Business stealing in imperfect cartels
PRELIMINARY DRAFT

B. Douglas Bernheim
Stanford University
and
NBER

Erik Madsen
Stanford GSB

November 2013

Abstract

Though economists have made substantial progress toward formulating theories of collusion in industrial cartels that account for important fact patterns, important puzzles remain. Standard models of repeated interaction formalize the observation that cartels keep participants in line through the threat of punishment, but they fail to explain two important factual observations: first, apparently deliberate cheating actually occurs; second, it frequently goes unpunished even when it is detected. We propose a theory of "equilibrium business stealing" in cartels to bridge this gap between theory and observation.

1 Introduction

An important objective of theoretical research in Industrial Organization is to achieve a conceptual understanding of the mechanisms through which actual price-fixing cartels arrive at collusive outcomes. Analyses of strategic models involving repeated interaction have yielded important insights but also leave significant gaps. A serious shortcoming of the simplest such theories is that neither deliberate cheating nor punishment occurs in equilibrium, even though both are observed in practice.[*] Extensions involving imperfect

[*]Examples of punishments have been widely discussed in the literature; see, e.g., Porter (1983), Green and Porter (1984), and Harrington (2006). Instances of cheating are frequently identified in the context of antitrust cases. For example, during litigation associated with the lysine cartel (which operated at least from 1992 through 1995), an Ajinomoto executive testified that cheating was common but that "the range of cheating is not so big . . . they kept their promise about 90 percent. Something like that." [Connor (2001).] With respect to business stealing, he noted that "taking the other's customers happens all the time." [Chicago Tribune, 7/23/1998.]

monitoring (beginning with Green and Porter, 1984) explain why punishments (such as price wars) are observed, but do not account for *deliberate cheating*: in equilibrium, punishments are triggered only by events beyond the control of the cartel members, and not by their intentional choices.[*] Moreover, according to those theories, if cheating did occur and was detected, it would *definitely* trigger punishment. Yet it appears that observed business stealing sometimes goes unpunished: instead of retaliating, cartel members urge each other to recall their common interests and let cooler heads prevail.[†] These unresolved empirical puzzles have important practical implications, in that attorneys for defendant companies often point to evidence of business stealing, and to a purported lack of retaliation, as "proof" that a cartel is ineffective.[‡]

Though intuitively plausible, the possibility that an imperfect but nevertheless effective real-world cartel might exhibit some degree of deliberate business stealing, and that such behavior might sometimes go unpunished despite detection, has as yet found no rigorous theoretical articulation. In this paper, we attempt to bridge this important gap between theory and observation by constructing a theory of equilibrium business stealing in imperfectly effective cartels. We formulate a model in which firms have natural advantages with respect to serving particular market segments and must make "relationship investments" (e.g., incur bid-preparation costs) to do business with specific customers. For intermediate discount factors, some collusion is feasible but perfect collusion is not. Our agenda is to study the the properties of imperfectly collusive equilibria in those settings.

Beginning with the case of two firms, our first main result shows that, under reasonably general conditions, imperfectly collusive equilibria *necessarily* involve business stealing, in

[*]To be clear, the theory of imperfect monitoring can in principle account for *unintended* cheating – e.g., apparent defections from collusive agreements attributable to "rogue employees" who are not involved in the conspiracy. In particular, one could construct a model with imperfectly controlled sales personnel whom, in equilibrium, each firm would instruct to quote some collusive prices (which means *deliberate* cheating would not occur). However, analogously to Green-Porter (1984), "rogue" salespeople would periodically grant price concessions (e.g., with the object of enhancing their own compensation given their false understanding of their employer's objectives). Because other firms would be unable to distinguish between actual and rogue defections, all defections would have to occasionally trigger punishments.

[†]Such sentiments are well captured in remarks made by ADM executive Terrance Wilson during a March 10, 1994 meeting of lysine executives: "[Lysine buyers] are not your friend. They are not my friend. And we gotta have 'em. Thank God we gotta have 'em, but they are not my friends. You're my friend. I wanna be closer to you than I am to any customer... [L]et's put the prices on the board. Let's all agree that's what we're gonna do and then walk out of here and do it." [Justice Department.]

