Optimal Collusion with Private Information

Susan Athey and Kyle Bagwell

First Draft: May, 1998. This Draft: February, 2001.

ABSTRACT: We consider an infinitely-repeated Bertrand game, in which prices are publicly observed and each firm receives a privately observed, i.i.d. cost shock in each period. Productive efficiency is possible only if high-cost firms are willing to relinquish market share. In the most profitable collusive schemes, firms implement productive efficiency, and high-cost firms are favored with higher expected market share in future periods. If types are discrete, there exists a discount factor strictly less than one above which first-best profits can be attained purely through history-dependent reallocation of market share between equally efficient firms. We provide further characterizations and several computational examples. We next examine different institutional features. We find that firms may regard explicit communication (smoke-filled rooms) as to costs beneficial after some histories but not others. Finally, if firms can make explicit side-payments and these entail any inefficiency (e.g., if they are illegal and bear some risk of detection), then optimal collusive equilibria involve the use of future market-share favors.

JEL Classification Numbers: C73, L13, L41.

Keywords: Collusion, cartel design, repeated games, private information, mechanism design, auctions, information-sharing, communication.

---

*M.I.T. and NBER (Athey) and Columbia University and NBER (Bagwell). We are especially grateful to Pavel Grigoriev for exceptional research assistance and to Eric Maskin for providing detailed comments on an earlier draft. We thank Peter Cramton, Glenn Ellison, Michihiro Kandori, Wally Mullin, David Pearce, Michael Riordan, Lars Stole, Jean Tirole, two anonymous referees and seminar participants at Barcelona (Pompeu Fabra), Berkeley, Cal Tech, Federal Reserve Bank Board of Governors, Harvard, M.I.T., Minneapolis Federal Reserve Bank, S.I.T.E., Stanford, Toulouse (I.D.E.I.), UCLA, UC-San Diego, University of Virginia, Wisconsin, the Summer Conference on I.O. in Toronto, the Stony Brook Game Theory Conference, and the Winter 2000 Meetings of the Econometric Society for helpful discussions. We thank the NSF (Athey: SBR-9631760 and SES-9983820; Bagwell: SES-9905460) for generous financial support. Athey thanks the Cowles Foundation at Yale University for hospitality in 1997-1998, when much of this research was conducted, as well as the National Fellows Program of the Hoover Institution for support in the 2000-1 academic year.

1. Introduction

Anti-trust law and enforcement vary widely over time, across countries and between industries. For example, as Stocking and Watkins (1946) detail, U.S. anti-trust policy was relatively permissive in the first part of the 20th century, and industry associations in which firms shared information and records, allocated market shares and fixed prices, and exchanged side-payments were commonly observed.[1] The recent U.S. policy, by contrast, is considerably more antagonistic. The U.S. Antitrust Division's Revised Amnesty Program (1993) provides incentives for firms to self-report collusive conduct, and this has led to the prosecution of a number of "hard-core" cartels, often operating in international markets and characterized by "price fixing, bid-rigging, and market- and customer-allocation agreements" (Griffin (2000)).[2] Levinsohn (1995) describes the significant variation in anti-trust law and enforcement that is found across countries. And significant variation also occurs within countries and between industries; for example, in many countries (including the U.S.), the legal stance toward cartels is more tolerant in export industries.

The different manifestations of anti-trust policies naturally affect the organizational structure of collusive activity. If the anti-trust environment is permissive, then firms may set up a formal organization, in which they set prices and allocate sales, communicate about current circumstances, keep records of past experiences and exchange side-payments. On the other hand, when the anti-trust policy is antagonistic, the organization of collusive activity may be more secretive and less formal. Firms may avoid direct meetings altogether. Or they may communicate surreptitiously, in "smoke-filled rooms." And firms might also refrain from direct side-payments, which leave a "paper trail."

The implications of anti-trust policies for collusive conduct are more subtle. In its perfected form, collusion enables a group of firms to conduct themselves as would a single firm: prices are set and market-shares are allocated in a manner that maximizes joint profits. In practice, however, the road to perfection contains obstacles. One important obstacle is impatience: high prices can be enjoyed only if firms are sufficiently patient that they resist the temptation to undercut. A further obstacle is that firms naturally possess private information as to their respective circumstances. At a given time, some firms may have high costs while others enjoy low costs, due to variations in local conditions, labor relations, inventory management and so on. The market-share allocation that achieves productive efficiency then may be feasible only if firms communicate cost information, and truthful communication may be possible only if higher-cost firms are assured of side-payments or some future benefit. In broad terms, anti-trust policy affects collusive conduct by influencing the "instruments" that firms may use when encountering such obstacles.

[1] Sophisticated cartels of this kind were found in the steel, aluminum, incandescent electric lamp and sugar (see Genesove and Mullin (1998, 1999)) and shipping (see Deltas, Serfes and Sicotte (1999)) industries.

[2] Prominent examples include the lysine, vitamin, graphite electrode, and citric-acid cartels.

This perspective suggests three questions concerning the optimal collusion of impatient firms. First, how does the presence of private information among firms affect collusive profits? In particular, is it possible for privately informed firms to construct a self-enforcing collusive scheme in which they act as would a single firm and thereby achieve first-best profits? Second, how does the presence of private information affect collusive conduct? In particular, when privately informed firms collude, what are the implications for market prices and shares? Finally, how do anti-trust policies affect collusive profits and conduct? In particular, what are the consequences of restrictions on communication and side-payments for collusive profits and conduct?

These are basic questions whose resolution might offer practical insights. For example, a theory that answers these questions might provide a lens through which to interpret observed (historic or current) collusive conduct in terms of the surrounding anti-trust environment. And it also might provide a framework with which to better predict the consequences of a change in anti-trust policies for collusive conduct. Nevertheless, in the literature on self-enforcing collusion, these questions are as yet unanswered. Indeed, as we explain below, even the most basic issues - e. g., how might communication among firms facilitate collusion? - are poorly understood.

Motivated by these considerations, we develop here a theory of optimal collusion among privately informed and impatient firms, and we examine how the level and conduct of collusion varies with the anti-trust environment. The modeling framework is easily described. We consider an infinitely repeated Bertrand game, in which prices are publicly observed and each firm receives a privately observed, i.i.d. cost shock in each period. We assume further that demand is inelastic, there are two firms and each firm's unit-cost realization is either "high" or "low." These assumptions simplify our presentation. Our main findings would emerge as well in a model with finite numbers of firms and cost types.

To understand our findings, it is helpful to recall the theory of the legalized cartel, in which side-payments can be enforced by binding contracts. As Roberts (1985), Cramton and Palfrey (1990) and Kihlstrom and Vives (1992) have shown, the central tradeoffs are then well-captured in a static mechanism design model. An important consideration for the cartel is that production is allocated efficiently over cartel members, but when firms are privately informed as to their respective costs of production this requires communication and transfers. Communication enables firms to establish before production the identity of the lowest-cost firm, while transfers (from this firm to the other cartel members) ensure that firms have the incentive to communicate truthfully.

Outside of a legalized cartel, however, the collusive relationship must be self-enforcing, and anti-trust policies may restrict the manner in which firms interact. Thus, we characterize optimal collusive conduct among privately informed firms that interact repeatedly in environments that are distinguished on the basis of restrictions on the instruments available to the firms. In our base model, we make the following assumptions: (i). firms can communicate with regard to current cost conditions; and (ii). firms cannot make side-payments (use "bribes"). We show that optimal collusion involves extensive use of "market-share favors," whereby individual firms are treated asymmetrically as a reward or punishment for past behavior. After studying this model in some detail, we then analyze the way in which optimal collusion changes as each of the two assumptions is relaxed.

Our modeling approach is to recast our repeated private-information game as a static mechanism, similar to that analyzed in the legalized-cartel literature. To this end, we follow Abreu, Pearce and Stacchetti (1990) and Fudenberg, Levine and Maskin (1994) and observe that Perfect Public Equilibrium (PPE) payoffs for the firms can be factored into two components: current-period payoffs and (discounted) continuation values. This suggests that PPE continuation values can play a role like that of side-payments in the legalized-cartel literature, although transfers are now drawn from a restricted set (namely, the set of PPE continuation values). In this way, we argue that firms who are prohibited from making side-payments can still implement a self-enforcing scheme, in which communication has potential value, where in place of a side-payment from one firm to another, the collusive mechanism specifies that one firm is favored over another in future play.

While this analogy is instructive, the two approaches have important differences. Suppose that firm 1 draws a low-cost type while firm 2 draws a high-cost type. In the legalized-cartel model, firm 2 would reveal its cost type and not produce, anticipating that it would then receive a transfer. In our base model, firm 2 would likewise report its high-cost type, expecting to receive its "transfer" in the form of a more favorable continuation value. In turn, this value can be delivered, if firm 2 receives future market-share favors, corresponding to future cost states in which firm 2's market share is increased. But here key differences appear. First, if the required transfer is too large, there may not exist a PPE that yields the necessary continuation value for firm 2. Second, even if the corresponding PPE value does exist, when the transfer is achieved through an adjustment in future play, the transfer may involve an inefficiency: the strategies that achieve this transfer may involve firm 2 enjoying positive market share in some future state in which it alone has high costs.

This second difference directs attention to an interesting feature of our base model. Future play is burdened with two roles: in a given future period, production must simultaneously (i). serve a transfer role, rewarding firms for past revelations of high costs, and (ii). serve an efficiency role, allocating production as efficiently as possible in the future period itself. These roles may conflict. We show, however, that no conflict emerges, so long as firms are sufficiently patient. In particular, our first general finding is as follows: for the base model, and for a wide range of parameter values, there exists a critical discount factor that is strictly less than one and above which the cartel can achieve first-best profits in every period. Intuitively, firms disentangle the two roles for future play, if they limit transfer activities to future "ties," in which both firms are equally efficient. If the discount factor is sufficiently high, the transfers so achieved are sufficient to ensure truth-telling.

This finding is of broader interest. It generalizes a related finding by Fudenberg, Levine and Maskin (1994), who consider a family of repeated private-information games and show that first-best payoffs can be reached in the limit as the discount factor goes to unity. By contrast, making use of our assumption of a finite number of types, we show that first-best payoffs can be achieved exactly, by firms that are not infinitely patient, and we offer an explicit construction of the efficient PPE. To our knowledge, this is the first construction of a first-best PPE in a repeated private-information game, when players are impatient.[3]

In addition, this finding raises an important qualification for a common inference that is drawn in empirical studies of market-share stability. In many studies, such as those offered by Caves and Porter (1978), Eckard (1987) and Telser (1964), an inference of greater collusive (competitive) conduct accompanies an observation of greater market-share stability (instability). Our analysis suggests that this inference may be invalid when colluding firms have private information. Indeed, when firms achieve first-best profits, a firm's future market share tends to be negatively correlated with its current market share.[4]

We consider next the possibility that firms are less patient. When the firms attempt to reward firm 2's honest report of high costs with favored treatment in future ties, a problem now arises: firm 1 may be unwilling to give up enough market share in the event of ties. More generally, if the disadvantaged firm's assigned market share is too low in a particular cost state, then it may undercut the collusive price and capture the entire market. When firms are less patient, therefore, productive efficiency today necessitates some inefficiency in the future. The firms, however, can choose the form that this inefficiency takes. For example, the collusive scheme may call for pricing inefficiency: the firms may lower prices when market-share favors are exchanged, in order to diminish the gain from undercutting. Or the scheme may require productive inefficiency: the disadvantaged firm may provide some of the transfer by giving up some market share in the state in which it is most efficient. Finally, in view of the these future inefficiencies, less patient firms may decide to implement less productive efficiency today (e.g., firm 2 may have positive market share today, even

[3] In related contexts, Athey, Bagwell and Sanchirico (1998) and Ayogai (1998) characterize particular asymmetric PPE, and Athey, Bagwell and Sanchirico (1998) characterize optimal symmetric PPE. The present paper, by contrast, characterizes optimal PPE. In the macroeconomics literature (e.g., Green (1987) or Atkeson and Lucas (1992)) on repeated games with private information, the game between a central planner and a continuum of agents is studied. A few papers (Wang (1994), Cole and Kocherlakota (1998)) consider small numbers of agents, but the focus is on existence or computational methods.

[4] An interesting case study is offered by McMillan (1985), who describes collusion among firms in the Japanese construction industry. Consistent with our formal analysis, McMillan reports that firms use future market share favors as a means of providing incentive for honest communication so that greater productive efficiency can be achieved. Future market share favors are also descriptive of other cartels, such as the citric acid cartel (*Business Week* (July 27, 1998)), in which any firm that sold beyond its budget in a given year purchased from "under-budget" firms in the following year.

when it alone has high costs), thus reducing the future transfer burden. Among these possibilities, we argue that pricing inefficiency is the least appealing. Our second general finding is the following: when firms are less patient, they give up productive efficiency (today or in the future) before lowering prices.

We next evaluate our two assumptions about the anti-trust environment. We begin with the role of communication. Our third general finding is that communication introduces potential benefits and costs to colluding firms. The benefit of communication is that it allows firms to smoothly divide the market on a state-contingent basis. Without communication, firms can only allocate market share with prices, and this decentralized approach limits significantly the range of market-sharing plans available. The cost of communication in our Bertrand model is subtle. Intuitively, when firms do not communicate, a given firm does not know its opponent's cost type when it chooses its price. Accordingly, if the opponent's price varies with cost, then the firm also does not know the exact price that its opponent will choose. This in turn diminishes the incentive that a firm has to undercut its prescribed price. Put differently, when firms communicate, the temptation to undercut may be exacerbated. Building off of this general cost-benefit tradeoff, we establish a number of specific results. We show that in the absence of communication, there again exists a discount factor strictly less than one above which first-best profits still can be achieved. For firms of moderate patience, however, restrictions on communication may diminish collusive profits. In addition, we show that firms may choose not to communicate in periods with significant market-share favors, as the absence of communication then serves to diminish the disadvantaged firm's incentive to undercut. More generally, we show that impatient firms may choose to avoid communication in some but not all periods.

To our knowledge, we are the first to identify benefits and costs from communication for colluding firms.[5] Communication offers no benefit in the standard (complete-information, perfect-monitoring) or public-monitoring (e.g. Green and Porter (1984)) collusion models. A potential benefit from communication is suggested in the emerging private-monitoring literature, wherein firms observe private and imperfect signals of past play. As Compte (1998) and Kandori and Matsushima (1998) explain, communication can then generate a public history on the basis of which subsequent collusion may be coordinated. But, as these authors acknowledge, they are unable to characterize optimal collusive conduct when communication is absent, and so they cannot determine when, or even whether, communication benefits colluding firms. In comparison, we assume that private information concerns current circumstances and past play is publicly observable. A public history is

[5] A role for communication also arises in the legalized-cartel and information-sharing literatures. In the basic information-sharing model (Shapiro (1986), Vives (1984)), firms can commit to share information before the play of a static oligopoly game. As Ziv (1993) shows, without this commitment, truth-telling incentives can be provided if firms exchange transfers (as in the legalized-cartel literature). Kuhn and Vives (1994) survey the competitive implications of communication.

thus generated whether firms communicate or not, and we may examine both cases.

Next, we consider anti-trust environments in which firms may entertain the exchange of bribes, though these must be self-enforcing and may incur inefficiencies, as through a risk of detection. Firms can potentially substitute current-period bribes for future market-share favors. In practice, bribes may be direct, with one firm paying other firms for the right to produce, or they may be associated with sophisticated and indirect processes.[6] Our fourth general finding is as follows: When detection by anti-trust officials is a concern, so that bribes are not fully efficient, bribes never fully replace future market-share favors as a means of transferring utility. Put differently, unless bribes are perfectly efficient, firms strictly prefer to keep track of history, using non-stationary equilibria that specify a future advantage to firms that admit high costs.

Our findings suggest that anti-trust policy can have perverse consequences. A recurring theme is that successfully colluding firms tolerate productive inefficiency before lowering prices. An antagonistic anti-trust policy, which limits firms' ability to communicate or exchange bribes, may thus limit productive efficiency without affecting prices. Such policies increase consumer welfare, though, if firms are sufficiently impatient that removing these instruments destroys their ability to collude at high prices. Overall, our findings provide some formal support for those (Bork (1965, 1966), Sproul (1993)) who are attentive to the efficiency gains that restraints of trade may afford.

2. The Model

We focus on a stylized model with two firms and two cost types, where firms 1 and 2 produce perfect substitutes and sell to a unit mass of customers with valuation r. Each firm has possible costs θ_L and θ_H and privately observes its realized costs prior to any pricing decisions. Thus, the state space in any period is denoted $\Omega = \{L, H\} \times \{L, H\}$, and we index these states as $(j, k) \in \Omega$, where the costs of firms 1 and 2 are given by $\theta^1 = \theta_j$ and $\theta^2 = \theta_k$ in state (j, k). The probability of the cost draw $j \in \{L, H\}$ in any period is denoted $\Pr(\theta^i = \theta_j) = \eta_j$, where $\eta_j > 0$ and $\eta_L + \eta_H = 1$; this is independent over time and across firms. To simplify the exposition of a few of the results, we assume $\eta_L > 1/2$.

The Nash equilibrium to the one-shot pricing game (without communication or transfer

[6] For example, an "over-budget" firm may compensate an "under-budget" firm, by purchasing the latter's output at the end of the budget period. Griffen (2000) reports that "compensation schemes" are common among international cartels (e.g., the lysine cartel). Or colluding suppliers may hold a "knockout" auction (among themselves) that determines the firm that is to win the procurement contract, and then rig the actual bids to ensure that this firm wins with a low bid (see, e.g., McAfee and McMillan (1992)). Finally, a firm that exceeds its production quota may contribute to a "common fund" while a firm that falls below its quota is permitted to withdraw from the fund. Common-fund arrangements appeared in the steel, aluminum and incandescent electric lamp cartels of the early 1900's (Stocking and Watkins, 1946). A similar arrangement was also found in the recent Garmet Box case (FTC Dockett 4777).

possibilities) is a symmetric mixed strategy equilibrium.[7] For each firm, the high type charges price equal to cost ($p = \theta_H$), while the low type mixes, receiving profit equal to $(\theta_H - \theta_L)\eta_H$, the expected profit from just undercutting the price charged by the high-cost type. Thus, ex ante expected profit to each firm in this equilibrium is equal to $\pi^{NE} \equiv (\theta_H - \theta_L)\eta_H\eta_L$. This payoff can be contrasted with the first-best level of profit to each firm, $\pi^{FB} \equiv \frac{1}{2}(r - E[\min(\theta^1, \theta^2)])$.

In our basic repeated-game model, firms can meet and communicate their types but cannot make side-payments. Formally, the firms play the following stage game in each period: (i) each firm i observes its type θ^i; (ii) each firm i makes an announcement $a^i \in A \equiv \{L, H, N\}$; (iii) each firm i then selects a price p^i and makes a market-share proposal q^i; (iv) for $\mathbf{p} \equiv (p^1, p^2)$ and $\mathbf{q} \equiv (q^1, q^2)$, market shares $m^i(\mathbf{p}, \mathbf{q})$ are allocated as follows: if $p^i > r$, then $m^i(\mathbf{p}, \mathbf{q}) = 0$; if $p^i < p^j \leq r$, then $m^i(\mathbf{p}, \mathbf{q}) = 1$; and if $p^i = p^j \leq r$, then $m^i(\mathbf{p}, \mathbf{q}) = \frac{1}{2}$ if $a_i = N$, $a_j = N$, or $q^i + q^j \neq 1$, while otherwise $m^i(\mathbf{p}, \mathbf{q}) = q^i$.

We interpret this stage game as describing an environment in which firms meet, make announcements concerning their respective cost types and then select prices and make market-share proposals. We allow each firm three possible announcements: a firm may announce that it has low (L) or high (H) costs, or it may choose to say nothing (N). We include the latter option, since, while we allow firms to meet and communicate, they are under no obligation to do so. Our formalization of market-share proposals permits firms to jointly determine their respective market shares when they set the same price. Since the market-share proposals follow the firms' announced cost positions, this formalization allows equally priced firms to allocate market share in a state-dependent fashion. We do not permit, however, both firms to produce positive quantities at different prices.[8] Beyond this restriction, the model grants firms considerable flexibility, and in principle they may mimic a centralized "mechanism" that gathers cost reports and determines prices and market shares. Our decentralized representation of interaction among firms, however, must incorporate further constraints that dissuade firms from deviations (e.g., undercutting the collusive price) that real-world firms might consider, but that would not be possible under the assumption that a mechanism sets prices.

