prisoner's Dilemma Supergame Payoffs as a Function of the Discount Factor." *Games and Economic Behavior*, Vol. 3 (1991), pp. 368–384.
- WEBER, E.U., SHAFIR, S., AND BLAIS, A.-R. "Predicting Risk-Sensitivity in Humans and Lower Animals: Risk as Variance or Coefficient of Variation." *Psychological Review*, Vol. 111 (2004), pp. 430–445.
- WOOLDRIDGE, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass.: MIT Press.