[‡]This strategy was employed, for instance, by defense attorneys during criminal prosecution of several top ADM executives following discovery of the lysine cartel. [Connor (2000).] News coverage of the trial described the defense's strategy as follows: "Top executives of Archer Daniels Midland Co. 'busted up a longstanding Asian cartel,' introducing 'fierce competition' into the market for a livestock-feed additive, a defense attorney said... [ADM] was 'stealing customers' and 'undercutting competitors' at the time prosecutors say it was carving up the lysine business... [C]ompetitors lied to each other routinely, he said. 'This is not Business Ethics 101. This is how you deal with the real world. You have to mislead the competition.' " [Chicago Tribune, 9/10/1998.]

the sense that no firm wins any customer with certainty (even though the model is deterministic). Our second main result demonstrates that the *best* collusive equilibria within an important class have the following properties. First, the cartel (in effect) attempts to divide the market according to the firms' relative advantages, and to establish a collusive price, which is necessarily below the monopoly price. Second, each firm sometimes attempts to extract even more profits from its "home market" by charging more than the agreed-upon price. Third, when a firm engages in this "surpa-collusive" pricing, its rivals sometimes successfully "steal business" in its home market. Fourth, when business stealing occurs in the aforementioned circumstances, it goes unpunished. Thus, we demonstrate that deliberate and unpunished business stealing (which would appear to observers as "cheating") can be critical to the healthy functioning of a cartel.

The intuition for our characterization of optimal imperfectly collusive equilibria runs as follows. When perfect collusion is unsustainable, a cartel will set a price below the monopoly level in order to reduce incentives for away firms to undercut rivals in their home markets. However, with sub-monopoly pricing, an implicit agreement to divide the market according to comparative advantage creates an incentive for each firm to defect to a supra-collusive price (indeed, the monopoly price) in its home market. Occasional entry into away markets eliminates the incentives for supra-collusive pricing in home markets. Of course, a firm will not enter an away market in equilibrium unless it receives adequate compensation for the costs of relationship investments for the customers in those markets. Accordingly, the away firm must sometimes steal business successfully. Because that form of business stealing is part of the equilibrium, it goes unpunished. However, other forms of business stealing will trigger punishments.

Notice that this theory has a potentially testable implication: unpunished business stealing occurs at prices *above* the cartel's price target, whereas all business stealing at prices *below* that target triggers punishment.* In practice, this implication may prove difficult to test because a cartel's price targets may not be known with precision (indeed, in practice, price agreements can sometimes appear somewhat fuzzy). For this purpose, one cannot use the average price as a proxy for the target price, because the theory implies that some business stealing will occur at below-average prices. A more robust testable implication of the theory is that, the higher the price at which business is stolen, the lower the likelihood that apparent "cheating" will trigger punishement. Testing this implication falls outside the scope of our current theoretical inquiry.

The rest of this paper is organized as follows. We present the basic model of industrial competition in section 2. Section 3 characterizes the properties of non-collusive equilibria (i.e., equilibria in one-shot play). Section 4 presents results for the case of two firms, and

*Technically, in our model, business stealing below the target price occurs with probability zero. However, in a more general model it could occur for unrelated reasons, such as imperfect control over "rogue" salespeople who have no knowledge of the cartel; see footnote 2.

section 5 explains how those results extend to settings with more than two firms. Some brief concluding remarks appear in section 6.

2 The model

A set of firms compete for the business of a collection of customers in an infinite sequence of discrete periods ($t = 1, 2, ...$). For the time being, we assume that there are two firms ($i = 1, 2$) and two markets. The firms are risk-neutral profit maximizers and share a common discount factor $\delta \in (0, 1)$. They are also differentiated, with each holding a (symmetric) comparative advantage in one of the two markets. For concreteness and simplicity, we model these advantages as pertaining to costs: the marginal cost of production is c_H in a firm's home market and c_A in its away market, where $\Delta c \equiv c_A - c_H > 0$. In each period, each firm decides whether to quote a price in each market. Price quotations are costly, requring the firm to incur a "relationship-maintenance" expense, c.