We now define firm strategies for the stage game. Letting $\Omega^i \equiv \{L, H\}$, the space of policies from which a firm might choose is given by:

$$S^i = \{\alpha^i \mid \alpha^i : \Omega^i \to A\} \times \{\rho^i \mid \rho^i : \Omega^i \times A \to \Re\} \times \{\varphi^i \mid \varphi^i : \Omega^i \times A \to \Re\}.$$

A typical policy for firm i is denoted $s^i(\theta^i, a^j) = \{\alpha^i(\theta^i), \rho^i(\theta^i, a^j), \varphi^i(\theta^i, a^j)\}$, where the

[7] We consider pure strategy equilibria in the repeated game. This creates no tension, since we emphasize Pareto optimal equilibria, and in the characterizations we highlight these are pure.

[8] We thus rule out the possibility that the firms divide the market (e.g., geographically) and charge different prices in each segment. While the stage game is somewhat ad hoc, it does offer a simple framework (e.g., all transactions occur at the same price, so a rationing rule is not needed) within which to allow that firms may communicate and allocate market share in a state-contingent fashion.

first component is the announcement function, the second is the pricing function, the third is the market-share proposal function and a^j is firm j's realized announcement. Further, letting $\boldsymbol{\theta} \equiv (\theta^1, \theta^2)$ and $\mathbf{a} \equiv (a^1, a^2)$, we define the following vectors:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) \equiv (\alpha^1(\theta^1), \alpha^2(\theta^2)); \quad \boldsymbol{\rho}(\boldsymbol{\theta}, \mathbf{a}) \equiv (\rho^1(\theta^1, a^2), \rho^2(\theta^2, a^1))$$

$$\boldsymbol{\varphi}(\boldsymbol{\theta}, \mathbf{a}) \equiv (\varphi^1(\theta^1, a^2), \varphi^2(\theta^2, a^1)); \quad \mathbf{s}(\boldsymbol{\theta}) \equiv (s^1(\theta^1, a^2), s^2(\theta^2, a^1))$$

A policy vector $\mathbf{s}(\boldsymbol{\theta})$ determines announcements as well as the price and market-share proposal responses to these announcements. A policy vector thus determines a path through the stage game, and we may write stage-game payoffs conditional on a realization of cost types as $\pi^i(\mathbf{s}, \boldsymbol{\theta})$, with expected stage-game payoffs then given as $\bar{\pi}^i(\mathbf{s}) = E_\theta[\pi^i(\mathbf{s}, \boldsymbol{\theta})]$.

Consider now the repeated game. The firms meet each period to play the stage game described above, where each firm has the objective of maximizing its expected discounted stream of profit, given the common discount factor δ. Upon entering a period of play, a firm observes only the history of: (i) its own cost draws and policy functions, and (ii) realized announcements, prices and market-share proposals. Thus, a firm does not observe rival types or rival policy functions. Following Fudenberg, Levine and Maskin (1994), we restrict attention to those sequential equilibria in which firms condition only on the history of realized announcements, prices and market-share proposals and not on their own private history of types and policy schedules. Such strategies are called *public strategies* and such sequential equilibria are called *perfect public equilibria (PPE).*

Formally, let $h_t = \{\mathbf{a}_t, \mathbf{p}_t, \mathbf{q}_t\}$ be the public history of realized prices, announcements and market-share proposals up to date t. Let H_t be the set of potential histories at period t. A strategy for firm i in period t is denoted $\sigma_t^i : H_t \rightarrow S^i$. Let σ_t be a strategy profile in period t, and let $\boldsymbol{\sigma}$ represent a sequence of such strategy profiles, $t = 1, ..., \infty$. Then, given a history h_t, the expected per-period payoff in period t for firm i is $\bar{\pi}^i(\sigma_t(h_t))$. Each strategy induces a probability distribution over play in each period, resulting in an expected payoff for firm i, $v^i(\boldsymbol{\sigma}) = E[\sum_{t=1}^{\infty} \delta^{t-1} \bar{\pi}^i(\sigma_t(h_t))]$, where h_1 is the null history.

We assume that after every period firms can observe the realization of some public randomization device and select continuation equilibria on this basis. This is a common assumption in the literature, and it convexifies the set of equilibrium continuation values.[9] We do not introduce explicit notation for the randomization process.

Following Abreu, Pearce and Stacchetti (1986, 1990), we can now define an operator $T(V)$ which yields the set of PPE values, V^*, as the largest invariant, or "self-generating,"

[9] While we believe that this assumption is fairly innocuous, convexity of the set of continuation values plays an important role in parts of our analysis. In Section 4.2, we discuss conditions under which convexity obtains without resorting to randomization.

set. Letting $S \equiv S^1 \times S^2$ and $\boldsymbol{\upsilon} \equiv (\upsilon^1, \upsilon^2)$, the operator is defined as follows:

$$\begin{gathered} T(V) = \{(u^1, u^2) :\ \exists\ \mathbf{s} \in S \text{ and } \boldsymbol{\upsilon} : A^2 \times \Re^4 \to co(V) \text{ such that:} \\ \text{for } i = 1, 2,\ u^i = \bar{\pi}^i(\mathbf{s}) + \delta E \upsilon^i(\mathbf{s}(\boldsymbol{\theta})); \\ \text{and, for each } i \text{ and } \tilde{s}^i = (\tilde{\alpha}^i, \tilde{\rho}^i, \tilde{\varphi}^i) \in S^i, \\ u^i \geq \bar{\pi}^i(\tilde{s}^i, s^j) + \delta E \upsilon^i[\tilde{s}^i(\theta^i, \alpha^j(\theta^j)), s^j(\theta^j, \tilde{\alpha}^i(\theta^i))]\} \end{gathered}$$

This operator effectively decomposes equilibrium play into two components: current-period strategies $\mathbf{s} \in S$ and continuation values $\boldsymbol{\upsilon}$ drawn from the convex hull of the set V. We record immediately a useful property of T:

Lemma 1. *T maps compact sets to compact sets..*

Proof. All of the constraints entail weak inequalities; the feasible set is compact; and utility and constraint functions are real-valued, continuous and bounded. ■

This property of T is the critical one for applying the methodology of Abreu, Pearce, and Stacchetti (1990). In particular, let V_0 be compact and contain all feasible, individually rational payoffs (e.g., $V_0 = [0, r/(1-\delta)] \times [0, r/(1-\delta)]$), and define $V_{n+1} = T(V_n)$, $n \geq 0$. Then the definition of T implies that $T(V_n) \subseteq V_n$. Using this and the fact that V_n is nonempty for each n (since $\pi^{NE}/(1-\delta)$ is in every V_n), $V^* = \lim_{n\to\infty} V_n$ is a nonempty, compact set. Following the arguments in Abreu, Pearce, and Stacchetti (1990), V^* is the largest invariant set of T, and thus it is equal to the set of equilibrium values of this game.

To present our findings, we distinguish between two kinds of equilibria. In an *informative PPE*, firms employ equilibrium strategies in which they always share their cost information with one another: for all $i \in \{1, 2\}$ and $j \in \{L, H\}$, if $\theta^i = \theta_j$, then $\alpha^i(\theta^i) = j$. By contrast, in an *uninformative PPE*, firms are unwilling (or unable) to communicate, and we capture this by focusing upon equilibria in which firms never share cost information: for all $i \in \{1, 2\}$ and $j \in \{L, H\}$, $\alpha^i(\theta_j) \equiv N$. We use the operators $T^I(V)$ and $T^U(V)$, respectively, to capture these additional restrictions on $\mathbf{s}$, where both operators are extended to include also the repeated-Nash payoffs, $\mathbf{u}^{NE} \equiv (\pi^{NE}/(1-\delta), \pi^{NE}/(1-\delta))$, which derive from mixed strategies and may be used as an off-equilibrium-path punishment. Informative and uninformative PPE are of independent interest, and the juxtaposition of these two classes of PPE highlights the benefits and costs of informative communication for colluding firms. The characterization of such equilibria also contributes to our understanding of the full PPE set, V^*, since optimal equilibria of the unrestricted PPE class may involve informative communication following some histories and not others.

3. The Mechanism Design Approach

In this section, we consider the class of informative PPE and show that the search for the optimal informative PPE can be recast in terms of a static mechanism design program. In addition, we establish the solution to this program in two benchmark cases.

3.1. Mechanism Notation and Incentive Constraints

The set of informative PPE values, V^I, is the largest invariant set of the operator T^I; therefore, every utility vector $\mathbf{u} \in V^I$ can be generated by associated current-period strategies and continuation-value functions, $\mathbf{s}$ and $\boldsymbol{\upsilon}$. When following these strategies, firms report their cost types truthfully and receive the corresponding prices and market-share allocations. Our approach in this section is to introduce notation for such state-contingent prices, market-share allocations and continuation values, and then formalize the corresponding incentive constraints that these must satisfy to be implementable as equilibrium play.[10]

We begin with a general description of the incentive constraints. In an equilibrium of the repeated game, there are two kinds of deviations. First, a firm with cost type θ^i may adopt the policy that the equilibrium specifies when its cost type is instead $\theta^{i'} \neq \theta^i$. Such an "on-schedule" deviation is not observable, as a deviation, to the rival firm. The equilibrium prices, market shares and continuation values therefore must be incentive compatible. Second, a firm also must not have the incentive to choose a price and market share that is not assigned to any cost type. Such an "off-schedule" deviation is observable to the rival firm as a deviation, and a sufficiently patient firm is deterred from a deviation of this kind if the collusive scheme then calls for a harsh "off-the-equilibrium-path" punishment. The on-schedule incentive constraints are reminiscent of truth-telling constraints in standard mechanism design theory, with continuation values playing the role of transfers. The off-schedule constraints are analogous to type-dependent participation constraints.

To make these analogies precise, we first define state-contingent prices, market shares and continuation values. In state (j,k), firm i serves q^i_{jk} customers at price p_{jk}.[11] The continuation value assigned to firm i in state (j,k) is denoted v^i_{jk}. Let $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{v}$ denote the associated vectors, and let $\mathbf{z}$=($\mathbf{p}$,$\mathbf{q}$,$\mathbf{v}$) be the "policy vector." Finally, we use $\mathcal{Z}(V)$ to denote the set of such vectors that are feasible and consistent with the extensive form game when continuation values are drawn from the set V :

$$\mathcal{Z}(V) = \{\mathbf{z} = (\mathbf{p}, \mathbf{q}, \mathbf{v}) : \text{For all } i = 1, 2,\ (j,k) \in \Omega,\ (v^1_{jk}, v^2_{jk}) \in co(V), \\ p_{jk} \leq r;\ q^i_{jk} \in [0,1] \text{ and } q^1_{jk} + q^2_{jk} = 1\}.$$

[10] The logic of our approach is analogous to the revelation principle. In our model, however, communication is an explicit part of the extensive-form game, unlike most standard applications of the revelation principle (for single-stage games), where the idea that firms "report" their costs to a mechanism is an abstraction. We represent each player's incentive constraints in direct form, where the incentive constraints protect against deviations at each stage (the communication and pricing stages). See Myerson (1986) for more discussion of the revelation principle in multi-stage communication games.

[11] Our Bertrand model ensures that in any state (j,k) a single transaction price p_{jk} prevails. Firm i therefore sets this price if it makes positive sales (i.e., if $q^i_{jk} > 0$). If firm i makes no sales in state (j,k), then firm i's price may differ from p_{jk}, but it cannot be lower.

Next, we denote expected market shares and continuation values for each firm, given a cost realization, by

$$\bar{q}_j^1 = \sum_{k \in \{L,H\}} \eta_k \cdot q_{jk}^1; \ \bar{v}_j^1 = \sum_{k \in \{L,H\}} \eta_k \cdot v_{jk}^1;$$

and likewise for firm 2. Consider now each firm's interim current-period payoff as a function of its announcement, assuming that the opponent announces truthfully and both firms adhere to the schedule. When firm 1 announces cost type $\hat{\jmath}$ when the true cost type is j, interim current-period profits are given by:

$$\Pi^1(\hat{\jmath}, j; \mathbf{z}) = \sum_{k \in \{L,H\}} \eta_k \cdot q_{\hat{\jmath}k}^1 \cdot (p_{\hat{\jmath}k} - \theta_j).$$

Adding on continuation values, we write interim and ex ante utilities as

$$U^1(\hat{\jmath}, j; \mathbf{z}) = \Pi^1(\hat{\jmath}, j; \mathbf{z}) + \delta \bar{v}_{\hat{\jmath}}^1; \ \bar{U}^1(\mathbf{z}) = \sum_{j \in \{L,H\}} \eta_j \cdot U^1(j, j; \mathbf{z}).$$

These functions are defined analogously for firm 2.

Using this notation, the on-schedule incentive constraints can be easily related. We distinguish "upward" from "downward" incentive constraints, since typically only the downward constraints are binding:

$$U^i(H, H; \mathbf{z}) \geq U^i(L, H; \mathbf{z}) \quad \text{(IC-On}i_D\text{)}$$

$$U^i(L, L; \mathbf{z}) \geq U^i(H, L; \mathbf{z}) \quad \text{(IC-On}i_U\text{)}$$

Our next task is to represent the off-schedule incentive constraints. In an informative PPE, there are two kinds of off-schedule constraints. The first concerns the incentive of a firm to deviate from the assigned price after communication takes place. If both firms are assigned a price less than firm 1's cost, firm 1 might like to price slightly above firm 2, to avoid producing in that state; alternatively, at higher prices, firm 1 might wish to slightly undercut firm 2's price and capture the entire market.[12] If the following constraint is satisfied, neither of these deviations is profitable:

$$\delta(v_{jk}^1 - \underline{v}^1) \geq \max(q_{jk}^2(p_{jk} - \theta_j), q_{jk}^1(\theta_j - p_{jk})) \quad \text{(IC-Off1}_{jk}^I\text{)}$$

where $\underline{v}^i = \underline{v}^i(V) \equiv \inf\{v^i : v \in V\}$.[13] As $\underline{v}^i$ is reached only off of the equilibrium path, we can essentially treat it as a parameter in the analysis. IC-Off2$_{jk}^I$ is defined analogously.

[12] Given that unit costs are constant in output, a firm best deviates by claiming all market share or relinquishing all market share. In either event, a small change in price serves the purpose. We therefore need not concern ourselves with the possibility that a firm deviates by maintaining the price and adjusting up or down its proposed market share.

[13] We write $\underline{v}^i$ rather than $\underline{v}^i(V)$ to conserve notation, and we take the off-schedule constraints relative to the set of values under consideration in a particular context.

The second kind of off-schedule deviation is an *interim* deviation. Suppose that the collusive scheme assigns a lower price in state (L, L) or (L, H) than in (H, H) or (H, L). If firm 1 draws a low cost, firm 1 might be tempted to report a high cost, in order to induce firm 2 to price high, so that firm 1 might then undercut firm 2's high price. Firm 1 might wish to learn the realization of firm 2's type before making a final decision to undercut. Deviations of this kind are dissuaded if:

$$U^1(L, L; \mathbf{z}) \geq \sum_{k \in \{L,H\}} \eta_k \cdot \max(q^2_{Hk}(p_{Hk} - \theta_L) + \delta \underline{v}^1, q^1_{Hk}(p_{Hk} - \theta_L) + \delta v^1_{Hk}) \tag{IC-Off-M1}$$

where the M is mnemonic for "misrepresentation." The constraint for firm 2 is defined analogously. Since a firm gains most from a market-share increase when its costs are low, it can be verified that if the other on- and off-schedule incentive constraints are satsified, then the high type never has the incentive to engage in this type of misrepresentation. Further, if prices are the same in each state (as in many of our characterizations below), then the other off-schedule constraints render IC-Off-Mi redundant.

3.2. The Repeated Game as a Mechanism

We introduce notation for the feasible set of policy vectors when firms use informative communication, given an arbitrary set of continuation values V :

$$\mathcal{F}^I(V) = \{\mathbf{z} = (\mathbf{p}, \mathbf{q}, \mathbf{v}) \in \mathcal{Z}(V) : \text{ For all } i = 1, 2, \text{ IC-On}i_D, \text{ IC-On}i_U, \\ \text{IC-Off}i^I_{jk} \text{ and IC-Off-M}i \text{ hold}\}.$$

With this notation in place, we present the following lemma.

Lemma 2. *Given a set $V \subset \Re^2$, let*

$$\tilde{T}^I(V) = \{(u^1, u^2) : \exists \mathbf{z} = (\mathbf{p}, \mathbf{q}, \mathbf{v}) \in \mathcal{F}^I(V) \text{ such that for } i = 1, 2,\ u^i = \bar{U}^i(\mathbf{z})\} \cup \mathbf{u}^{NE}.$$

Then $\tilde{T}^I(V) = T^I(V)$.

The lemma follows by a comparison of constraints (see Athey and Bagwell (1999) for details). For the class of informative PPE, Lemma 2 formalizes the relationship between the repeated game and the mechanism design problem we have just defined. It states that we can characterize the operator T^I as generating the set of all utilities that satisfy the constraints of a fairly standard mechanism design problem, with the addition of the unusual restriction $(v^1_{jk}, v^2_{jk}) \in V$. An important consequence of this result is that for any informative PPE utility vector $\mathbf{u}$, there exists a policy vector $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ that "implements" $\mathbf{u}$, in the sense that it satisfies the conditions in the definition of $\tilde{T}^I(V)$.

3.3. Benchmark Cases

In this section, we characterize the Pareto frontier of $\tilde{T}^I(V)$ for two examples of sets V. These examples are motivated by the static mechanism design literature where V is the set of available monetary transfers. In the first example, V is a line of slope -1; this represents "budget-balanced" transfers of utility that incur no efficiency loss. In the second example, we consider sets of the form $V = \{(v^1, v^2) : v^1, v^2 \leq K\}$; for such sets, all continuation values except (K, K) are Pareto inefficient. The cases are illustrated in Figure 1. These benchmarks allow us to develop some basic intuition, on which we build when we later consider sets V with more general shapes, such as the convex set illustrated in Figure 2.

To draw most clearly the analogy to the static mechanism design literature, we ignore the off-schedule incentive constraints in this section. We then refer to the set of constraints excluding off-schedule incentive constraints as $\mathcal{F}^I_{On}(V)$, and we define:

$$\tilde{T}^I_{On}(V) = \{(u^1, u^2) : \ \exists \mathbf{z} = (\mathbf{p}, \mathbf{q}, \mathbf{v}) \in \mathcal{F}^I_{On}(V) \text{ such that for } i = 1, 2, \ u^i = \bar{U}^i(\mathbf{z})\}.$$

In discussing schemes, we say that a scheme uses *productive efficiency* if for every state $(j, k) \in \Omega$, $q^1_{LH} = q^2_{HL} = 1$. We say that a scheme uses *efficient pricing* if $p_{jk} = r$ for all $(j, k) \in \Omega$. Similarly, the scheme is characterized by *Pareto efficient continuation values* if for every (j, k), there does not exist a continuation value pair $(\tilde{v}^1, \tilde{v}^2) \in V$ that Pareto dominates (v^1_{jk}, v^2_{jk}).

To begin, we record the following standard lemma:

Lemma 3. *Any* $\mathbf{z}$ *satisfying IC-On*i_D *and IC-On*i_U *also satisfies* $\bar{q}^i_H \leq \bar{q}^i_L$*. If IC-On*$i_D$ *binds, then*

$$U^i(H, H; \mathbf{z}) = U^i(L, H; \mathbf{z}) = U^i(L, L; \mathbf{z}) - \bar{q}^i_L(\theta_H - \theta_L). \quad (3.1)$$

Market-share monotonicity follows since our model satisfies a single-crossing property: the low-cost type has a higher marginal return to market share. The representation of the relationship between the interim utilities follows directly and says that the low-cost type earns an "efficiency rent" of $\bar{q}^i_L(\theta_H - \theta_L)$ over the high-cost type.

By Lemma 3, when IC-Oni_D binds for each firm, the ex ante utility for firm i is

$$\bar{U}^i(\mathbf{z}) = U^i(H, H; \mathbf{z}) + \eta_L \bar{q}^i_L(\theta_H - \theta_L) = \Pi^i(H, H; \mathbf{z}) + \delta \bar{v}^i_H + \eta_L \bar{q}^i_L(\theta_H - \theta_L). \quad (3.2)$$

Among the set of allocation rules where IC-Oni_D binds, firm i is indifferent between providing incentives with low prices or low continuation values for its low-cost type. Intuitively, in contrast to market share, neither the price nor the continuation value interacts directly with the firm's type in the firm's objective function; thus, the cartel has a preference over low-cost prices and continuation values for which a firm's on-schedule incentive constraint binds, only insofar as these instruments generate utility losses or gains for the other firm.

Lowering price decreases the utility of the other firm. In contrast, when cross-firm transfers of utility are available, lowering one firm's continuation value may allow an increase in that of the other firm. Continuation values are then a superior instrument.

To better highlight some of these themes, we turn now to two special cases. First, we suppose that the set of feasible continuation values is a line of slope -1.[14]

Lemma 4. *For $K \in \Re$, suppose that $V(K) = \{(v^1, v^2) : v^1 + v^2 = 2K\}$. Then, for any $K \geq 0$, the Pareto frontier of $\tilde{T}^I_{On}(V(K))$ is $\{(u^1, u^2) : u^1 + u^2 = 2\pi^{FB} + \delta 2K\}$, and this frontier can be implemented with a policy vector $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ that satisfies the following properties: productive efficiency, pricing efficiency and Pareto efficient continuation values ($v^1_{jk} + v^2_{jk} = 2K$ for all j, k); $v^1_{HL} - v^1_{LH} = (r - \theta^H)/\delta$; IC-Oni$_D$ binds for each i; and $v^1_{LH} < v^1_{jj} < v^1_{HL}$ for $j \in \{L, H\}$.*

As expected, first-best is attained. The downward on-schedule constraints bind, since it is the low-cost type who has the higher market share, and market share is desirable for both firms. Thus, the relevant consideration is to dissuade the high-cost type from mimicking the low-cost type; as lower cost types have a higher marginal benefit to high market share, if the high-cost type is just indifferent between the high and low announcement, the low-cost type strictly prefers the low-cost announcement. The optimal mechanism requires transfers through continuation values that reward a firm for announcing high costs.

Next, we consider a second special case, wherein the firms receive continuation values from a rectangular set in which each firm receives at most K. The continuation-value Pareto frontier is then a single point, and efficient continuation-value transfers across firms are thus unavailable. To state the result, we refer to the following condition:

$$(r - \theta_H)/(\theta_H - \theta_L) > \eta_H \tag{3.3}$$

Lemma 5. *Suppose that $V(K) = \{(v^1, v^2) : v^1, v^2 \leq K\}$. (i) Suppose that (3.3) holds. Then, for any K, the Pareto frontier of $\tilde{T}^I_{On}(V(K))$ is*

$$\{(u^1, u^2) : \ u^1 + u^2 = r - E[\theta] + \delta 2K, \ u^i \geq 0\},$$

and this frontier can be implemented with a policy vector $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ that satisfies the following properties: pricing efficiency, Pareto efficient continuation values ($v^i_{jk} = K$ for all i, j, k) and productive inefficiency with $\bar{q}^i_H = \bar{q}^i_L$ for $i = 1, 2$. (ii) Suppose that (3.3) fails. Then the Pareto frontier of $\tilde{T}^I_{On}(V(K))$ is given by

$$\{(u^1, u^2) : \ u^1 + u^2 = \eta_H(r - \theta_H) + \eta_L(1 + \eta_H)(\theta_H - \theta_L) + \delta 2K, \ u^i \geq 0\}.$$

[14] For the public goods problem with private information, d'Aspremont and Gerard-Varet (1979) show that the first-best can be attained using budget-balanced transfers, when the mechanism must satisfy incentive compatibility but not participation constraints. In their analysis of "strong" bidding cartels, where side-payments are allowed, McAfee and McMillan (1992) specialize this result for the case of first-price auctions, showing that participation constraints can be satisfied. The following result is a two-type specialization, rephrased to allow for continuation values that sum to a constant other than zero.

This can be implemented with a policy vector $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ *that satisfies the following properties: productive efficiency, Pareto efficient continuation values* ($v^i_{jk} = K$ *for all* i, j, k), *pricing efficiency in state* (H, H) ($p_{HH} = r$) *and a price of* $\frac{\eta_H}{1+\eta_H}(r - \theta_H) + \theta_H$ *in other states.*

Lemma 5 refers to an environment in which the only instruments available (reduced continuation values, low prices) with which to achieve productive efficiency are wasteful. When (3.3) holds, so that the profit to the high-cost type is large relative to the efficiency advantage of the low-cost type, Lemma 5 establishes that the Pareto frontier entails productive inefficiency: the loss in profit from either Pareto inefficient continuation values or inefficient pricing overwhelms any potential productive efficiency gain.

To see the role of (3.3), consider raising productive efficiency by increasing q^2_{HL} (and therefore decreasing q^1_{HL}). The subtle aspect of the intuition entails understanding the effects of this change when prices and continuation values must adjust to maintain the on-schedule constraints. The change decreases firm 1's ex ante utility by $\eta_L(r - \theta_H)$, since firm 1's high type bears the cost directly and firm 1's low type must now charge a lower price or receive a lower continuation value to avoid violating IC-On1$_D$. The change increases firm 2's ex ante utility by $\eta_L\eta_H(\theta_H - \theta_L)$, the higher "efficiency rent" $(\theta_H - \theta_L)$ available to firm 2's low-cost type in state (H, L). Then (3.3) guarantees that the cost to firm 1, incurred across both states (H, L) and (L, L), outweighs the efficiency benefit to firm 2 in state (H, L). This result introduces a theme that will recur throughout our analysis. There is a "tax" on productive efficiency: improving productive efficiency tightens on-schedule constraints, leading to further distortions. If instead (3.3) fails, with a rectangular continuation value set, it is always possible to achieve the optimal collusive payoffs using the highest available continuation values and low prices for the low-cost types.[15]

Whether firms choose to produce efficiently or not, cartel profit is not improved by moving from $V = \{(K, K)\}$ to $V = \{(v^1, v^2) : v^1, v^2 \leq K\}$. Wasteful continuation values are not useful for providing incentives. With this observation, Lemma 5 may be related to other findings for continuum-type models. In their analysis of "weak" bidding cartels, McAfee and McMillan (1992) show that when transfers are prohibited ($V = \{(0, 0)\}$) and the distribution over types, $F(\theta)$, is log-concave, the optimal cartel uses identical bidding at the seller's reservation value. This is the bidding cartel analog of pricing efficiency and productive inefficiency. Athey, Bagwell and Sanchirico (1998) consider collusion among sellers where $V = \{(v^1, v^2) : v^1 = v^2\}$. In a repeated game, this corresponds to Symmetric PPE. They find that wasteful continuation values ("price wars") are not used, while pricing efficiency and productive inefficiency obtain when $F(\theta)$ is log-concave.[16]

[15] Notice that the pricing scheme outlined in the lemma can be implemented decentrally: each firm charges a price of r when its own cost is high, and selects a price $\hat{p}$ when its own cost is low. This allocates market share efficiently and achieves the price of $\hat{p}$ in all states except (H, H).

[16] The continuum- and two-type models may be further related using an N-type model. Let η_n be the probability of cost type n. Then, the following conditions replace (3.3): $(r - \theta_N)\eta_m - \eta_N(\theta_{m+1} -$

4. Characterization of Informative PPE

We next characterize the set of informative PPE values. Our analysis builds on the insights developed in the benchmark cases of Section 3.3. Throughout, we develop analytically some key findings, and we then illustrate additional subtleties with computational examples.

Before beginning the formal analysis, we outline the central tradeoffs. Suppose the firms attempt to implement first-best profits. In the first period of the game, a first-best scheme must implement productive efficiency and pricing efficiency; thus, from the perspective of current-period profits, high-cost firms are tempted to mis-report their costs in order to achieve greater market share. To ensure truthful reporting, the agreement therefore must provide that firm 2 receives future market-share favors from firm 1 following a realization of the state (L, H). Suppose then that (L, H) is realized in the first period, and consider the scheme in the second period. In a first-best collusive scheme, productive efficiency is again required; consequently, if state (L, H) is once more realized, then firm 2 must again receive zero market share. On the other hand, if the firms experience the same costs in the second period, then the collusive arrangement may favor firm 2 while simultaneously delivering first-best profits. This is achieved by giving firm 2 more than 1/2 of the market in the second period when the (L, L) and (H, H) states are realized. If these market shares are appropriately chosen, both firms still have the incentive to report truthfully. What might prevent such a scheme from succeeding? The firms must be sufficiently patient so that firm 1 is dissuaded from undertaking an off-schedule deviation following a realization of (L, L), when its assigned market share is low. What if this cannot be accomplished? Then, asymmetric treatment introduces new inefficiencies. In particular, the scheme may require low prices, or it may call upon firm 1 to relinquish some market share in period 2 in the (L, H) state, even though it is most efficient, as its temptation to undertake an off-schedule deviation is low when its assigned market share is high.

Pulling these themes together, we may summarize the central tradeoffs as follows. If in a given period, the firms seek productive efficiency "today," then asymmetric treatment is required "tomorrow." Productive and pricing efficiency tomorrow, however, can then be maintained only if the asymmetric treatment is implemented through asymmetric market-share assignments among equally efficient firms pricing at the reservation value. In turn, this is possible only if tomorrow the disadvantaged firm is sufficiently patient to endure its assigned low market share; if not, some inefficiency may be required. In view of these tradeoffs, a cartel comprised of moderately patient firms may assign market shares today without achieving full productive efficiency, in order to lessen the future transfer burden and thus reduce future inefficiency.

$\theta_m) \sum_{n=1}^{m} \eta_n > 0$ for all $m < N$; and $(\theta_{m+1} - \theta_m) \sum_{n=1}^{m} \eta_n / \eta_m$ is nondecreasing in m. The first expression is the analog of (3.3); the second condition is the analog of log-concavity of $F(\theta)$. If $r > \theta_N$, the first expression is satisfied in the continuum-type limit, when the η_n's go to zero at a common order.

In the next two subsections, we derive conditions on the discount factor under which firms are able to implement a given level of efficiency (such as first-best) in every period of the game. Subsequently, we explore in greater depth the optimal resolution of the tradeoffs between current and future efficiency faced by firms of moderate patience.

4.1. A Linear Informative PPE Set With First-Best Profits

In this subsection, we identify a discount factor strictly less than one above which the cartel can achieve first-best profits in every period. Recall that Section 3.3 analyzes Pareto optimal schemes for an exogenous set of continuation values. We now confront the endogenous nature of the continuation-value set. Our goal is to establish the existence of a set of Informative PPE values, where (i). each utility pair yields first-best profits to the cartel, and (ii). when implementing any point in the set, only other elements of the set are used as continuation values on the equilibrium path. A "self-generating" set of values supporting first-best profits must be a line segment with slope -1, together with the "punishment" value(s) that serve as threats to deter off-schedule deviations.

We attempt to construct such a line segment of equilibrium values, where the endpoints are denoted (x, y) and (y, x). We focus on finding a policy vector $\mathbf{z}$ that implements the endpoint (x, y) using pricing and productive efficiency and continuation values taken only from the line segment $[(x, y),(y, x)]$, while satsifying all feasibility and incentive constraints. If this can be accomplished, then there exists a $\mathbf{z}'$ that exchanges the roles of the two players and implements (y, x). Any convex combination of (x, y) and (y, x) can be attained using a convex combination of $\mathbf{z}$ and $\mathbf{z}'$.

We proceed in two steps. First, we consider the implementation of the endpoint (x, y) when off-schedule constraints are ignored. This step can be challenging. If monopoly profit for a high-cost firm, $r - \theta_H$, is too large, it may be difficult to achieve the desired level of profit for firm 1, $\bar{U}^1(\mathbf{z}) = x$, while maintaining $v^1_{jk} \geq x$. Intuitively, firm 1's average profit today then must be worse than its per-period profits derived from each of its continuation values: $E[\Pi^1(j, j; \mathbf{z})] \leq v^1_{jk}(1 - \delta)$ for each (j, k). Further, firm 1 has incentive to reveal a high-cost type only if the future looks *relatively* better after a realization of (H, L): following the logic of Lemma 4, the on-schedule constraints can be satisfied only if $v^1_{HL} - v^1_{LH} \geq (r - \theta_H)/\delta$. This requirement places additional downward pressure on today's expected profit. But productive and pricing efficiency impose a lower bound on today's profit. Similarly, if the efficiency-rent term, $\theta_H - \theta_L$ is too small, it can be difficult to implement $\bar{U}^2(\mathbf{z}) = y$ while maintaining $v^2_{jk} \leq y$. Intuitively, firm 2's average profit today then must be greater that its per-period profits derived from each of its continuation values. Recalling (3.2), this is more easily achieved when the efficiency rent $\theta_H - \theta_L$ is large.

This discussion suggests a restriction under which $\kappa \equiv (r - \theta_H)/(\theta_H - \theta_L)$ is not too large. Recalling our assumption $\eta_L > 1/2$, we consider the following restriction:[17]

$$\eta_L^2 > \kappa(2\,\eta_L - 1). \tag{4.1}$$

We may verify that (4.1) is satisfied if $\kappa < 1$; more generally, it holds if η_L is sufficiently close to 1/2.[18] Under (4.1), and in the absence of off-schedule constraints, we show in the Appendix that the implementation of (x, y) is feasible if δ exceeds a critical value, δ^{Fon}, which is less than unity and defined as follows:

$$\delta^{Fon} = \frac{\kappa}{\eta_L^2 + 2\kappa(1 - \eta_L)}.$$

The second step is to assume (4.1) and consider restrictions implied by the off-schedule constraints. Of course, if firms are sufficiently patient, then an off-schedule deviation is unattractive. But the associated critical discount factor is difficult to compute, since the exact value depends on the worst punishment available. Fortunately, our qualitative results do not depend on a closed-form calculation. Instead, we proceed as follows.

First, for any given δ, we let $\underline{v}_1(\delta)$ denote the worst equilibrium value for firm 1. From the folk theorem of Fudenberg, Levine, and Maskin (1994), we know that $\underline{v}_1(\delta)$ approaches 0 as δ approaches 1; furthermore, since the repeated play of the static Nash equilibrium is a feasible punishment, we also know that $\underline{v}_1(\delta) \leq \pi^{NE}/(1-\delta)$. We thus may define $\lambda^I(\delta) \in [0, 1]$ by $\underline{v}_1(\delta) \equiv \lambda^I(\delta)\pi^{NE}/(1-\delta)$, so that $\lambda^I(\delta)$ gives the fraction of the static Nash profits that can be sustained, on average, in the worst equilibrium for firm 1. The function $\lambda^I(\delta)$ is nonincreasing and satsifies $\lambda^I(0) = 1 > 0 = \lambda^I(1)$. Second, for any given λ and associated punishment value $\lambda\pi^{NE}/(1-\delta)$, we solve for the critical discount factor for supporting first-best profits, denoted $d^F(\lambda)$. The function $d^F(\lambda)$ is nondecreasing, where $d^F(0) < 1$ and $d^F(1)$ are the critical discount factors for implementing first-best profits when the punishment entails zero and repeated-Nash profits, respectively. The critical discount factor is thus determined as the fixed point of the equation $\delta = d^F(\lambda^I(\delta))$, and it must lie in $(d^F(0), d^F(1))$.

Consider now the derivation of $d^F(\lambda)$. We seek the smallest δ such that the values (x, y) can be sustained as an equilibrium, using only values on $[(x, y), (y, x)]$ on the equilibrium path and $\lambda\pi^{NE}/(1-\delta)$ as the off-schedule punishment. The program is formalized in the Appendix. In describing its solution, a subtlety arises: for different parameter values,

[17] Our assumption that $\eta_L > \frac{1}{2}$ determines which continuation value, v^1_{HH} or v^1_{LL}, is lower and thus more likely to drop below x when we try to implement $\bar{U}^1 = x$ with $v^1_{LH} = x$ and $v^1_{HL} \leq y$. For $\eta_L < \frac{1}{2}$, it can be shown that a different but analogous condition must hold.

[18] If (4.1) fails, the firms may not be able to implement exactly first-best profits. But it is possible to construct self-generating sets composed of three connected line segments, where the interior line segment has slope -1, and all points on that segment are implemented using productive efficiency. In contrast, some productive inefficiency is used on the exterior segments. As firms become more patient, the width of the interior line segment grows, and so first-best can be approximated as δ approaches 1.

different constraints bind, and so the formula for $d^F(\lambda)$ changes. Rather than enumerating all possible cases, we derive an upper bound for $d^F(\lambda)$ that applies for all parameter values. As we discuss further in the Appendix, to construct this upper bound, we impose that IC-Off1$_{LL}$ is binding, and we set the punishment at its "softest" level with repeated-Nash play (i.e., $\lambda = 1$). With this, we may report a (conservative) upper bound for the critical discount factor that suffices for an informative PPE that achieves first-best profits:

$$\delta^{FB} = \max\left(\delta^{Fon}, \frac{\eta_L + \kappa(1-\eta_L)}{\eta_L + \kappa(1-\eta_L) + \eta_L^2\kappa}\right).$$

Observe that $\delta^{FB} < 1$ when (4.1) is satisfied.

Proposition 1. *Assume (4.1). Then, for all $\delta \in (\delta^{FB}, 1]$, there exist values $y > x > 0$ such that $x + y = 2\pi^{FB}/(1-\delta)$, and the set $[(x,y),(y,x)] \cup \mathbf{u}^{NE}$ is a self-generating set of informative PPE values.*

Proposition 1 can be thought of as a generalization of Fudenberg, Levine and Maskin's (1994) folk theorem. Instead of resorting to taking the limit as $\delta \to 1$, we compute a discount factor strictly less than one where first-best is achieved. Our result further provides an explicit characterization of the behavior associated with this first-best arrangement. The following specific example illustrates how this is accomplished.

4.1.1. Example: Achieving First-Best Collusion

To understand how first-best collusion unfolds over time, consider a particular example, where $r = 2.5$, $\theta_H = 2$, $\theta_L = 1$ and $\eta_L = .6$, so that $\pi^{FB} = .67$.

Consider first the critical discount factors. For these parameter values, we find that $\delta^{Fon} = .66$ and $\delta^{FB} = .816$. As described above, these bounds are in general conservative. Given specific parameter values, however, the program defined in the Appendix for $d^F(\lambda)$ can be readily solved. In the present example, for all $\lambda \in [0,1]$, $d^F(\lambda)$ is achieved using a policy vector whereby the following constraints bind: $p_{jk} = r$ for all $(j,k) \in \Omega$, $q^1_{LH} = 1$, $q^1_{HL} = 0$, $q^1_{HH} = 0$, $v^1_{HH} = x$, $v^1_{LH} = x$, $v^1_{HL} = y$, IC-On1$_D$, IC-On2$_D$, and IC-Off1$_{LL}$. We find that $d^F(\lambda) = 12.5(\sqrt{1087 - 216\lambda} - 3)/(108\lambda - 539)$, which yields $d^F(1) = .769$ and $d^F(0) = .695$. That is, when the firms use repeated-Nash play as the off-schedule punishment, first-best profits can be sustained if and only if $\delta \geq .769$.

Now consider the collusive strategies that support these payoffs. We take $\delta = .769$ and $\lambda = 1$, so that the equilibrium we describe is sure to exist; for lower levels of λ, the qualitative description of play is similar. In implementing a first-best equilibrium, the history of past play can always be summarized by one of five states, numbered 1 to 5, where state 1 is best for firm 1 and state 5 is best for firm 2. Figure 3a summarizes the policy vectors that implement each state (recalling that prices always equal r, and letting s_{jk} denote the state reached in continuation play following a realization of (j,k)).

After the null history, play begins in state 3. In that state, firms are treated symmetrically. The firms implement productive efficiency and share the market otherwise. Following a realization of (L, H), the firms proceed to state 5, while following a realization of (H, L) they proceed to state 1. Otherwise, they return to state 3.

Suppose now that the cost types are (L, H) in the first period. The firms proceed to state 5, where payoffs are asymmetric but productive efficiency is still achieved. The asymmetries are most pronounced in state (H, H): $q^1_{HH} = 0$, and if (H, H) is realized, the firms return to state 5 in the next period. The constraint IC-Off1$_{LL}$ binds, and so to mitigate the incentive to cheat, $q^1_{LL} = .152$; and after the realization of state (L, L), the firms proceed to a better state for firm 1, state 4. Firm 1 is induced to admit when it draws a high cost, by the prospect of a future reward: if the cost realizations are (H, L), firm one receives no market share, but in the next period the firms proceed to state 1.

Observe that the firms never make use of "review" strategies, where they try to infer the likelihood of a sequence of reported cost draws.[19] Because the collusive scheme provides firms the incentives to report truthfully in each period, the firms are not concerned with the possibiltity of past misrepresentations. Even after a history where (L, H) is realized 10 periods in a row, firms start period 11 by following the strategies specified in state 5, without worrying about how long they have been there.

4.1.2. Example: Obstacles to First-Best Collusion

Suppose now that firms are less patient and consider the factors that limit their ability to sustain first-best profits. In the example above, the limiting factor is that IC-Off1$_{LL}$ binds when implementing state 5. When firms are less patient, therefore, firm 1 would undercut the collusive price, charging $r - \varepsilon$, in state 5 when it draws a low cost.

What is to be done when δ is too small to support first-best profits? One possibility is to reduce productive efficiency in all states. This would yield a line segment of equilibrium values, where the total profit is less than first-best. But such a solution may be too drastic. A more profitable equilibrium can be attained if firms are treated symmetrically and use productive efficiency in the first period, but then use productive inefficiency in the subsequent asymmetric states. In general, this approach dominates one where the firms implement asymmetric states by lowering prices.