Each consumer views the products of the two firms as perfect substitutes, and has unit demand for all prices up to some privately known reservation value $v \sim G(\cdot)$, a CDF. (Equivalently, one can think of the market as encompassing two consumer segments, each containing a continuum of consumers with unit measure, with deterministic but privately known reservation values distributed according to G.) Consumers purchase the good from the firm offering the lowest price provided it does not exceed their reservation value v, with ties broken randomly.

Aggregate demand at price p is given by $D(p) \equiv 1 - G(p-)$.* We use $\underline{p}_i$ ($i = H, A$) to denote the single-firm break-even price (accounting for both c_i and c), and p_i^* to denote the monopoly price. To avoid uninteresting technical complications, we assume that these prices are well-defined and unique, and that a monopolist's profits would increase monotonically for prices between $\underline{p}_i$ and p_i^*. The assumptions can be derived from more primitive restrictions on G, but neither the restrictions nor their derivation are enlightening in this context. We also assume that $\underline{p}_A < p_H^*$, i.e. that Δc isn't too large, to ensure that the low-cost firm can't simply blockade entry at the optimal single-firm price.

For each market in each period, each firm chooses a pair (π, F), where π represents the probability of quoting a price and F is a CDF governing the distribution of the price, conditional upon a quotation. Both decisions (whether and what to quote) are made at the same time, without knowledge of the other firms' choice.

The following example concretely illustrates the type of setting our model is intended to capture. Boeing and Airbus compete to supply commercial aircraft to airlines worldwide. Suppose that, due to plant location, familiarity with regulatory regimes, and so forth,

*By $G(p-)$, we mean the limit of $G(p')$ as p' approaches p from below.

Boeing has a natural advantage in the US (for airlines such as Delta), while Airbus has a natural advantage in Europe (for airlines such as British Airways). It is also natural to assume that the process of preparing and submitting proposals for fulfilling carriers' new aircraft requirements entails substantial costs, irrespective of whether the firm wins. Our analysis concerns the nature of imperfect collusion in such industries.

3 The non-collusive outcome

Before studying collusion, we first characterize non-collusive outcomes by studying equilibria in a one-shot version of the game. The following theorem provides this characterization:

Theorem 1. *There exists a unique Nash equilibrium of the stage game. In this equilibrium:*

1. *The home firm always quotes a price, while the away firm quotes a price with probability between 0 and 1.*
2. *The home firm makes profits $\Pi_H = D(\underline{p}_A)\Delta c$, while the away firm makes zero profits.*
3. *Each firm's price distribution has full support on $[\underline{p}_A, p^*_H]$, and is atomless except at p^*_H, which the home firm chooses with strictly positive probability.*

Several features of this equilibrium merit emphasis. First, the consumer ends up paying the price $\underline{p}_A$ with zero probability. This is significant because it is natural to interpret $\underline{p}_A$ as the fully competitive price, given that it reflects the lowest price that the less efficient firm would be willing to charge. Still, the expected profits earned by each firm are the same as if both set $\underline{p}_A$ and the consumer purchased from the home firm. Second, because the distribution of prices for both firms has full support on $[\underline{p}_A, p^*_H]$, the *ex post* outcome can appear arbitrarily collusive. Third, "business stealing" (defined here as the away firm winning sales) occurs with strictly positive probability.

4 Optimal collusion

Industry profits are maximized when each market is monopolized by the low-cost firm posting price p^*_H. This outcome, which we refer to as *perfect collusion*, serves as a natural benchmark against which to measure the effectiveness of a collusive arrangement.* We will

*To be sure, there exist other points along the firms' Pareto frontier. But such points are technologically inefficient and not robust to the introduction of side transfers.

refer to any equilibrium yielding firm profits strictly above the stage-game Nash but below perfectly collusive levels as an *imperfectly collusive* arrangement.