To be more precise, we return to the parameter values of the last subsection, except we now take $\delta = .768$. Again, we set $\lambda = 1$, noting that the qualitative features of the solution are maintained under more severe punishments. First, consider constructing an

[19] See Radner (1981) for a first-best result for infinitely patient firms that use review strategies in a "hidden-action" game. Our firms achieve first-best profits, even though they are not infinitely patient. In addition, when firms are less patient, real inefficiencies may be required to provide incentives, and we characterize below the optimal manner in which to provide such incentives.

equilibrium set that is a line segment. To this end, we may impose the binding constraints described in the last subsection, while setting $\delta = .768$ and now allowing q^1_{LH} and q^1_{HL} to vary. It is straightforward to calculate that the best equilibrium with these features has $q^1_{LH} = .992$ and $q^1_{HL} = 0$ when implementing the endpoint (x, y), and it yields per-period expected profits of .66903 for each firm.

Now consider a more sophisticated equilibrium, illustrated in Figure 3b. To simplify the description of the equilibrium, we allow the firms to randomize among continuation equilibria, although it is possible to achieve the same payoffs without such randomization by introducing new states, where q^1_{LL} and q^1_{HH} are chosen appropriately. To denote the continuation play where the firms proceed to state 2 with probability .83 and to state 4 with probability .17, we write "(2,4), (.83, .17)."

In this equilibrium, productive efficiency is used in states 2, 3, and 4, while productive inefficiency is used in states 1 and 5. The firms use productive efficiency in state 3, but then productive inefficiency is used in implementing rewards and punishments following realizations of either (L, H) or (H, L). Subsequently, productive inefficiency is used in some periods but not others, depending on history. The sum of firm profits in states 1 and 5 is strictly less (by .004) than the sum of profits in states 2, 3 and 4. Ex ante expected firm profits in this equilbrium are .66964, higher than those in the simpler equilibrium described above. This illustrates a theme that we will return to below in our theoretical characterizations: colluding firms of moderate patience use greater productive efficiency to implement fairly symmetric equilibrium values, and reduced productive efficiency when implementing highly asymmetric equilibrium values.

What would happen if, instead of reducing productive efficiency in states 1 and 5, the firms were to lower p^1_{LL}? Reducing p^1_{LL} allows the firms to further reduce q^1_{LL} without violating the off-schedule constraint, and, it relaxes the on-schedule constraints, since the low-cost type gets lower profits. Continuing with our parametric example, if firms follow an equilibrium with the same structure as Figure 3b, except they always use productive efficiency, but reduce prices and q^1_{LL} in state 5 (and symmetrically in state 1), the highest ex ante expected profits per period for each firm that can be supported are .66930,[20] lower than the equilibrium of Figure 3b. Intuitively, lowering q^1_{LH} increases profits for firm 2, while lowering p^1_{LL} makes both firms worse off.

[20] These profits are computed by solving a system of equations, analogous to the system used to compute the equilibrium of Figure 3b. Letting the profits in state 5 be (x, y) and those in state 4 be (t, w), the system imposes productive efficiency in all states, and pricing efficiency except in states 1 and 5. When implementing states 4 and 5, the following constraints bind: $q^1_{HH} = 0$, $v^1_{HH} = t$, $v^1_{LH} = x$, $v^1_{HL} = w$, IC-On1$_D$, IC-On2$_D$, and IC-Off1$_{LL}$.

4.2. The Shape of the Pareto Frontier

The examples above demonstrate some of the tradeoffs that firms face when they are too impatient to implement the first best. We now provide more general characterizations of the informative PPE utility set. We begin by characterizing the shape of the Pareto frontier of the informative PPE utility set. As discussed at the start of Section 4, when firms attempt to implement highly asymmetric equilibrium values, the off-schedule incentive constraints bind and some inefficiency may be required. We thus anticipate that total cartel profits fall as values become more asymmetric, indicating that the frontier is typically nonlinear. In the present subsection, we establish conditions under which a *subset* of the Pareto frontier of the set of informative PPE values is a line with slope equal to -1. In addition, we characterize the manner in which the off-schedule constraints determine the boundaries of this linear subset (as well as the boundaries of the frontier itself).

To begin, we recall our assumption that firms can randomize between continuation equilibria, which ensures that firms have available a convex set of continuation values at any point in time. Figure 2 illustrates the general shape of a symmetric, convex set of continuation values. The set has four "corners," labelled as North, South, East, and West, or v_N, v_S, v_E, v_W, where $v_N = (v_N^1, v_N^2)$ and likewise for the other corners. Between two corners, the boundary of the set is monotone. The part of the boundary between v_N and v_E is of particular interest to us, since it represents set of Pareto efficient continuation values.

When describing the Pareto frontier of the set of feasible continuation values given a set V, we use the notation

$$f(v_{jk}^1) = \begin{cases} \max\{v_{jk}^2 : (v_{jk}^1, v_{jk}^2) \in co(V)\} & \text{if } v_{jk}^1 \in [v_N^1, v_E^1] \\ v_N^2 & \text{if } v_{jk}^1 < v_N^1 \\ -C \cdot (v_{jk}^1 - v_E^1) & \text{if } v_{jk}^1 > v_E^1 \end{cases}$$

for some large constant C. Of course, convexity of the set V implies concavity of the frontier f. We define the function f outside the domain of the Pareto frontier in order to simplify the statement of some of our results about the slope of the frontier.

Given our assumption that firms are symmetric, $f(v) + v$ is maximized at $v^1 = v_s^1$, where $f(v_s^1) = v_s^1$. We may thus say that a scheme is characterized by *future inefficiency* if $v_{jk}^1 + f(v_{jk}^1) < 2v_s^1$ for some (j,k), so that under some state the continuation values fail to maximize total cartel future profits. As mentioned above, future inefficiencies are associated with highly asymmetric values, and represent an efficiency cost that is incurred when firms attempt to provide incentives with such values. Thus, it is important to identify conditions under which a subset of the Pareto frontier has slope of -1, so that the firms may make some use of future market-share favors without efficiency loss.

For the informative PPE set V^I, let v_s^I be the point on the Pareto frontier of V^I that provides equal utility to both firms. Consider a policy vector that implements v_s^I, and assume that the off-schedule constraints do not bind in states (L, L) and (H, H). Suppose for simplicity that pricing efficiency is used. By lowering firm 1's market share in state (L, L) by ε/η_L and (H, H) by ε/η_H, it is possible to transfer market share from firm 1 to firm 2 without upsetting any of the on-schedule incentive constraints. This new scheme is also feasible. While firm 1's profit is lower, total cartel profit is unchanged, and so the Pareto frontier has an interval with slope equal to -1.

How can we ensure that the off-schedule constraints do not bind in states (L, L) and (H, H)? Without loss of generality, when implementing v_s^I, we may specify that $q_{jj}^1 = 1/2$ and $v_{jj} = v_s^I$ for $j \in \{L, H\}$. With this specification in place, and observing that the most demanding circumstance from the perspective of off-schedule constraints arises in state (L, L) when $p_{LL} = r$, we see that it suffices to check the following condition:

$$(r - \theta_L)/2 < \delta(v_s^{1I} - \underline{v}^1). \tag{4.2}$$

Of course, v_s^{1I} is endogenously determined. We illustrate a range of discount factors where (4.2) holds in our computational examples. However, to define a lower bound on v_s^{1I} that depends only on exogenous parameters, we present the following lemma, which describes a self-generating set of equilibrium values that exists for a range of discount factors.

Lemma 6. *There exists a $\delta^{lin} < 1$ such that, for all $\delta > \delta^{lin}$, there exist values $y > x > 0$ such that the set $[(x, y), (y, x)] \cup \mathbf{u}^{NE}$ is a self-generating set of informative PPE values. Each utility pair $\mathbf{u}$ on the segment can be implemented using a policy vector $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ such that pricing efficiency holds, $v_{jk} \in [(x, y), (y, x)]$ for each (j, k), and $q_{LH}^1 + q_{HL}^2 = 1 + \frac{\delta(\kappa+\eta_L^2)}{1+\kappa-\delta\eta_H}$.*

For the parameter values used in our examples ($r = 2.5$, $\theta_H = 2$, $\theta_L = 1$ and $\eta_L = .6$), $\delta^{lin} \approx .7$, and at that discount factor $q_{LH}^1 + q_{HL}^2 \approx 1.5$, less than the first-best value of 2. More generally, this result establishes a lower bound for v_s^{1I} : since the set $[(x, y), (y, x)]$ described in the lemma must be contained in V^I, v_s^{1I} must be greater than $(x + y)/2$. Using straightforward computations, if $\kappa > .275$, then (4.2) holds for all $\delta > \delta^{lin}$.[21]

When (4.2) holds, we have an initial characterization of the Pareto frontier:

Proposition 2. *Assume (4.2).*
(i) The Pareto frontier of V^I has an open interval with slope equal to -1.
(ii) Let (x, y) and (y, x) denote the endpoints of this open interval, and let $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ implement (x, y). Then at least one of the following holds: (a) for some $j \in \{L, H\}$, IC-Off1_{jj}^I binds; (b) $v_{jj}^1 \leq x$ for some $j \in \{L, H\}$, and either $p_{jj} \leq \theta_j$ or $q_{jj}^1 = 0$ for some $j \in \{L, H\}$.

Part (i) confirms the existence of a subset of the Pareto frontier with slope -1. Part (ii) then identifies the factors that limit this subset. Formally, when implementing an endpoint

[21] The lower bound on κ is calculated when $\eta_L = 1/2$. This bound is decreasing in η_L, and so the bound is relaxed when $\eta_L > 1/2$.

of this subset, either an off-schedule constraint binds, or else the firms run out of market-share favors and the ability to shift continuation values in the event of ties. In either case, the firms cannot implement any further transfer of utility away from firm one and towards firm two without a loss of efficiency. For firms of moderate patience ($\delta < \delta^{FB}$), the off-schedule constraints typically bind first.

Next, we observe that the off-schedule constraints also determine the endpoints of the entire Pareto frontier and that they force the firms to bear inefficiency when implementing those endpoints.

Proposition 3. *Suppose that* $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ *implements* v_N^I. *(i) If IC-On*i_U *is slack for each* i, *then there is either pricing inefficiency, productive inefficiency or both. (ii) If* $\bar{q}_L^i > \bar{q}_H^i$ *for each firm* i, *at least one of the following holds: (a) for some* $j \in \{L, H\}$, *IC-Off*1_{jj}^I *binds; (b)* $v_{jj}^1 \leq v_N^{1I}$ *for some* $j \in \{L, H\}$, *and either* $p_{jj} \leq \theta_j$ *or* $q_{jj}^1 = 0$ *for some* $j \in \{L, H\}$.

To understand part (i), suppose that the firms implement some equilibrium value while using productive and pricing efficiency. Firm 1's off-schedule constraints are then slack in state (L, H); therefore, so long as the upward on-schedule incentive constraints are slack,[22] firm 1 could give up some market share in state (L, H) without violating any incentive constraints. The feasibility of this utility transfer indicates that the firms originally could not have been implementing the corner, v_N^I. Part (ii) is similar to Proposition 2 (ii).

Finally, we consider whether the set of equilibrium values is itself convex. Since payoffs and constraints are nonlinear in market shares and prices (they depend on q_{jk}^1, $q_{jk}^1 \cdot p_{jk}$ and $(1 - q_{jk}^1) \cdot p_{jk}$), $\mathcal{F}^I(V)$ is not generally convex, and V^I may not be convex either. When prices are the same in two distinct equilibria, however, the nonlinearity does not pose a problem, and the convex combination of two equilibrium values can be implemented using a convex combination of the two associated policy vectors. In the next subsection, we analyze conditions under which prices are always equal to the reservation value r when implementing values on the Pareto frontier of the equilibrium set.

4.3. Pricing and Continuation Value Efficiency

We now consider the implementation of Pareto efficient informative PPE values, and we establish important circumstances under which the implementation of such values requires that pricing and continuation-value Pareto efficiency are used. These results indicate that, even if firms are only moderately patient, when they collude optimally, they often maintain pricing and continuation-value Pareto efficiency. We explain as well that these properties imply that the downward on-schedule constraints are typically binding.

[22] We observe also that, if IC-Oni_D binds for firm i, then IC-Oni_U can bind only if $\bar{q}_L^i = \bar{q}_H^i$, indicating productive inefficiency. We discuss below conditions under which IC-Oni_D binds.

To begin, we consider the implementation of any Pareto efficient equilibrium value such that: (i) the off-schedule constraints are slack, and (ii) the Pareto frontier is sufficiently wide, and the equilibrium value is sufficiently far from the corners of the Pareto frontier, v_N^I and v_E^I, that the firms implement the equilibrium value using continuation values strictly between v_N^I and v_E^I. As Proposition 1 indicates, these properties hold when implementing values in the neighborhood of v_s^I for discount factors that exceed δ^{FB}. However, conditions (i) and (ii) apply in a wider set of circumstances. In particular, for a range of more moderate discount factors, there is a set of equilibrium values on the interior of the Pareto frontier that can be implemented with slack off-schedule constraints (though these constraints bind in subsequent periods, when implementing values that are sufficiently asymmetric).

Proposition 4. *Let* $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ *implement a Pareto efficient utility pair in* V^I. *Suppose that the off-schedule constraints hold with slack, and* $v_N^{1I} < v_{jk}^1 < v_E^{1I}$ *for all* (j,k). *Then: (i) continuation values are Pareto efficient, and (ii) prices are efficient.*

Part (i) is proved by showing that it is possible to adjust continuation values in pairs (moving them closer together, farther apart, or increasing both) in ways that do not affect the incentive constraints, but move the values closer to the Pareto frontier. Part (ii) follows because it is always possible to raise prices and lower continuation values to keep the firms indifferent, and we establish in Part (i) that continuation values below the frontier can be strictly improved upon.

We next establish conditions under which pricing and continuation-value Pareto efficiency are necessary, even when we relax the constraint that the continuation values lie strictly between the corners of the Pareto frontier.

Proposition 5. *Suppose that (3.3) holds. Suppose that* $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ *implements a Pareto efficient utility pair in* V^I*, the off-schedule constraints hold with slack, and further either (a) both* v_{HH} *and* v_{LL} *are on the interior of the line segment on the Pareto frontier of* $co(V)$ *with slope equal to* -1*, or (b)* $0 < q_{jj}^1 < 1$ *for* $j \in \{L, H\}$*. Then: (i) continuation values are Pareto efficient, and (ii) prices are efficient.*

This result generalizes Proposition 4, under additional restrictions. To understand the restrictions, recall Lemma 5, which establishes that when (3.3) holds and the valuation set is rectangular, firms increase cartel profit by decreasing productive efficiency and increasing the efficiency of prices and continuation values.[23] As an extension of Lemma 5, this result implies that when the continuation value frontier is narrow and firms maximize cartel

[23] It can be shown that when (3.3) fails, cartel profits are maximized using productive efficiency, even at the expense of inefficient prices and continuation values. Since both instruments have limited flexibility, we expect that when implementing asymmetric utility vectors, the firms may use both pricing and continuation-value inefficiency. We will not pursue this case further here.

profit, they choose productive inefficiency over inefficient prices and continuation values.[24] It further implies that under (4.2), continuation value efficiency and pricing efficiency hold at the start of the game.

So far, we have considered only the on-schedule constraints. We now make two observations about the effects of off-schedule constraints on pricing and continuation-value efficiency. First, off-schedule constraints place downward pressure on the prices of low-cost firms, in order to reduce the incentive that such firms have to undercut the equilibrium price; however, we argue below that in some circumstances firms choose to give up productive efficiency before lowering price. Second, in establishing continuation-value Pareto efficiency, we employ arguments in which firms shift profits across states of the world. When off-schedule constraints bind, the profit of a firm in a particular state of the world may be constrained, and our characterizations are more limited.

Finally, in Athey and Bagwell (1999), we consider the implications of pricing and continuation-value Pareto efficiency. Returning to the case in which the off-schedule constraints are slack, we find that if a Pareto efficient utility pair $\mathbf{u} \in V^I$ can be implemented by some policy vector for which pricing and continuation-value Pareto efficiency holds, then there must exist a policy vector that implements $\mathbf{u}$ for which each firm's downward on-schedule constraint binds.[25] Intuitively, given that the frontier is concave, asymmetric continuation values are associated with future inefficiency; therefore, unless behavior is constrained by off-schedule considerations, if asymmetric continuation values are used to provide incentives for greater productive efficiency, then they should be used to the minimal extent possible. More generally, this finding confirms that the relevant on-schedule concern is indeed the incentive of high-cost firms to mimic low-cost firms.

4.4. Productive Efficiency

We establish above conditions under which firms use pricing and continuation-value Pareto efficiency, when implementing Pareto efficient utilities. We argue now that the case for productive efficiency is weaker. The key reason is that, when the efficiency frontier is concave, a tradeoff arises between productive efficiency today and future inefficiency.

To begin, we establish sufficient conditions under which productive efficiency *is* used. Our result applies to all points on the Pareto frontier, with no additional restrictions. In

[24] When off-schedule incentive constraints are slack, a natural question is whether the Pareto frontier is itself self-generating. As Proposition 5 Part (i) indicates, for an efficient utility pair, Pareto inefficient continuation values would be used to transfer utility, only if other mechanisms for transferring utility were exhausted. When such instruments are exhausted, however, it is conceivable that an optimal cartel might implement a utility transfer with an inefficient continuation value, particularly in the (L, L) state, so as to draw utility from a firm while simultaneously relaxing that firm's downward on-schedule incentive constraint. We note, though, that this possibility does not arise in any of our computational examples.

[25] This statement allows that firms might be indifferent among a range of equally desirable implementation schemes (as might occur if continuation values lie on a linear segment of the frontier).

particular, off-schedule constraints may or may not bind. The result generalizes Lemma 4 beyond budget-balanced transfers: since low-cost firms find high market share relatively more attractive, firms will use asymmetric continuation values to provide incentives for productive efficiency, if the future inefficiency is not too great.

Proposition 6. *Choose any Pareto efficient utility pair in V^I and let $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ be the policy vector that implements this pair. Then productive efficiency holds in state (L, H) (i.e., $q^1_{LH} = 1$) if $p_{LH} > \theta_H$ and if there exists $\varepsilon > 0$ such that*

$$1 + \Delta^-_\varepsilon f(v^1_{LH}) \equiv 1 + [f(v^1_{LH}) - f(v^1_{LH} - \varepsilon)]/\varepsilon < (\theta_H - \theta_L)/(p_{LH} - \theta_L). \tag{4.3}$$

Productive efficiency holds in state (H, L) (i.e., $q^1_{HL} = 0$) if $p_{HL} > \theta_H$ and if there exists $\varepsilon > 0$ such that

$$1 + \Delta^+_\varepsilon f(v^1_{HL}) \equiv 1 + [f(v^1_{HL} + \varepsilon) - f(v^1_{HL})]/\varepsilon > -(\theta_H - \theta_L)/(p_{HL} - \theta_L). \tag{4.4}$$

As suggested, productive efficiency is used when implementing Pareto efficient utilities, if the continuation values in the (L, H) and (H, L) states are drawn from regions of the frontier at which the frontier slope does not depart too greatly from -1. Since (4.2) implies that the frontier has a linear portion, it then follows that some productive efficiency is used at the start of the game, when implementing v^I_s.[26]

Proposition 6 provides sufficient but not necessary conditions for productive efficiency. We now tighten the characterization under the assumptions that the off-schedule constraints are slack, the upward on-schedule constraints are slack, and utility is transferrable without efficiency loss (as in conditions (a) and (b) of Proposition 5). For example, under (3.3), these assumptions imply pricing and continuation-value Pareto efficiency (by Proposition 5); yet, as we now confirm, the case for productive efficiency is weaker.

Proposition 7. *Suppose that $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ implements an equilibrium value $\mathbf{u}$ on the Pareto frontier of V^I and either assumption (a) or (b) of Proposition 5 is satisfied. Further, suppose that the off-schedule constraints and IC-Oni$_U$ are slack for each i. Finally, select an equilibrium such that $p_{LH} = p_{HL} > \theta_H$, $q^1_{LH} = q^2_{HL}$, $v^1_{LH} = v^2_{HL}$, and that no other such policy vector implements $\mathbf{u}$ using a larger q^1_{LH}.*[27] *Then $q^1_{LH} \in (1/2, 1)$ and $q^1_{HL} \in (0, 1/2)$ if*

[26] Under (4.2) a symmetric scheme may be implemented, with $q^i_{jk} = 1/2$, $p_{jk} = r$, and $v_{jk} = v^I_s$ for all (j, k). This is a best scheme with no productive efficiency. But by Proposition 2, the interior of the Pareto frontier then has slope equal to -1. Proposition 6 then implies that the described scheme is Pareto dominated by an alternative scheme with at least some productive efficiency, implemented using v_{LH} and v_{HL} (at least weakly) outside the interval of the Pareto frontier with slope -1.

[27] To state necessary and sufficient conditions for productive efficiency, we must confront the possibility that a range of policy vectors implements the same equilibrium values. Then, we focus on equilibria that satisfy two criteria. First, we select the policy vector with the highest level of productive efficiency. Second, we focus on schemes that are symmetric across states for which the cost types are unequal. To see that such a scheme exists, observe that as long as utility is transferrable, the optimal scheme maximizes the sum of firm profits and transfers utility in states (L, L) and (H, H); the symmetry of the model then implies the desired symmetry of policy vectors.

and only if for all $\varepsilon > 0$,

$$1 + \Delta_\varepsilon^+ f(v_{LH}^1) \leq \frac{\theta_H - \theta_L}{p_{LH} - \theta_H + \eta_L(\theta_H - \theta_L)} < 1 + \Delta_\varepsilon^- f(v_{LH}^1). \tag{4.5}$$

Further, $q_{LH}^1 = 1/2$ and $q_{HL}^1 = 1/2$ if and only if the second inequality holds at $v_{HL}^1 = v_{LH}^1 = v_s^{I1}$, while $q_{LH}^1 = 1$ and $q_{HL}^1 = 0$ if and only if the second inequality fails.

One implication of this result is that if $v_E^{1I} - v_N^{1I} < (r - \theta_H)/\delta$ (the minimum width required to implement productive efficiency using pricing and continuation-value Pareto efficiency), there is productive inefficiency even in the first period of play, when implementing v_s^I, so long as the off-schedule constraints do not bind. Further, we see that if future inefficiency is extreme, so that $\Delta_\varepsilon^- f(v_s^{I1}) = 0$, the second inequality in (4.5) holds when $p_{LH} = r$ if and only if (3.3) holds. Thus, we can interpret Proposition 7 as a generalization of Lemma 5. In general, firms implement some productive efficiency; however, they stop short of full productive efficiency if the slope of the frontier gets too steep or too flat, and in particular, if the frontier is too narrow.[28]

Let us now summarize our characterizations of Pareto efficient collusive schemes for firms of moderate patience. First, we find that firms are willing to bear a moderate future inefficiency to gain productive efficiency in the present. Second, when off-schedule constraints do not bind, and either the Pareto frontier is wide enough or (3.3) holds, the firms maintain pricing and continuation-value Pareto efficiency, even at the possible expense of productive efficiency. Finally, the firms may sacrifice even pricing and continuation-value Pareto efficiency when they attempt to implement asymmetric equilibrium values, if the off-schedule incentive constraints directly or indirectly prevent the use of future market-share favors in the event of ties. Thus, for firms of moderate patience, we expect to start the game using fairly efficient schemes (at worst, there is some productive inefficiency), but the schemes may incorporate additional inefficiencies following a series of one or more cost realizations whereby one firm has lower cost than another.

4.5. Computational Examples

In this subsection, we develop a computational example to illustrate some of the tradeoffs and themes from Section 4. We begin by offering some remarks about our computational approach. Motivated by Abreu, Pearce, and Stacchetti (1990), we specify a set V^0 and then compute $V^t = \tilde{T}^I(V^{t-1})$ for $t = 1, ...,$ iterating until the distance between the sets

[28] Additional characterizations can be provided. For example, if the upward on-schedule constraints are slack, then for ε sufficiently small, if $\Delta_\varepsilon^- f(v_{HH}^1) < -1$, then $q_{HH}^1 = 1$, and if $\Delta_\varepsilon^+ f(v_{HH}^1) > -1$, then $q_{HH}^1 = 0$. In other words, the firms take the market shares to the extreme before incurring future inefficiency in state (H, H). Also, when off-schedule constraints are slack, f is differentiable, and all continuation values are interior, Pareto efficient points require $f'(v_{LH}^1) \cdot f'(v_{HL}^1) = f'(v_{LL}^1) \cdot f'(v_{HH}^1)$. Intuitively, the continuation values are chosen to balance the inefficiencies incurred in each state of the world.

becomes lower than a given tolerance level (.0001 in our computations). To operationalize this algorithm, a natural method is to divide each set V^t into a grid, and then check which members of this grid survive to become members of V^{t+1}. This approach is slow, however. Following Wang (1994), we use a trick which speeds up the computations.[29] At the start of the algorithm, we divide the set $[0, r/(1-\delta)]$ into a fixed grid, where we let $\mathbf{x}$ denote the vector of points in this grid. The grid represents the set of feasible continuation values for firm 2, and these are the only values ever permitted for firm 2. On each iteration of the algorithm, we compute the set of continuation values for firm 1 that can be sustained for each element of the grid.

To further ease the computational burden, we impose two restrictions. First, we assume that firms punish off-schedule deviations by reverting to the static Nash equilibrium. This restriction does not directly affect the qualitative characterization of the efficiency frontier, since in the computations firms only leave the efficiency frontier off of the equilibrium path. Repeating the computations using lower punishments affects only the level of the discount factor at which different types of equilibria can be supported. Second, we consider only equilibria where firms use pricing efficiency on the equilibrium path. This restriction certainly matters for impatient firms, but without it the computation becomes much more complex. Given the restrictions we have imposed, the equilibrium sets we construct should be interpreted as lower bounds on the Pareto frontier of equilibria.

Figures 4 and 5 illustrate equilibrium sets for particular parameter values.[30] Consider Figure 4 in relation to Proposition 4. Neither the off-schedule constraints nor the constraints on the width of the Pareto frontier are binding for the policy vectors that implement states 9 to 22, and thus our characterizations from Proposition 4 apply for those states. Continuation-value Pareto efficiency indeed holds: in every state, after every realization of cost types, the firms move to another state on the Pareto frontier. Given that, the diagram only labels and represents the Pareto frontier. Similar results hold in Figure 5, where the conditions of Proposition 5 are satisfied when implementing states 9-17. Observe that the Pareto frontier is narrow, and the implementation of Pareto efficient utilities is achieved with continuation values that fall on the corners (following the (L, H) and (H, L) states). Nevertheless, as Proposition 5 requires, the continuation values are always Pareto efficient.

Now consider productive efficiency in Figure 4. Notice that for a wide range of states (9

[29] Wang's (1994) approach builds on Phelan and Townsend (1991). Recently, Judd and Conklin (1995) have developed approaches to computation that, if extended to this model, could be more efficient. As our aim is only to illustrate the theoretical results, we do not pursue this here.

[30] Observe that the computations yield slightly asymmetric continuation values across the two firms; this arises as a result of the computational algorithm, which treats one firm's profit as discrete and the other's as continuous. Further, on the region where the continuation value frontier is approximately linear, there are often many ways to implement a given value; but, due to the discretization of the frontier, the firms may have a strict preference among alternative collusive schemes giving approximately the same utility. Thus, behavior sometimes "jumps" drastically among nearby states. This does not qualitatively affect the computation of the equilibrium set.

to 22), productive efficiency is (approximately) implemented, as predicted by Proposition 6: at the extreme continuation values (in states (L, H) and (H, L)) associated with these states, the slope of the frontier is always within the bounds specified in (4.3) and (4.4). Further, states 9 to 21 use the same extreme continuation values ($v_{LH} = 24$ and $v_{HL} = 3$); due to concavity of the frontier, increasing or decreasing both continuation values would reduce total utility across the two firms. Notice also that the firms use productive inefficiency when implementing the asymmetric values of the Pareto frontier. In particular, states 3 and 24 use productive inefficiency; since these states (or less efficient ones) are reached with positive probability from every starting point, even the most profitable points on the Pareto frontier yield less than the first-best profits. This is consistent with Proposition 6: when implementing state 24, $v_{LH} = 26$, the corner of the Pareto frontier, and so (4.3) fails, indicating that productive efficiency is not necessarily optimal.

Now consider Figure 5. Across all but the most extreme states, the overall level of productive efficiency is approximately the same, with $q^1_{LH} + q^2_{HL}$ approximately equal to 1.57, and incentives for truth-telling are provided in a similar fashion, with the firms going to state 1 following a realization of (H, L) and to state 20 following a realization of (L, H). These states correspond to the corners of the frontier, and so the fact that firms achieve only partial productive efficiency is consistent with Proposition 6. However, consistent with Proposition 5, the firms do *not* use continuation-value inefficiency to implement higher levels of productive efficiency. Finally (and similar to Figure 3b), to implement the extreme states, somewhat greater productive inefficiency is required. Thus, colluding firms capture some productive efficiency in the first period of the game, but at the cost of greater inefficiency in the future. The result is a concave Pareto frontier.

5. Informative v. Uninformative Communication

We now consider the role of communication. We begin by contrasting the case of informative communication with the opposite possibility, where firms are unable or unwilling to communicate. Building on this analysis, we then discuss the qualitative features of the set of unrestricted PPE, where firms choose whether to use informative communicate in any period as a function of the history of play.

Recall the extensive-form game defined in Section 2. Any communication occurs first and then firms make pricing decisions and market-share proposals, where the market-share proposals affect outcomes only when prices are equal. In this game, to capture a situation in which firms are unable or unwilling to communicate, we simply specify that firms use the uninformative announcement $\alpha^i(\theta^i) = N$ in all states of the world. Recall that following this announcement, if firms charge the same price, they must share the market.[31]

[31] As we show in Athey and Bagwell (1999), if firms were allowed to withhold quantity in a decen-

In this context, how are the firms affected by requiring announcements to be uninformative? This requirement has both a cost and a benefit. The cost is easily understood: in the absence of informative communication, the set of market-sharing arrangements that can be implemented is restricted, since state-contingent arrangements are then feasible only when they are compatible with decentralized decision making. But how severe is this restriction?

In our Bertrand setting, the restriction is less severe than one might expect. For example, a simple no-communication scheme sets $\rho^2(\theta_H) = \rho^1(\theta_H) = r$ and $\rho^2(\theta_L) = \rho^1(\theta_L) = r - \Delta$. This yields productive efficiency, equal market shares in ties and approximate pricing efficiency (for $\Delta > 0$ and small). Similarly, by setting $\rho^2(\theta_H) = r$, $\rho^1(\theta_H) = r - \Delta$, $\rho^2(\theta_L) = r - 2\Delta$ and $\rho^1(\theta_L) = r - 3\Delta$, the firms may continue to achieve productive and approximate pricing efficiency, but now firm 1 wins all ties.

Despite these examples, the restriction is real. First, in our Bertrand setting, many market-sharing arrangements are infeasible without informative communication. For example, any arrangement with $q^i_{jk} \notin \{0, .5, 1\}$ requires informative communication. Second, our Bertrand model understates the actual cost of decentralized behavior, as it abstracts from a variety of benefits to communication and "advanced planning" that naturally arise in other models. In the Bertrand model, firms only bear cost for realized market share, and it is costless to be "prepared" to serve those consumers. Other models, such as Cournot, would entail much greater costs to decentralization.[32] To capture costs of this kind, we define $\Delta \geq 0$ as the minimum price difference that can be perceived by consumers (e.g. pennies or dollars), so that $\Delta > 0$ provides a crude means of representing the cost of allocating market shares decentrally, through price differences. We interpret $\Delta = 0$ as an approximation for the case where Δ can be arbitrarily small.

The absence of informative communication also has a benefit, once the off-schedule constraints are considered. When informative communication is absent, each firm must be dissuaded from deviating after observing its own type, but before knowing the type of the other firm. In other words, the off-schedule incentive constraints bind at the interim stage. For example, suppose that an equilibrium specifies $q^1_{LL} = 1/2$, $q^1_{LH} = 1$, $q^1_{HL} = 0$ and $q^1_{HH} = 1/2$. This market-share allocation can be achieved without informative communication. If the firms communicate, IC-Off1^I_{LL} might bind, because after communication, firm 1 knows that the state is (L, L) and is tempted to cut price slightly and pick up the remaining $1/2$ of the market. By contrast, if the firms do not communicate, a low-cost firm 1 is unaware of firm 2's cost type, so that its expected market share is $\eta_H + (1 - \eta_H)(1/2) > 1/2$, leaving

tralized way, the range of outcomes would expand somewhat, but the qualitative results of this section would not change. We do continue to allow firms to randomize among continuation equilibria. Formally, randomization does not require communication.

[32] In a Cournot model, without communication colluding firms typically suffer large inefficiencies: a high-cost firm produces, just in case the opponent is high-cost as well, but this is wasteful if the opponent turns out to be low cost. In contrast, when the firms can communicate, they are able to select the monopoly output for the cost type of the firm that actually produces, and other firms refrain from producing.

less to gain from a deviation. The absence of informative communication can thus promote cooperation, by preserving uncertainty about opponent play and softening the off-schedule incentive constraint. Notably, uninformative communication relieves some of the pressure to give up productive efficiency that is present in an informative PPE, since it becomes easier to maintain $q^1_{LH} = 1$ and lower q^1_{LL} without violating the off-schedule constraints.

To formally represent the incentive constraints under uninformative communication, let $p_{jk} = \min\{\rho^1(\theta_j), \rho^2(\theta_k)\}$ be the transaction price for state (j, k). The market share received by firm 1, q^1_{jk}, is determined as described in Section 2. Finally, let v^1_{jk} represent the continuation value for firm 1 that is induced by the price selections $\rho^1(\theta_j)$ and $\rho^2(\theta_k)$. The on-schedule incentive constraints are again represented by IC-Oni_D and IC-Oni_U. To define the off-schedule constraints, it is somewhat easier to refer directly to the decentralized pricing strategies. For $j \in \{L, H\}$, IC-Off1^U_j is defined as:

$$\sum_{k\in\{L,H\}} \eta_k[q^1_{jk}(p_{jk} - \theta_j) + \delta v^1_{jk}] \geq \max\Big((\rho^2(\theta_L) - \Delta - \theta_j),\ \eta_H(\rho^2(\theta_H) - \Delta - \theta_j),\ 0\Big) + \delta\underline{v}^1$$

while the corresponding constraint for firm 2, IC-Off2^U_k, is defined analogously.

For a given continuation-value set V, we now define a function $C(\mathbf{p}, \mathbf{q}, \mathbf{v})$, where $C : \Re^4_+ \times [0, 1]^4 \times V^4 \to \{0, 1\}$. We let $C(\mathbf{p}, \mathbf{q}, \mathbf{v}) = 0$ if there exist decentralized pricing strategies $(\rho^1(\cdot), \rho^2(\cdot))$ that can induce the specified prices, market shares and continuation values, while $C(\mathbf{p}, \mathbf{q}, \mathbf{v}) = 1$ if informative communication is necessary.[33] When $C(\mathbf{p}, \mathbf{q}, \mathbf{v}) = 1$, the off-schedule constraints defined previously, IC-Offi^I and IC-Off-Mi, are appropriate, while the IC-Offi^U constraints are appropriate if $C(\mathbf{p}, \mathbf{q}, \mathbf{v}) = 0$. The feasible set when firms use uninformative communication is written:

$$\mathcal{F}^U(V) = \{\mathbf{z} = (\mathbf{p}, \mathbf{q}, \mathbf{v}) \in \mathcal{Z}(V) :\ C(\mathbf{p}, \mathbf{q}, \mathbf{v}) = 0;\ \text{For all } i = 1, 2,\ j \in \{L, H\},\ \text{IC-On}i_D,\ \text{IC-On}i_U \text{ and IC-Off}i^U_j \text{ hold.}\}$$

As the off-schedule constraints are different from the case of informative PPE, neither $\mathcal{F}^U(V)$ nor $\mathcal{F}^I(V)$ is a subset of the other. The set of uninformative PPE, V^U, can then be characterized as the largest invariant set of the following operator:

$$\tilde{T}^U(V) = \{(u^1, u^2) :\ \exists \mathbf{z} = (\mathbf{p}, \mathbf{q}, \mathbf{v}) \in \mathcal{F}^U(V) \text{ such that for } i = 1, 2,\ u^i = \bar{U}^i(\mathbf{z})\} \cup \mathbf{u}^{NE}.$$

An initial observation is that in an uninformative PPE, even if firms collude at the reservation price and the off-schedule constraints are slack, the set of feasible policy vectors $\mathcal{F}^U(V)$ is not convex. Thus, the set of equilibrium values may not be convex, so that we rely more heavily on our assumption that firms can randomize among continuation equilibria.

We now discuss circumstances under which restricting communication might hurt firms if the off-schedule constraints do not bind. Consider the choice of q^1_{LH}. In regions where the

[33] Notice that, when $C(\mathbf{p}, \mathbf{q}, \mathbf{v}) = 0$, continuation values can only be state-contingent to the extent that the state of the world is revealed by the prices.

continuation value frontier is too steep or too flat, or if the continuation value frontier is too narrow, Propositions 6 and 7 establish that the firms implement productive inefficiency. In such cases, intermediate values of q^1_{LH} are optimal, so that the restriction to uninformative communication may be costly. More formally, observe the tradeoff between productive efficiency and future inefficiency can be characterized in a manner analogous to Proposition 6. A scheme $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ where $q^1_{LH} = 1/2$ can be improved upon by a scheme where $\tilde{q}^1_{LH} = 1$ and $\tilde{v}^1_{LH}$ is chosen to satisfy IC-On1$_D$ (holding the rest of the scheme fixed), if

$$1 + [f(v^1_{LH}) - f(\tilde{v}^1_{LH})]/(v^1_{LH} - \tilde{v}^1_{LH}) < (\theta_H - \theta_L)/(p_{LH} - \theta_L).$$

However, if the frontier is too narrow, or if it eventually becomes too steep, firms sacrifice productive efficiency even if (4.3) holds so that, were it available, a *small* increase in q^1_{LH} (holding the market shares in other states fixed) would increase profits. Thus, restricting communication may lead to greater productive inefficiency.

Circumstances may exist, therefore, under which cartel profits are reduced when firms are prohibited from informative communication. But is the possibility of such losses eliminated when firms are sufficiently patient? We establish next that, even when communication is prohibited, there exists a critical discount factor strictly less than 1 above which first-best is attained when $\Delta = 0$. In particular, the linear self-generating segment constructed in Proposition 1 can be implemented without communication, but at a lower discount factor (provided in the Appendix). The absence of informative communication is beneficial when implementing the "corner" value of the equilibrium set, (x, y). Since the firms use productive efficiency, firm 1 has no incentive to deviate in state (L, H); by contrast, in state (L, L), firm 1 produces less than unity (in the implementation we use, $q^1_{LL} = 0$), and the off-schedule constraint binds under informative communication. By refraining from communication, the firms pool the (L, L) off-schedule constraint with the non-binding (L, H) off-schedule constraint.

Proposition 8. *Assume (4.1). Then there exists $\delta^{NC} < 1$ such that, for $\delta \in [\delta^{NC}, 1]$ and $\Delta = 0$, there exists an uninformative PPE that yields first-best profits to the cartel: $(\pi^{FB}/(1-\delta), \pi^{FB}/(1-\delta)) \in V^U$.*

Restrictions on communication thus hurt collusive ventures only if firms are moderately patient or Δ is large. At the same time, it is important to emphasize that the proof of Proposition 8 exploits the assumed ability of non-communicating firms to randomize over continuation play. Absent this ability, for a range of discount factors firms could achieve first-best profits only if some histories were followed with informative communication.

Finally, consider unrestricted PPE. In each period, the firms first choose whether or not to use informative communication. If so, they reveal their types and face the IC-Offi^I constraints; otherwise, they choose from a restricted set of market-share and price policies, but they face the relaxed IC-Offi^U constraints. Formally:[34]

$$\tilde{T}(V) = \{(u^1, u^2) : \exists \mathbf{z} = (\mathbf{p}, \mathbf{q}, \mathbf{v}) \in \{\mathcal{F}^U(V) \cup \mathcal{F}^I(V)\} \text{ s.t. for } i = 1, 2,\ u^i = \bar{U}^i(\mathbf{z})\} \cup \mathbf{u}^{NE}.$$

In general, the firms will use communication in order choose a policy vector from $\mathcal{F}^I(V)$, so as to implement market-sharing arrangements that are not available using decentralized schemes; however, when there is a significant gain from relaxing off-schedule constraints (e.g., $q^1_{LL} << q^1_{LH}$) *and* when the "ideal" market shares are close enough to a scheme that can be implemented without communication, firms refrain from communicating, choosing a policy vector from $\mathcal{F}^U(V)$. Such "breaks" in communication are especially likely when firms attempt to implement a very asymmetric utility pair.

So long as (4.1) holds and $\delta \geq \delta^{FOn}$ is not the limiting factor, for $\Delta > 0$ there will be a region of discount factors (which contains $[\delta^{FB}, 1]$) such that firms choose to communicate on the equilibrium path. On the other hand, there will be a lower region of discount factors where, for Δ sufficiently small, firms often avoid communication on the equilibrium path, and collusive profits are equal to first-best profits, less the distortion due to Δ. For this region, the option to refrain from communication allows strictly higher profits than a purely informative PPE. The following example illustrates.