When firms are sufficiently patient, a folk theorem result obtains and perfect collusion is sustainable as a subgame perfect Nash equilibrium. More precisely, there exists a critical discount factor $\delta^M < 1$ above which perfect collusion is sustainable and below which it is not. We study collusion for discount factors below δ^M.

We first establish that imperfect collusion is sustainable for a range of discount factors below δ^M.* Our main result of this section goes on to characterize optimal collusion within a natural class of equilibria. In this equilibrium, firms set a floor price p^* and divide the market along cost lines in each period, with firm i capturing all the business of market i whenever he charges p^*. Any undercutting of p^* leads to a price war yielding lifetime continuation profits $\underline{\Pi}$. On the other hand, entry into a firm's away market at prices above p^* is left unpunished. In this equilibrium, each firm occasionally enters its away market at prices above p^* and wins the other firm's customer. Our model thus predicts *equilibrium business stealing.*

4.1 Stationary equilibria

Characterizing the Pareto frontier of a repeated game for fixed discount factors is a difficult task. Several notable papers illustrate the complexities inherent in this task. Abreu, Pearce, and Stacchetti (1990) construct a set-valued mapping whose largest fixed point is the set of SPNE payoffs, but they provide few analytic properties of the set. Mailath, Obara, and Sekiguchi (2002) characterize player-optimal pure-strategy equilibria of the repeated prisoner's dilemma. They show that these equilibria are often non-stationary and cyclic of arbitrarily long period. Abreu and Sannikov (2012) examine extremal pure-strategy equilibria of two-player finite-action games. They show that the number of extremal equilibria is finite and bounded by the size of the action set, and that extremal payoffs are characterized by a system of nonlinear equations. We are aware of no work attempting a similar characterization for mixed-strategy equilibria on continuum action spaces.

To retain tractability and provide sharp characterizations of optimal collusion, we therefore restrict attention to SPNEs exhibiting a natural stationary structure.

Definition 1. *Fix an SPNE σ. We call σ a* stationary equilibrium *if there exists a stage-game strategy profile τ and a set of game histories $\widetilde{\mathcal{H}}^\infty \subset \mathcal{H}^\infty$, the* equilibrium path*, such that: 1) $\widetilde{\mathcal{H}}^\infty$ occurs with probability 1 under σ; and 2) $\sigma(h^t) = \tau$ after all partial histories*

*This result is not entirely obvious a priori. For instance, in symmetric Bertrand competition with no fixed costs, the equilibrium set has a "bang-bang" structure with perfect collusion sustainable for $\delta \geq 1/2$ and no collusion sustainable otherwise.

h^t lying on the equilibrium path.[*]

This definition singles out equilibria in which behavior looks "the same" in all periods so long as no deviations have taken place. It imposes no restrictions on equilibrium structure off-path, as optimal penal codes are generally non-stationary. In practice the equilibrium path is easy to identify, as it is generated by a set of "acceptable" action profiles that lead to a repetition of the on-path strategy profile at each stage. The on-path strategy profile τ will then consist of mixtures over this allowable action set, though possibly not with full support over the set.

We restrict attention to stationary equilibria primarily for reasons of tractability. While we can't rule out the possibility that non-stationary equilibria would improve upon an optimal stationary equilibrium, it is not obvious that non-stationarity provides any advantages in our setting. Further, even within this class equilibria have a surprisingly rich structure that matches anecdotal features of real-world cartels. We are therefore confident that stationary equilibria deliver the essential qualitative features of collusion in our model.

4.2 Optimal collusion with stationary equilibria

The main result of this part is that an equilibrium of the form described at the beginning of the section is optimal within the class of stationary equilibria. Further, we can characterize the maximally collusive profits that can be sustained by such an arrangement.