5.1. Selective Communication: An Example

For the parameter values from our example, when firms use Nash reversion to punish off-schedule deviations ($\lambda = 1$), we compute $\delta^{NC} = .704$, which is less than .769, the lowest discount factor that supports first-best using informative communication. The difference in critical discount factors persists for each value of λ. Maintaining $\lambda = 1$ and $\Delta > 0$, firms strictly prefer to communicate in every period when $\delta \in$[.769,1), but for each $\delta \in (.704, .769)$, there is a Δ small enough such that firms strictly prefer a regime of no communication on the equilibrium path to a scheme of communication in every period.

However, the firms can do better still by using a strategy of selective communication. Suppose $\lambda = 1$, $\delta \in (.704, .769)$ and Δ is small but positive. Then, firms will prefer to communicate (saving Δ) when implementing fairly symmetric equilibrium values (for example, following a realization of (L, L), as in Figure 3a), but will always avoid communication when implementing asymmetric equilibrium values, for example following realizations of (L, H) and (H, L). Under this scheme of selective communication, total profits are lower for asymmetric equilibrium values.

[34] Observe that the firms have available the same "worst punishment" irrespective of whether they choose to communicate on the equilibrium path.

6. Bribes

In this section, we extend the base model to allow for bribes. The following stage is added to the extensive form stage game: (v) firm i sends $b^i \geq 0$ to firm j; firm j receives γb^i. The extended model is called the Bribes model. Communication is not necessary to implement bribes, since the firms can condition bribes on the market shares realized ex post. However, to simplify the exposition, we restrict attention to informative communication. The exogenous parameter $\gamma \in [0,1]$ describes the inefficiency of the bribe: $\gamma = 1$ corresponds to the use of money without any transaction costs; $\gamma = 0$ corresponds to no transfers; and $\gamma \in (0,1)$ corresponds to the case where there is some probability that a bribe will be detected by antitrust authorities, or where firms can only make in-kind transfers that have some inherent inefficiency.

Formally, the utility function with bribes for firm 1 is denoted:

$$U^{B1}(\hat{\jmath}, j|z, b) = U^1(\hat{\jmath}, j|z) + \sum_{k \in \{L,H\}} \eta_k(\gamma b^2_{\hat{\jmath}k} - b^1_{\hat{\jmath}k})$$

and likewise for firm 2. The on-schedule constraints, denoted IC-Oni^B_U and IC-Oni^B_D, are redefined using U^{Bi} as the interim expected utility function. To represent the off-schedule constraints, we observe that optimal collusion never requires a state in which both firms send bribes, since, with $\gamma \leq 1$, the desired net transfer can be achieved most efficiently if a single bribe is made. Then the off-schedule constraint for firm 1 is[35]

$$\gamma b^2_{jk} - b^1_{jk} + \delta(v^1_{jk} - \underline{v}^1) \geq \max(q^2_{jk}(p_{jk} - \theta_j), q^1_{jk}(\theta_j - p_{jk})). \tag{IC-Off1^B_{jk}}$$

IC-Off2^B_{jk} is defined analogously; likewise, IC-Off-Mi^B is constructed from IC-Off-Mi in the natural way. Let $\mathcal{F}^{IB}(V)$ be defined as $\mathcal{F}^I(V)$, once the utility functions and constraints from the base model are replaced with those in the Bribes model. The policy vector is now $(\mathbf{z}, \mathbf{b})$, where $\mathbf{z} = (\mathbf{p}, \mathbf{q}, \mathbf{v})$. Finally, with ex ante utility given as $\bar{U}^{Bi}(\mathbf{z}, \mathbf{b})$, let

$$\tilde{T}^B(V) = \{(u^1, u^2) : \exists(\mathbf{p}, \mathbf{q}, \mathbf{v}, \mathbf{b}) \in \mathcal{F}^{IB}(V) \text{ such that for } i = 1, 2,\ u^i = \bar{U}^{Bi}(\mathbf{z}, \mathbf{b})\} \cup \mathbf{u}^{NE}.$$

We denote the set of PPE values in the Bribes model as V^B, which following our previous arguments is the largest invariant set of $\tilde{T}^B$. Let v^B_s be the point on the Pareto frontier of V^B that gives equal utility to both agents.

We establish first that bribes do not fully replace market-share favors in implementing Pareto efficient equilibrium values:

[35] Notice that when IC-Off1^B_{jk} holds, if firm 1 is assigned to send a bribe to firm 2, it never has the incentive to withhold the bribe after production takes place. After production, firm 1 wishes to adhere to the equilibrium play if $\gamma b^2_{jk} - b^1_{jk} + \delta(v^1_{jk} - \underline{v}^1) \geq 0$, which holds by IC-Off$1^B_{jk}$. It is more tempting to deviate from the agreement before production takes place, thereby capturing the market and avoiding the bribe, than after production, when the firm can only avoid paying the bribe.

Proposition 9. *(i) Suppose that* $(r-\theta_L)/2 < \delta(v_s^{1B} - \underline{v}^1)$. *For all* $\gamma < 1$, *if* $(\mathbf{p}, \mathbf{q}, \mathbf{v}, \mathbf{b})$ *implements* v_s^B *and uses any productive efficiency* ($q_{LH}^1 > 1/2$), *then the associated PPE is non-stationary. If* $\gamma = 1$, *there exists a non-stationary PPE that implements* v_s^B. *(ii) Assume (4.1). For all* $\gamma < (=)1$, *there exists* $\delta^B < 1$ *such that, for all* $\delta \in [\delta^B, 1]$, *bribes are never used (resp. not necessary) along the equilibrium path in the most profitable equilibria for the firms.*

Proof. (i) Under the stated conditions, the frontier has an open interval with slope -1, and thus it is more efficient to use continuation values rather than inefficient bribes. (ii) Proposition 1 establishes a critical discount factor beyond which first-best can be attained without bribes; thus, if bribes are at all inefficient, they are not used. ■

Intuitively, if the off-schedule constraints do not bind at the symmetric point on the Pareto frontier of V^B, then the frontier has an open interval of slope -1. Thus, it is initially more efficient to use market-share favors than inefficient bribes. Further, when sufficiently patient firms can attain first-best without bribes, inefficient bribes are never used.[36]

Next, we characterize the use of bribes to provide incentives for productive efficiency, when firms are not patient enough to implement first-best:

Proposition 10. *Fix* δ *and* γ, *and suppose that* $(\mathbf{p}, \mathbf{q}, \mathbf{v}, \mathbf{b})$ *implements a Pareto efficient value* $\mathbf{u} \in V^B$. *Suppose that* $p_{HL}, p_{LH} > \theta_H$. *If* $1-\gamma < (\theta_H - \theta_L)/(p_{LH} - \theta_L)$ *and* $1-\gamma < (\theta_H - \theta_L)/(p_{HL} - \theta_L)$, *then the scheme uses productive efficiency.*

The proof of this result follows as in Proposition 6, which established that even if the off-schedule constraints bind, firms use productive efficiency unless the continuation values are not available or require too much future inefficiency. Figure 6 illustrates how bribes augment the set of continuation values.

Thus, so long as bribes are suitably efficient, firms use productive efficiency. This analysis highlights an important theme: the main factor limiting productive efficiency is the availability of an instrument for efficiently transferring utility. In the absence of bribes, if firms achieve productive efficiency today, then the utility transfer is effected through market-share favors tomorrow; furthermore, as Proposition 3 establishes, this utility transfer entails a future inefficiency if in tomorrow's tied states the off-schedule constraints bind or the firms run out of market-share favors. When bribes are available, however, firms have a less-constrained instrument with which to achieve the desired utility transfer. Bribes can thus enable a substantial improvement in productive efficiency, provided that the direct inefficiency of bribes, as measured by $(1-\gamma)$, is sufficiently small.[37]

---

[36] In a continuum-type model, the Pareto frontier has slope equal to -1 at the Pareto efficient equilibrium providing equal utility to both firms, if firms are sufficiently patient such that the off-schedule constraints permit small future market-share favors. Thus, for any $\gamma < 1$, even in a continuum-type model, bribes do not fully replace future market-share favors.

[37] In Athey and Bagwell (1999), we consider the use of bribes in Symmetric PPE. If $1-\gamma < (\theta_H -$

These results have two main implications for applied analysis of collusion. First, we observe that market-share favors are a robust feature of collusive ventures, so long as bribes are inefficient and individual firm behavior can be tracked over time. Second, our results have a somewhat perverse policy implication. To the extent that bribes are used, they increase the productive efficiency of the cartel. For many discount factors and parameter values, firms can sustain collusion at high prices, and the only issue for the cartel is the extent to which they can implement productive efficiency. Thus, for many discount factors, prohibiting bribes reduces welfare. On the other hand, there do exist moderate discount factors such that prohibiting bribes lowers the collusive profits enough that collusion takes place only at substantially lower prices. For moderately patient firms, a prohibition on bribes may raise consumer welfare.

7. Conclusions

¿From a methodological perspective, our analysis offers several contributions. First, our paper is the first of which we are aware to provide tools for characterizing the optimal use of market-share favors by impatient firms. Depending on the anti-trust environment, different instruments are available, and impatient firms may face real tradeoffs among those instruments. We identify these tradeoffs and explore them both theoretically as well as using computational examples. Second, we develop the precise connections between static and dynamic analyses of collusion, making clear the similarities and differences, and laying the groundwork for treating other repeated-game problems within the mechanism design framework. Third, our work motivates some new questions for static mechanism design, and takes some initial steps towards addressing them.[38]

The results in this paper are motivated by the problem of collusion, but they apply also in other contexts. At a general level, our model considers interactions between agents – such as family members, workers in a firm, or politicians – who must repeatedly take actions in an environment with two main characteristics: first, each agent's cost or benefit of taking the action changes from period to period, where the actual change is private information; and second, there are limits on the agents' ability to use side-payments. Essentially, the repeated play of any of the standard multi-agent mechanism design problems (public goods, auctions, bargaining) fits into the framework, with the additional assumption of restricted transfers. Private information is easy to motivate. Family members may be privately

$\theta_L)/(r-\theta_L)$ and firms are sufficiently patient, the optimal symmetric collusive scheme is stationary, and it entails productive efficiency, pricing efficiency and the use of bribes. This scheme can be implemented without informative communication.

[38]For example, we examine how restrictions on transfers (for instance, to a convex set) affect optimal mechanisms. In the literature on collusion, only a limited class of restrictions on transfers have received attention. See McAfee and McMillan (1992).

informed about how tired they are on a particular day, and thus how costly it is to perform household work. Likewise, division heads within a firm may have private information about the efficiency of access to a resource, and politicians may have private information about the costs of legislation. The scope for transfers is also often limited: families may share a common budget, division heads may share a common resource and payments for votes may be illegal. Social norms may also prohibit monetary transfers.[39]

In the context of the collusion application, our analysis suggests several directions for further research. For example, we show that a more antagonistic anti-trust policy may have perverse welfare effects: successfully colluding firms tolerate productive inefficiency before lowering prices. This conclusion, however, is perhaps sensitive to our Bertrand model, and it would be interesting to consider this feature further in a model with differentiated products or Cournot competition. Additionally, our work suggest new empirical directions. Allowing for a sophisticated cartel design, we find here that optimal collusion is complex, with considerable market-share instability. By contrast, Athey, Bagwell and Sanchirico (1998) restrict attention to symmetric equilibria and show that a simple price-fixing agreement with stable market shares is then optimal for patient firms. In combination, this work may be useful in providing a theoretical framework with which to interpret the empirical factors that influence the cartel organizational form.[40] As a further example, we note that the collusion literature does not distinguish well between market-share allocation schemes that implement productive efficiency and those that do not. For example, bid-rotation schemes are common in auctions, and Comanor and Schankerman (1975) analyzed all prosecuted cases of bid rigging over a 50-year period, but they did not distinguish between "standard" bid rotation schemes and "sophisticated" bid rotation schemes that might achieve productive efficiency. Further study of the legal testimony may identify those schemes that made use of market-share favors or bribes to implement productive efficiency.

8. Appendix

Proof of Lemma 4: Imposing pricing efficiency, productive efficiency and Pareto efficient continuation values (i.e., $v^1_{jk} + v^2_{jk} = 2K$ for all (j,k)), it is straightforward to show that IC-On1$_D$ and IC-On2$_D$ respectively bind if and only if

$$\begin{aligned} 0 &= [r-\theta_H]\{\eta_H(q^1_{HH}-1)-\eta_L q^1_{LL}\}+\delta\{\eta_H(v^1_{HH}-v^1_{LH})+\eta_L(v^1_{HL}-v^1_{LL})\}, \text{ and} \\ 0 &= -[r-\theta_H]\{\eta_H q^1_{HH}+\eta_L(1-q^1_{LL})\}-\delta\{\eta_H(v^1_{HH}-v^1_{HL})+\eta_L(v^1_{LH}-v^1_{LL})\} \end{aligned}$$

[39] In a related paper, Holmstrom and Kreps (1996) study the use of "promises" in repeated games. Our analysis differs from theirs in that we bring together the tools of dynamic programming and mechanism design to characterize optimal equilibria for firms for a given discount factor, and we explicitly model the tradeoff between different kinds of side-payments (future favors versus bribes).

[40] There is little existing empirical work on the determinants of cartel organizational form. See, however, the empirical analysis of American shipping cartels that Deltas, Serfes and Sicotte (1999) present. They find that some cartels used simple price-fixing agreements while other cartels were considerably more complex.

Adding the constraints yields the necessary condition $v^1_{HL} - v^1_{LH} = (r - \theta^H)/\delta$, and we may choose the remaining market shares and continuation values to respect the additional conditions in the lemma while satisfying each of the above constraints. ■

Proof of Lemma 5: We posit that IC-Oni_D binds for all i, and substitute in from (3.1) for $U^i(L, L; \mathbf{z})$. We solve a relaxed program:

$$\max_{\substack{q^2_{LL}, q^2_{LH},\, q^2_{HH},\, q^2_{HL} \in [0,1] \\ p_{LH},\, p_{HH},\, p_{HL} \leq r;\ \bar{v}^1_H, \bar{v}^2_H \leq K}} \quad \sum_{j \in \{L,H\}} \eta_j \cdot q^2_{jH} \cdot (p_{jH} - \theta_H) + \delta \bar{v}^2_H + \eta_L \cdot \bar{q}^2_L (\theta_H - \theta_L)$$

$$+\lambda \left[\sum_{k \in \{L,H\}} \eta_k \cdot (1 - q^2_{Hk}) \cdot (p_{Hk} - \theta_H) + \delta \bar{v}^1_H + \eta_L \cdot (1 - \sum_{k \in \{L,H\}} \eta_k q^2_{Lk})(\theta_H - \theta_L) - u^1 \right]$$

$$+\psi^1 \cdot \left[\sum_{k \in \{L,H\}} \eta_k q^2_{Hk} - \sum_{k \in \{L,H\}} \eta_k q^2_{Lk} \right] + \psi^2 \cdot \left[\bar{q}^2_L - \bar{q}^2_H \right]$$

Let λ be the multiplier on firm 1's utility constraint, which is non-negative on the Pareto frontier. The multipliers on the monotonicity constraints are denoted ψ^i, and these are also non-negative. By inspection, it is clearly optimal to set $p_{Hk} = r$, $p_{jH} = r$ and $\bar{v}^i_H = K$; then, differentiating with respect to the market-share variables, we get:

$$\frac{\partial}{\partial q^2_{LL}} : \ (1 - \lambda) \cdot \eta^2_L (\theta_H - \theta_L) - \psi^1 \eta_L + \psi^2 \eta_L; \ \frac{\partial}{\partial q^2_{HH}} : \ (1 - \lambda)[\eta_H (r - \theta_H)] + \psi^1 \eta_H - \psi^2 \eta_H$$

$$\frac{\partial}{\partial q^2_{LH}} : \ \eta_L (r - \theta_H) - \lambda \eta_L \eta_H (\theta_H - \theta_L) - \psi^1 \eta_H - \psi^2 \eta_L$$

$$\frac{\partial}{\partial q^2_{HL}} : \ \eta_L \eta_H (\theta_H - \theta_L) - \lambda \eta_L (r - \theta_H) + \psi^1 \eta_L + \psi^2 \eta_H$$

(i) Notice first that if we maximize the sum of the firms' utilities ($\lambda = 1$), (3.3) implies that unless $\psi^1 > 0$ or $\psi^2 > 0$, $\frac{\partial}{\partial q^2_{LH}} > 0$ and $\frac{\partial}{\partial q^2_{HL}} < 0$, which implies a boundary solution that (as can be directly verified) is dominated by a symmetric solution with $q^i_{jk} = 1/2$ for all (j, k). Now suppose that we weight the firms evenly ($\lambda = 1$) and consider asymmetric solutions. Still, there will be no productive efficiency unless $\psi^1 + \psi^2 > 0$. Suppose that $\psi^1 > 0$ and $\psi^2 = 0$. Then, the objective is increasing in q^2_{HH} and decreasing in q^2_{LL}, so we take $q^2_{HH} = 1$ and $q^2_{LL} = 0$. But then, firm 2's monotonicity constraint implies $\eta_H + \eta_L q^2_{LH} < \eta_H q^2_{HL}$, a contradiction.

Thus, we have established that the largest joint utility available to the firms is achieved by "pooling," where $\bar{q}^i_L = \bar{q}^i_H$, and that allowing for asymmetric allocations of utility will not improve the sum of utilities. This scheme satisfies all of the constraints in $\mathcal{F}^I_{On}(V)$. So, an upper bound for the sum of utilities is given by $r - E[\theta]$. Now observe that for any $\alpha \in [0, 1]$, we can allocate $\alpha(r - E[\theta])$ to firm 1 and $(1 - \alpha)(r - E[\theta])$ to firm 2 by simply changing the market shares of the firms while maintaining $\bar{q}^i_L = \bar{q}^i_H$. Since this satisfies the on-schedule constraints, the Pareto frontier is given as in the statement of the lemma.

(ii) Under the alternative assumption that (3.3) fails, inspection of the program shows that profits are highest when $q^2_{LH} = 0$ and $q^2_{HL} = 1$. The monotonicity constraints do not bind. At these values, the relaxed program is independent of p_{LH} and p_{HL}. This scheme can be implemented by using pricing efficiency in state (H, H) ($p_{HH} = r$), $v^i_{jk} = K$ for all i, j, k, and productive efficiency. The truth-telling constraints can be satisfied as follows: find $\hat{p} < r$ to be used by all low-cost types. With $q^i_{LL} = q^i_{HH} = 1/2$, truth-telling by a high-cost firm requires:

$$\frac{1}{2}\eta_H(r-\theta_H) = (\eta_H + \frac{1}{2}\eta_L)(\hat{p}-\theta_H)$$

yielding the price stated in the lemma. It is now direct to derive the utility frontier. ■

Proof of Proposition 1: The formal program for determining $d^F(\lambda)$ is given by:

$$d^F(\lambda) \equiv \arg\min_{\delta\in[0,1]} \delta$$

$$\text{s.t. } \mathbf{z} \in \mathcal{Z}(V);\ \text{for all } i,j,k,\ \text{IC-On}i_D \text{ and IC-On}i_U \text{ hold;}$$
$$(x,y) = (\bar{U}^1(\mathbf{z}), \bar{U}^2(\mathbf{z}));\ x+y = 2\pi^{FB}/(1-\delta);$$
$$x \le v^i_{jk} \le y;\ \text{IC-Off}i_{jk} \text{ holds, letting } \underline{v}^i = \lambda\pi^{NE}/(1-\delta).$$

To determine δ^{FB}, which is an upper bound on $d^F(\lambda)$ that holds for all parameter values, we solve a set of equations. Consider the following system (imposing pricing efficiency, productive efficiency, and $v^2_{jk} = 2\pi^{FB}/(1-\delta) - v^1_{jk}$ for all (j,k)): IC-On1$_D$, IC-On2$_D$, $\bar{U}^1 = x$, $\bar{U}^2 = y$, $v^1_{LH} = x$, $v^1_{HH} = x$, $q^1_{HH} = 0$, and $q^1_{LL} = 0$. It can be verified that under our assumption that $\eta_L > 1/2$, $v^1_{LH} < v^1_{LL}$. In particular, $v^1_{LL} - v^1_{LH} = \frac{2\eta_L - 1}{\eta_L}\cdot\frac{r-\theta_H}{\delta}$. Since $\frac{2\eta_L-1}{\eta_L} \in (0,1)$, this implies $v^1_{LH} < v^1_{LL} < v^1_{HL}$, where the latter inequality follows since the downward on-schedule constraints imply that $v^1_{HL} = v^1_{LH} + \frac{r-\theta_H}{\delta}$. It remains to verify that given our solutions to these equations, $v^1_{HL} = x + \frac{r-\theta_H}{\delta} < y$. We can compute:

$$y - x = \frac{\eta_L^2(\theta_H - \theta_L) - (r-\theta_H)\ (2\eta_L - 1)}{1-\delta}.$$

This expression is always positive under our restrictions $\eta_L > 1/2$ and (4.1). δ^{Fon}, as stated in the text, solves $y - x = \frac{r-\theta_H}{\delta}$. It can be verified that $y - x > \frac{r-\theta_H}{\delta}$ for all $\delta > \delta^{Fon}$.

Consider now the off-schedule constraints. We observe first that IC-Off1$^I_{LH}$ is slack, as is IC-Off-M1. Further, using the implementation described above, IC-Off1$^I_{HH}$ implies IC-Off1$^I_{HL}$. We then substitute in the values for q^1_{jj} and v^1_{jj} computed above, and verify (using tedious algebraic manipulation) that IC-Off1$^I_{LL}$ binds and IC-Off1$^I_{HH}$ is slack when $\delta \ge \frac{\eta_L + \kappa(1-\eta_L)}{\eta_L + \kappa(1-\eta_L) + \eta_L^2\kappa}$ and (4.1) holds. Finally, since for the endpoints of the interval, we have described a policy vector that meets all of our constraints and uses as continuation values other values on the same interval, we can then construct the remainder of the line segment using convex combinations of policy vectors to implement convex combinations of equilibrium values. This is possible because, when pricing efficiency is imposed, the constraints and utilities are linear in market shares and continuation values. Thus, we have constructed a self-generating set of equilibrium values with first-best profits to the cartel. ■

Proof of Lemma 6: To implement (x,y), let $p_{jk} = r$, $v_{LH} = v_{LL} = (x,y)$, $v_{HH} = v_{HL}$, and $q^1_{HL} = q^1_{HH} = 0$. Further, set $q^1_{LH} = q^1_{LL} = \frac{\delta(\kappa+\eta_L^2)}{1+\kappa-\delta\eta_H}$. Finally, use the static Nash equilibrium as the off-schedule threat point. Notice first that IC-On2$_D$ and IC-On2$_U$ both hold with equality given these values. Given the specified market shares and off-equilibrium-path threat points, the v^1_{LH} that satisfies IC-Off1$^I_{LL}$ and IC-Off1$^I_{LH}$ is uniquely determined. Next, it can be shown that IC-On1$_D$ holds with these parameter values if and only if $v^1_{HL} - v^1_{LH} = q^1_{LH}(r-\theta_H)/\delta$. Thus, it remains to verify that this distance is feasible using $v^1_{HL} \le y$, where y is determined as firm 2's profit in this scheme. Given pricing efficiency and the specified level of productive efficiency for the cartel, and since all continuation values lie on a line with slope -1 through (x, y) , it is possible to compute the sum of firm profits at (x, y), $K = x + y$, as a function of the exogenous parameters. Then, since $v^1_{LH} = x$, $v^1_{HL} \leq y$ holds if and only if $v^1_{LH} + (q^1_{LH} + q^1_{LL})(r - \theta_H)/(2\delta)$ $\leq K - v^1_{LH}$. It can be verified that this constraint becomes relaxed as δ increases. Substituting and solving, the constraint binds for $\delta = \delta^{lin} \equiv$

$$\frac{\eta_L^2 - \kappa(\kappa + \eta_H) + \sqrt{(\kappa(1+\kappa) - \eta_L(\eta_L + \kappa))^2 + 4\kappa(1+\kappa)(\kappa^2 + \eta_L^2(1+\eta_L^2) + \kappa\eta_L(1+2\eta_L))}}{2(\kappa^2 + \eta_L^2 + \eta_L^4 + \kappa\eta_L(1+2\eta_L))}$$

■

Proof of Proposition 2: (i) The symmetric point of the Pareto frontier of $\tilde{T}(V)$ can be implemented with $q^i_{jj} = 1/2$ and $v^1_{jj} = v^2_{jj}$. Before beginning, we observe that we can take $p_{jj} > \theta_j$ without loss of generality. To see why, observe that if $\theta_j - p_{jj} > 0$, we can raise p_{jj} by ε and lower v^i_{jj} by $\varepsilon/(2\delta)$ until we arrive at $\hat{p}_{jj}$ and $\hat{v}_{jj}$, where $\theta_j = \hat{p}_{jj}$, without affecting any utilities or incentive constraints (since an optimal off-schedule deviation would ensure zero market-share in state jj). To see that the resulting $\hat{v}^i_{jj}$ is feasible, observe that given market share of 1/2, our assumption that IC-Offi^I_{jj} is slack implies that $\frac{1}{2}(\theta_j - p_{jj}) < \delta(v^i_{jj} - \underline{v}^i)$; since the adjustments preserve this inequality, the new continuation value $\hat{v}^i_{jj}$ must satisfy $\hat{v}^i_{jj} \geq \underline{v}^i$. Since the set of feasible continuation values is convex and symmetric, it is feasible. Finally, if $\theta_j - p_{jj} = 0$, we may employ a similar adjustment, unless $v^i_{jj} = \underline{v}^i$. But in this case (4.2) is violated.

Starting from this point, our approach is to implement an alternative utility pair, with no efficiency loss, in which $\bar{U}^1$ is decreased and $\bar{U}^2$ is increased. We define two perturbations. In Perturbation 1, we lower q^1_{HH} by $\varepsilon/((p_{HH} - \theta_L)\eta_H)$ and lower q^1_{LL} by $\varepsilon/((p_{LL} - \theta_L)\eta_L)$. For each firm i, IC-Oni_U is unaltered by this perturbation. In Perturbation 2, we lower q^1_{HH} by $\varepsilon/((p_{HH} - \theta_H)\eta_H)$ and lower q^1_{LL} by $\varepsilon/((p_{LL} - \theta_H)\eta_L)$. For each firm i, this perturbation leaves unaltered IC-Oni_D. Both perturbations lower $\bar{U}^1$ and increase $\bar{U}^2$.

There are several cases to consider. Suppose first that, for a given $\psi \in \{U, D\}$, IC-Oni_ψ is slack for each i. If $\psi = U$, we use Perturbation 2 to engineer the desired utility transfer without violating on-schedule incentive constraints. Likewise, if $\psi = D$, we use Perturbation 1. Next, we modify the argument for the case where the on-schedule constraints are slack in different directions. First, take the case where IC-On1_D and IC-On2_U are slack. If $p_{LL} \leq p_{HH}$, we use Perturbation 1, which relaxes IC-On2_D by $(p_{HH} - \theta_H)/(p_{HH} - \theta_L) - (p_{LL} - \theta_H)/(p_{LL} - \theta_L)$, which is positive by our assumption that $p_{LL} \leq p_{HH}$. If $p_{LL} \geq p_{HH}$, we use Perturbation 2. This relaxes IC-On1_U by $(p_{HH} - \theta_L)/(p_{HH} - \theta_H) - (p_{LL} - \theta_L)/(p_{LL} - \theta_H)$, which is positive. Similarly, in the second case, where IC-On1_U and IC-On2_D are slack, we proceed as follows: if $p_{LL} \leq p_{HH}$, we use Perturbation 2, which relaxes IC-On2_U by $(p_{LL} - \theta_L)/(p_{LL} - \theta_H) - (p_{HH} - \theta_L)/(p_{HH} - \theta_H)$, which is positive; and if $p_{LL} \geq p_{HH}$, we use Perturbation 1, which relaxes IC-On1_D by $(p_{LL} - \theta_H)/(p_{LL} - \theta_L) - (p_{HH} - \theta_H)/(p_{HH} - \theta_L)$, which is positive.

(ii) Suppose that (a) and the first part of (b) fail: IC-Off1^I_{LL} and IC-Off1^I_{HH} are slack, and $v^1_{LL} > x$ and $v^1_{HH} > x$. Then, lower v^1_{HH} by ε/η_H and lower v^1_{LL} by ε/η_L, and raise the corresponding values for firm 2 by the same amount (this is possible because $v^1_{jj} > x$ and the set of available continuation values is convex). This does not affect any on-schedule constraints, relaxes firm 2's off-schedule constraints, and decreases $\bar{U}^1$ and increases $\bar{U}^2$ with no efficiency loss, contradicting the hypothesis that (x, y) is the end point of the region with slope equal to -1. Now consider the case where (a) and the second part of (b) fail: IC-Off1^I_{LL} and IC-Off1^I_{HH} are slack, $q^1_{LL} > 0$ and $q^1_{HH} > 0$, and $p_{jj} > \theta_j$ for each j. We show below in Proposition 6 that under the assumptions of this proposition, $\bar{q}^i_L > \bar{q}^i_H$, which implies that for each i, one of IC-Oni_D and

IC-Oni_U is slack. We may now apply the algorithm used in the proof of Part (i) to implement a utility pair that yields lower (higher) utility for firm 1 (2), without efficiency loss, contradicting the hypothesis that (x, y) is the endpoint of a region with slope equal to -1. ■

Proof of Proposition 3: (i) If each IC-Oni_U is slack, prices are efficient, and $q^1_{LH} = 1$, then IC-Off1^I_{LH} and IC-Off-M1 are slack, and we can decrease q^1_{LH} and give the market share to firm 2 without violating any on-schedule constraints. But this makes firm 1 worse off and firm 2 better off, violating the hypothesis that the scheme implements a corner of v^I_N.

(ii) Suppose that (a) and the first part of (b) fail: IC-Off1^I_{LL} and IC-Off1^I_{HH} are slack, and $v^1_{LL} > v^{1I}_N$ and $v^1_{HH} > v^{1I}_N$. Then, for ε_1 small enough, there exists an $\varepsilon_2 > 0$ such that it is possible to lower v^1_{HH} by ε_1/η_H and lower v^1_{LL} by ε_1/η_L, and raise v^2_{HH} by ε_2/η_H and raise v^2_{LL} by ε_2/η_L (this is possible because the set of available continuation values is convex and by the definition of v^I_N as the north corner of the Pareto frontier). This does not upset any on-schedule constraints, and makes firm 1 worse off and firm 2 better off. There is potentially an efficiency loss, however. Next, we consider the case where (a) and the second part of (b) fail. We may then argue as in the proof of Proposition 2 and arrive at a contradiction. ■

Proof of Proposition 4: The proof proceeds in a series of lemmas. Part (i) follows by Lemma 8, and part (ii) follows by Lemma 9 ■

Lemma 7. *Consider a scheme* $(\mathbf{p}, \mathbf{q}, \mathbf{v})$. *(T1j) If we subtract* $\eta_L\varepsilon$ *from* v^1_{jH} *and add* $\eta_H\varepsilon$ *to* $v^1_{jL}, \bar{U}^1$*, IC-On1*$_D$ *and IC-On1*$_U$ *are unaffected. (T2j) If we add* $\eta_H\varepsilon$ *to* v^2_{Lj} *and subtract* $\eta_L\varepsilon$ *from* $v^2_{Hj}, \bar{U}^2$*, IC-On2*$_D$ *and IC-On2*$_U$ *are unaffected.*

Lemma 8. *If* $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ *satisfies* $\mathcal{F}^I_{On}(V)$ *and* $v^1_N > v^1_{jk} > v^1_E$ *for all* (j, k)*, then if* $v^2_{jk} < f(v^1_{jk})$ *for any* (j, k)*, this scheme is Pareto dominated by another scheme that satisfies* $\mathcal{F}^I_{On}(V)$*, has all continuation values on the Pareto frontier of* V *and uses the same prices.*

Proof. Suppose that for some j, $v^2_{jH} < f(v^1_{jH})$ and $v^2_{jL} < f(v^1_{jL})$. Then, we can hold fixed firm 1's continuation values and increase v^2_{jH} and v^2_{jL} by the same amount without affecting IC-On2, thus increasing $\bar{U}^2$. Then, suppose that $v^2_{jH} = f(v^1_{jH})$ and $v^2_{jL} < f(v^1_{jL})$. Then, apply Lemma 7, (T1j), so that neither continuation value is on the frontier. Then, both v^2_{jH} and v^2_{jL} can be increased, again increasing $\bar{U}^2$ without affecting $\bar{U}^1$. The other case is analogous. ■

Lemma 9. *Suppose that* $(\mathbf{p}, \mathbf{q}, \mathbf{v})$ *implements a Pareto efficient point in* $\tilde{T}^I_{On}(V)$*. For all* (j, k)*, if* $v^1_N > v^1_{jk} > v^1_E$*, then* $p_{jk} = r$*.*

Proof. Suppose $p_{jk} < r$. Then we can increase p_{jk} by ε and decrease v^1_{jk} and v^2_{jk} by $\frac{1}{\delta}\varepsilon$ without affecting payoffs or on-schedule IC constraints. Then, we can improve utility by returning the continuation values to the frontier as in Lemma 8. ■

Proof of Proposition 5: Lemma 5 establishes that when the continuation value set has the shape $\{(u^1, u^2) : u^i \leq K\}$, the total cartel profits go down when firms use Pareto inefficient continuation values or prices. This logic can be applied directly here, observing that we are maximizing total profits because utility can be transferred across the firms (as in Proposition 2) under the conditions stated in the proposition. ■

Proof of Proposition 6: First, suppose $q^1_{LH} < 1$. Add $\frac{\delta}{p_{LH}-\theta_L}\varepsilon$ to q^1_{LH} and subtract ε from v^1_{LH}. If v_{LH} is on the Pareto frontier and $\Delta^-_\varepsilon f(v^1_{LH}) > -1$, move v^2_{LH} along the frontier of V. Otherwise, raise v^2_{LH} by ε (this is possible by convexity of the set of feasible continuation values, and since satisfaction of (4.3) implies $v^1_{LH} > v^1_N$). This does not affect $\bar{U}^1$. Consider now the effect on the interim expected utility of both firms: $U^1(H,H;\mathbf{z})$, $U^1(H,L;\mathbf{z})$, and $U^1(L,L;\mathbf{z})$ are unchanged; and $U^1(L,H;\mathbf{z})$ decreases. $U^2(L,H;\mathbf{z})$ and $U^2(L,L;\mathbf{z})$ are unchanged. $U^2(H,H;\mathbf{z})$ goes up if $-(p_{LH}-\theta_H)/(p_{LH}-\theta_L)-\max(-1,\Delta_\varepsilon^- f(v_{LH}^1))>0$, which when rearranged gives (4.3). $U^2(H,L;\mathbf{z})$ goes down if v_{LH}^2 increases by no more than ε, which it does by construction. Thus, all of the on-schedule incentive-compatibility constraints are relaxed. Finally, none of firm 1's off-schedule constraints are affected by this shift, and firm 2's off-schedule constraints are relaxed. To see the result for q_{HL}^1, we perform an analogous trick, subtracting $\frac{\delta}{p_{HL}-\theta_H}\varepsilon$ from q_{HL}^1 and adding ε to v_{HL}^1, and noting that satisfaction of (4.4) implies that $v_{HL}^1 < v_E^1$ (recalling that in the definition of f, we specified a large negative slope for f when $v_{jk}^1 \geq v_E^1$). ■

Proof of Proposition 7: Under the assumptions of the proposition, utility is fully transferable across the firms, and we can simply maximize the sum of firm utilities. Doing so leads to a symmetric scheme across states (L,H) and (H,L), with one firm being favored over another in states (H,H) and (L,L), if at all. Now imagine lowering q_{LH}^1 and raising q_{HL}^1 by ε, and then adjusting v_{LH}^1 and v_{HL}^2 upward by ζ until both firms' downward IC constraints bind again. The opponents' continuation values are moved along the frontier. Solving for ζ, we obtain:

$$\frac{1}{1-\eta_L(1-\Delta_\varepsilon^+ f(v_{LH}^1))}\frac{\varepsilon(p_{LH}-\theta_H)}{\delta}.$$

It can be shown using algebra that the first inequality in the statement of the proposition is necessary and sufficient for this change to lower total firm profit. The second inequality is necessary and sufficient for the firms' joint profit to decrease if we reverse this change.

Proof of Proposition 8: The linear self-generating set of equilibria constructed in the proof of Proposition 1 implements the endpoints of the segment, (x,y) and (y,x), using schemes that have market shares $q_{LH}^1=1$, $q_{jk}^1=0$ for all other (j,k). Consider a scheme whereby firm 1 chooses $\rho^1(\theta_H)=r$ and $\rho^1(\theta_L)=r-2\Delta$, and $\rho^2(\theta_H)=r-\Delta$ and $\rho^2(\theta_L)=r-3\Delta$. Using this scheme, the market shares are assigned appropriately. Further, each firm's announced price differs by state, so that continuation values can be contingent purely on prices. Thus, communication is not required to implement the scheme. Since firms can draw from a convex set of continuation values, all continuation values in between (x,y) and (y,x) are available to the firms, and the linear set is self-generating. To compute the critical discount factor, we follow Section 4.1 and parameterize the worst available punishment using λ to represent the fraction of the static Nash equilibrium profits that can be sustained as a punishment. Letting $\kappa=(r-\theta_H)/(\theta_H-\theta_L)$ taking the limit as $\Delta\to 0$, using the appropriate off-schedule constraints for uninformative communication, and setting $\lambda=1$, we compute the following bound:

$$\delta^{NC}=\max\left(\delta^{Fon},\frac{\eta_L(2-\eta_L)+\sqrt{(2-\eta_L)^2\eta_L^2+8\kappa(\kappa+\eta_L)}}{4(\kappa+\eta_L)}\right).$$

■

9. References

Abreu, Dilip, David Pearce and Ennio Stacchetti, "Optimal Cartel Equilibria with Imperfect Monitoring", *Journal of Economic Theory*, 39.1, June 1986, pp. 251-69.

Abreu, Dilip, David Pearce and Ennio Stacchetti, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58.5, September 1990, pp. 1041-63.

Aoyagi, Masaki, "Bid Rotation in a Model of Repeated Auctions," Mimeo, 1998, University of Pittsburgh.

Athey, Susan and Kyle Bagwell, 1999, "Optimal Collusion with Private Information," MIT Working Paper no. 99-12.

Athey, Susan, Kyle Bagwell, and Chris Sanchirico, 1998, "Collusion and Price Rigidity," MIT Working Paper no. 98-23.

Atkeson, Andrew and Robert E. Lucas, "On Efficient Distribution with Private Information," *Review of Economic Studies*, 59.3, July 1992, pp. 427-53.

Bork, Robert H., "The Rule of Reason and The Per Se Concept: Price Fixing and Market Division," *Yale Law Journal*, 74.5, April 1965, 775-847.

Bork, Robert H., "The Rule of Reason and The Per Se Concept: Price Fixing and Market Division," *Yale Law Journal*, 75.3, January 1966, 373-475.

Business Week, "Justice Cartel Crackdown," July 27, 1998, pp. 50-51.

Caves, Richard and Michael Porter, "Market Structure, Oligopoly and Stability of Market Shares," *Journal of Industrial Economics*, 1978, 27, 289-312.

Cole, Harold L., and Narayana Kocherlakota, "Dynamic Games with Hidden Actions and Hidden States," Federal Reserve Bank of Minneapolis Research Department Staff Report 254, 1998.

Comanor, W.S. and M.A. Schankerman, "Identical Bids and Cartel Behavior," *Bell Journal of Economics*, 7, 1976, 281-6.

Compte, Olivier, "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66.3, May 1998, 597-626.

Cramton, Peter C. and Thomas R. Palfrey, "Cartel Enforcement with Uncertainty about Costs." *International Economic Review*; 31 (1) February 1990, pp. 17-47.

d'Aspremont, Claude and L.A. Gerard-Varet, "Incentives and Incomplete Information," *Journal of Public Economics*, 11, 1979, 25-45.

Deltas, George, Konstantinos Serfes and Richard Sicotte, "American Shipping Cartels in the Pre-World War I Era," March 1999.

Eckard, E. W., Jr., "Advertising, Competition and Market Share Instability," *Journal of Business*, 60.4, October 1987, 539-52.

Fudenberg, Drew, David J. Levine and Eric Maskin, "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62.5 September 1994, pp. 997-1039.

Genesove, David and Wallace Mullin, "Narrative Evidence on the Dynamics of Collusion: The Sugar Institute Case," Mimeo, 1998, Michigan State University.

Genesove, David and Wallace Mullin, "The Sugar Institute Learns to Organize Information Exchange," in *Learning by Doing in Markets, Firms, and Countries*, Naomi Lamoreux, Daniel M.G. Raff, and Peter Temin, eds., University of Chicago Press, 1999.

Green, Edward J., "Lending and the Smoothing of Uninsurable Income." in Prescott, Edward C. and Neil Wallace, ed. *Contractual Arrangements for Intertemporal Trade.* Minnesota Studies in Macroeconomics series, vol. 1 Minneapolis: University of Minnesota Press 1987, pp. 3-25.

Green, Edward J. and Robert H. Porter, "Noncooperative Collusion under Imperfect Price Information," *Econometrica*; 52.1 January 1984, pp. 87-100.

Griffen, James M., "An Inside Look at the Cartel at Work: Common Characteristics of International Cartels," address before the American Bar Association, April 6, 2000. (www.usdoj.gov/atr/public/speeches/4489)

Holmstrom, Bengt and David Kreps, "Promises," mimeo, MIT, 1996.

Judd, Kenneth, and James Conklin, "Computing Supergame Equilibria," Mimeo, Hoover Institution, Stanford University, 1995.

Kandori, Michihiro and Hitoshi Matshushima, "Private Observation, Communication and Collusion," *Econometrica,* 66.3, May 1998, 627-52.

Kihlstrom, Richard and Xavier Vives, "Collusion by Asymmetrically Informed Firms," *Journal of Economics and Management Strategy*, 1 (2) 1992, pp. 371-96.

Kuhn, Kai-Uwe and Xavier Vives, "Information Exchanges Among Firms and Their Impact on Competition," mimeo, Barcelona: Institut d'Analisi Economica, June 1994.

Levinsohn, James, "Competition Policy and International Trade," in *Fair Trade and Harmonization* , J. Bhagwati and R. Hudec, eds., MIT Press, pp. 329-356, 1995.

McAfee, Preston and John McMillan, "Bidding Rings," *American Economic Review*; 82.3, June 1992, pp. 579-599.

McMillan, John , "Dango: Japan's Price-Fixing Conspiracies," *Economics and Politics*, November 1991, (3): 201-218.

Myerson, Roger, "Multistage Games with Communication." *Econometrica* 54 (2), March 1986, 323-358.

Phelan, Christopher, and Robert M. Townsend, "Computing Multi-period, Information-Constrained Optima." *Review of Economic Studies* 58 (5), 1991, pp. 853-81.

Radner, Roy, "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship," *Econometrica*, 49, 1981, pp. 1127-1148.

Roberts, Kevin, "Cartel Behavior and Adverse Selection," *Journal of Industrial Economics*, 33, 1985, pp. 401-413.

Shapiro, Carl, "Exchanges of Cost Information in Oligopoly," *Review of Economic Studies*, 53, 1986, pp. 433-446.

Sproul, Michael, "Antitrust and Prices," *Journal of Political Economy*, 101.4, 1993, pp. 741-754.

Stocking, George W. and Myron W. Watkins, *Cartels in Action. Case Studies in International Business Diplomacy*, New York: The Twentieth Century Fund, 1946.

Telser, Lester, "Advertising and Competition," *Journal of Political Economy*, 72.6, December 1964, 537-62.

Vives, Xavier, "Duopoly Information Equilibrium: Cournot and Bertrand," *Journal of Economic Theory*, 34, 1984, 71-94.

Wang, Cheng, "Dynamic Insurance with Private Information and Balanced Budgets," *Review of Economic Studies* ; 62.4, October 1995, pp. 577-95.

Ziv, Amir, "Information Sharing in Oligopoly: The Truth-Telling Problem," *RAND Journal of Economics*, 24.3, Autumn 1993, 455-65.

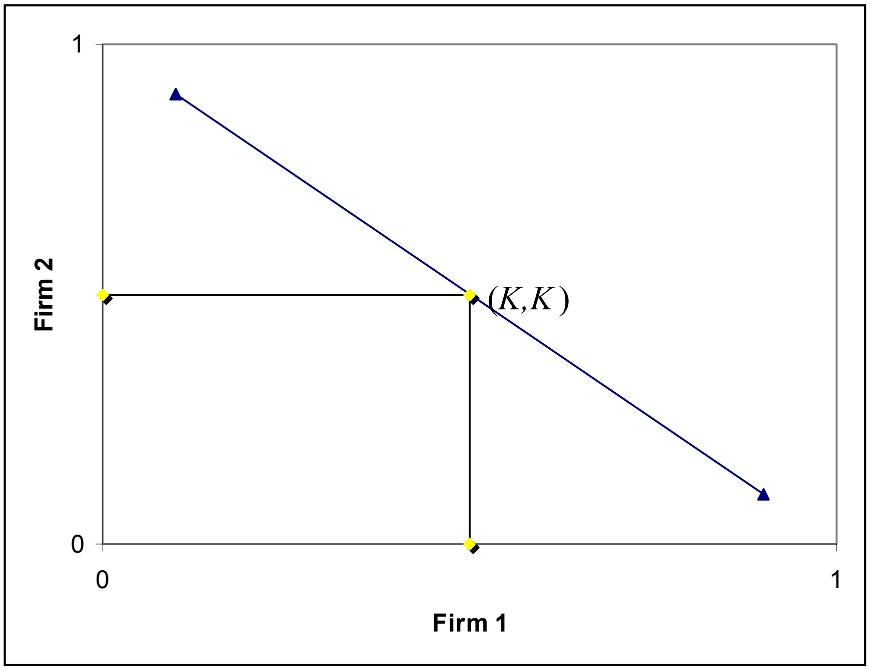

Figure 1: Sets of Feasible Continuation Values for Benchmark Cases

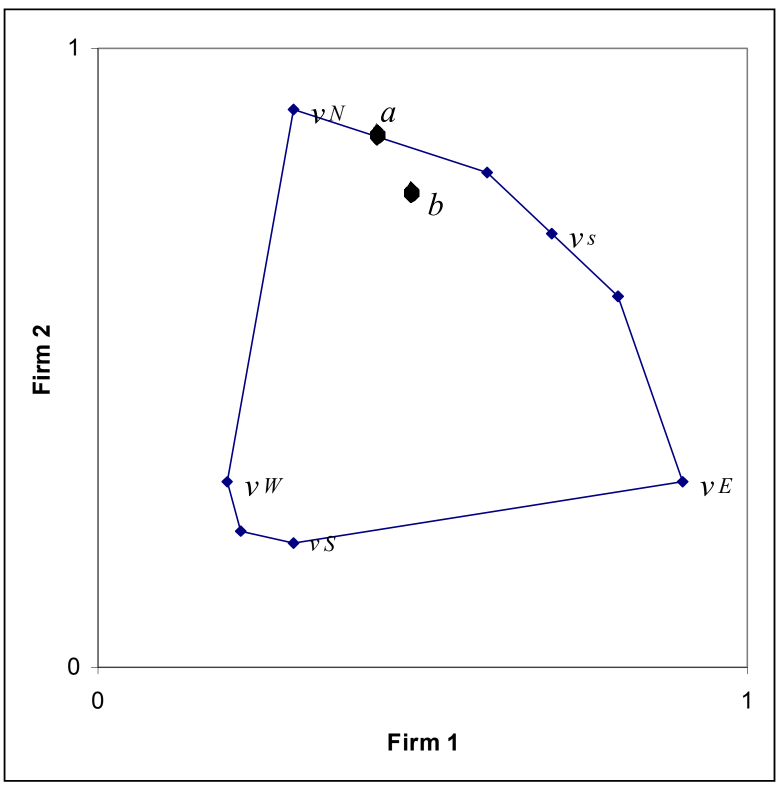

Figure 2: A convex set of continuation values. Point *a* has "future inefficiency," while point *b* is a Pareto inefficient continuation value.

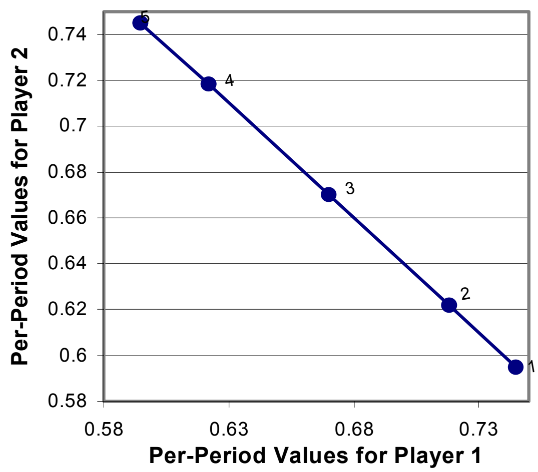

Figure 3a: A equilibrium set which is self-generating (together with the static Nash equilibrium) and achieves first-best profits. Parameters: r=2.5, θ_H=2, θ_L=1, Pr(θ_L)=.6, δ=.769.

		Transitions (States)				Player 1 Market Shares			
State.	Values	(L,L)	(L,H)	(H,L)	(H,H)	(L,L)	(L,H)	(H,L)	(H,H)
1	(.745, .595)	2	5	1	1	0.848	1.000	0.000	1.000
2	(.718, .622)	2	5	1	1	0.679	1.000	0.000	0.733
3	(.670, .670)	3	5	1	3	0.500	1.000	0.000	0.500
4	(.622, .718)	4	5	1	5	0.331	1.000	0.000	0.267
5	(.595, .745)	4	5	1	5	0.152	1.000	0.000	0.000

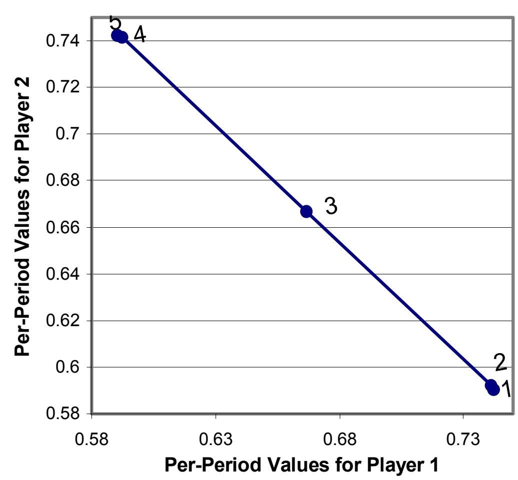

Figure 3b: An equilibrium set which is self-generating (together with the static Nash equilibrium) and achieves close to first-best profits. Parameters: r=2.5, θ_H=2, θ_L=1, Pr(θ_L)=.6, δ=.769.

		Transitions (States, Probabilities)				Player 1 Market Shares			
State	Values	(L,L)	(L,H)	(H,L)	(H,H)	(L,L)	(L,H)	(H,L)	(H,H)
1	(.742,.590)	(2,4) (.83,.17)	(2,4) (.02,.98)	1	2	0.838	.984	0.000	1.000
2	(.741, .592)	(2,4) (.82,.18)	4	1	2	0.841	1.000	0.000	1.000
3	(.667, .667)	3	(4,5) (.5,.5)	(1,2) (.5,.5)	3	0.500	1.000	0.000	0.500
4	(.592, .741)	(2,4) (.18,.82)	5	2	4	0.159	1.000	0.000	0.000
5	(.590, .742)	(2,4) (.17,.83)	5	(2,4) (.98,.02)	4	0.162	1.000	0.016	0.000

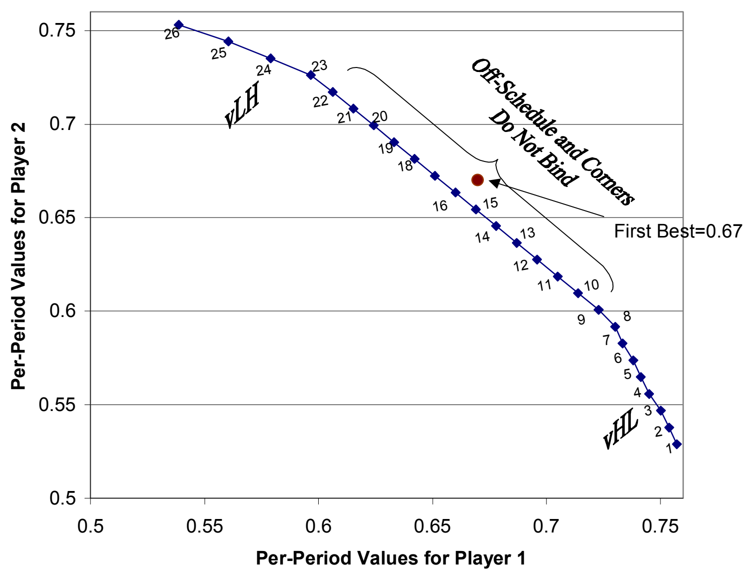

Figure 4: An equilibrium set which is self-generating (together with the static Nash equilibrium), for firms of moderate patience. Parameter values: r=2.5, θ_H=2, θ_L=1, Pr(θ_L)=.6, δ=.74.

	Transitions				Player 1 Market Shares			
State/Realiz.	**(L,L)**	**(L,H)**	**(H,L)**	**(H,H)**	**(L,L)**	**(L,H)**	**(H,L)**	**(H,H)**
1	7	12	2	13	0.649	0.983	0.556	1.000
2	10	14	3	15	0.700	0.998	0.529	1.000
3	9	15	1	13	0.675	0.994	0.407	1.000
4	7	17	3	11	0.648	0.994	0.345	1.000
5	7	18	1	10	0.641	0.996	0.231	1.000
6	8	20	2	10	0.659	1.000	0.184	1.000
7	10	21	1	11	0.691	0.995	0.104	1.000
8	8	23	1	8	0.660	0.999	0.003	1.000
9	12	24	3	12	0.733	0.989	0.000	0.995
10	13	24	3	13	0.704	0.989	0.000	0.926
11	13	24	3	13	0.652	0.989	0.000	0.847
12	13	24	3	13	0.600	0.989	0.000	0.769
13	13	24	3	13	0.548	0.989	0.000	0.691
14	13	24	3	13	0.495	0.989	0.000	0.612
15	13	24	3	13	0.443	0.989	0.000	0.534
16	13	24	3	13	0.391	0.989	0.000	0.456
17	13	24	3	13	0.339	0.989	0.000	0.377
18	13	24	3	13	0.286	0.989	0.000	0.299
19	13	24	3	13	0.234	0.989	0.000	0.221
20	13	24	3	13	0.182	0.989	0.000	0.142
21	18	24	3	18	0.246	0.989	0.000	0.110
22	20	25	5	20	0.271	0.998	0.000	0.057
23	21	26	8	21	0.288	0.971	0.000	0.022
24	22	26	11	21	0.305	0.825	0.000	0.037
25	23	26	14	20	0.323	0.672	0.000	0.012
26	21	26	18	18	0.286	0.468	0.000	0.110

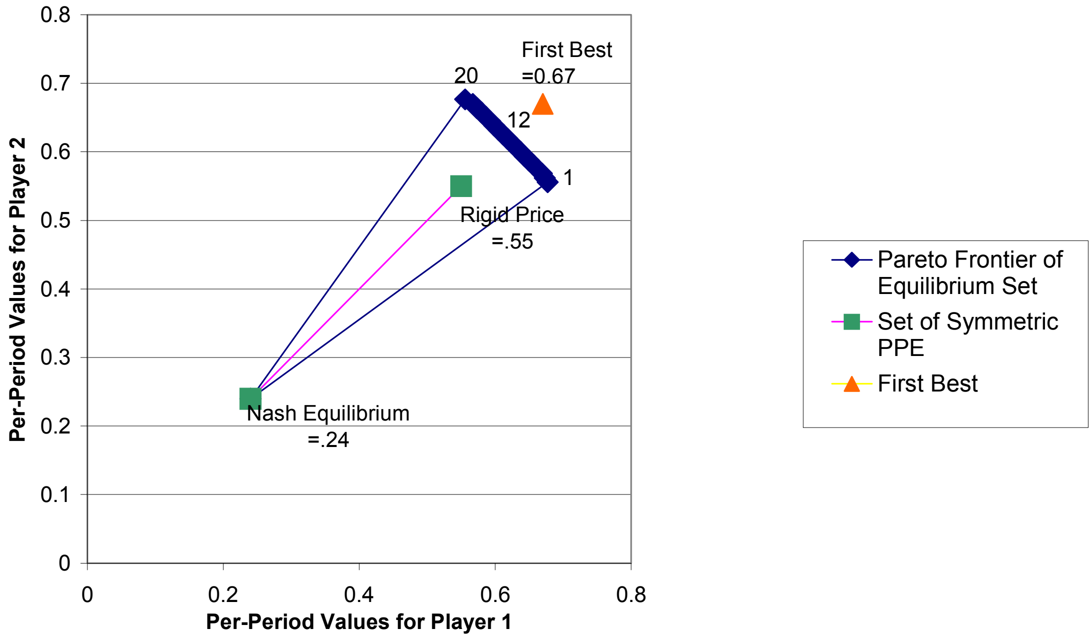

Figure 5: A self-generating equilibrium set. The discount factor is equal to the critical discount factor for supporting the best Symmetric equilibrium (which entails a rigid price of r for all cost types). Parameter values: r=2.5, θ_H=2, θ_L=1, Pr(θ_L)=.6, δ=.707.

	Transitions				Player 1 Market Shares			
State/Realiz.	**(L,L)**	**(L,H)**	**(H,L)**	**(H,H)**	**(L,L)**	**(L,H)**	**(H,L)**	**(H,H)**
1	3	18	1	18	0.529	0.996	0.485	1.000
2	3	19	1	18	0.518	0.994	0.453	1.000
3	3	20	1	18	0.511	0.992	0.414	1.000
4	3	20	1	16	0.527	0.918	0.340	1.000
5	3	20	1	16	0.516	0.885	0.307	1.000
6	6	20	1	18	0.553	0.843	0.265	1.000
7	3	20	1	14	0.521	0.778	0.201	1.000
8	3	20	1	13	0.524	0.725	0.147	1.000
9	5	20	1	14	0.550	0.679	0.101	1.000
10	5	20	1	14	0.539	0.646	0.068	1.000
11	5	20	1	14	0.528	0.613	0.035	1.000
12	5	20	1	9	0.518	0.577	0.000	0.850
13	5	20	1	9	0.485	0.577	0.000	0.800
14	15	20	1	12	0.592	0.577	0.000	0.591
15	5	20	1	9	0.420	0.577	0.000	0.701
16	5	20	1	9	0.387	0.577	0.000	0.652
17	5	20	1	9	0.354	0.577	0.000	0.603
18	10	20	1	12	0.391	0.577	0.000	0.520
19	16	20	1	12	0.442	0.577	0.000	0.319
20	19	20	4	10	0.479	0.497	0.000	0.202

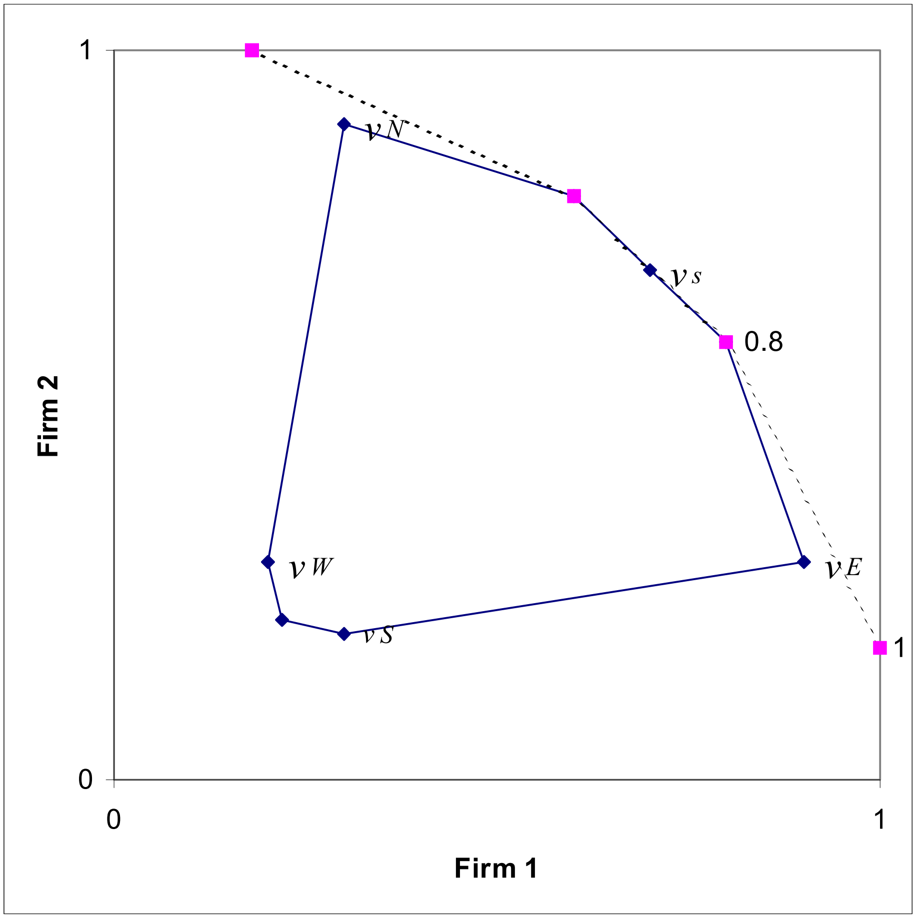

Figure 6: Bribes augment the set of feasible continuation values.