We first introduce a bit of notation that will be used frequently in formuating our results. Let $\Pi^C \equiv D(\underline{p}_A)(\underline{p}_A - c_A) - c = \Delta c D(\underline{p}_A)$ be the "competitive" profits of the home firm in the unique Nash equilibrium of the stage game. Similarly, let $\Pi^M \equiv D(p^*_H)(p^*_H - c_H) - c$ be the stage profits of a monopolist in his home market. Our interest lies with stationary equilibria sustaining lifetime profits in the interval $[\Pi^C, \Pi^M]$.[†] At the other extreme, let

[*] For a pure-strategy equilibrium, the equilibrium path is a singleton set consisting of the unique game history picked out by σ. With mixed strategies and a continuum action space, many "on-path" partial histories might occur with probability zero, so we must proceed more delicately.

σ interpreted as a behavioral strategy profile induces, for each t, a probability measure μ_t on the set of partial histories $\mathcal{H}^t$ (endowed with Lebesgue measure). By the Kolmogorov extension theorem, we can uniquely extend the collection $\{\mu_t\}_{t=1}^{\infty}$ to a measure μ on the set of complete histories $\mathcal{H}^\infty$. There then exists an equivalence class of measurable subsets of $\mathcal{H}^\infty$ with measure 1 under μ.

A stationary equilibrium singles out some member $\widetilde{\mathcal{H}}^\infty$ of this class as the "equilibrium path." A partial history $h^t \in \mathcal{H}^t$ lies on the equilibrium path if there exists some $h^\infty \in \widetilde{\mathcal{H}}^\infty$ such that $\zeta_t(h^\infty) = h_t$, where ζ_t is the projection operator onto the first t dimensions.

[†] As is standard convention in the game theory literature, we normalize the NPV of a firm's income stream by $1-\delta$ to obtain lifetime profits. Lifetime profits are interpretable as a weighted average per-period profitability, with weight $\delta^t(1-\delta)$ on the tth period. In a stationary equilibrium, lifetime profits are equal to the stage profits in each period on path.

$\underline{\Pi}(\delta)$ be the infimum of SPNE-supportable lifetime profits for a firm when the discount factor is δ.

For profits $\Pi \in [\Pi^C, \Pi^M]$, it will be useful to have notation for the price a monopolist would charge in his hoome market to achieve profits Π. Define $p^*(\Pi)$ to be this price; formally it is the solution in $[\underline{p}_A, p^*_H]$ to $\Pi = D(p)(p - c_H) - c$. Given our assumptions on $D(\cdot)$, $p^*(\cdot)$ is single-valued, continuous, and strictly increasing.

Finally, let δ^M be the minimal discount factor at which perfect collusion is sustainable. The results of this part establish that $\delta^M < 1$ and that perfect collusion is always sustainable above δ^M. Meanwhile for a range of discount factors below δ^M, we show that optimal stationary equilibria exhibit a simple intertemporal structure consisting of two phases:

1. In the cooperative phase, firms agree to maintain prices above a target p^* in each market. Firms stay in the cooperative phase so long as no firm undercuts p^*. Following undercutting, firms transition to the punishment phase.

2. In the punishment phase, firms engage in a price war yielding negative stage profits to each firm. The punishment phase continues until both firms participate in the price war for a single period, and firms then revert to the cooperative phase.

It is intuitively clear that the amount of collusion sustainable in the cooperative phase depends on the harshness of the price war that can be promised following a deviation. But at the same time, the harshest sustainable price war depends on the profitability of the reward promisable as recompense for participating in the price war. Therefore, as is typical in the analysis of repeated games, we will have to simultaneously characterize sustainable cooperative and punishment outcomes.

We first characterize the highest cooperative payoffs taking as given a harshest possible punishment continuation $\underline{\Pi}(\delta)$, whose value is as yet unknown. Because we restrict attention to stationary equilibria, we don't need to know anything more about the set of equilibrium payoffs: on-path continuations must be equal to the equilibrium payoff, while off-path continuations might as well be set as harshly as possible to deter deviations. The following theorem describes optimal stationary equilibrium payoffs for given $\underline{\Pi}(\delta)$:

Theorem 2. *For all $\delta \le \delta^M$, the unique Pareto-optimal stationary equilibrium profit vector (Π^*, Π^*) satisfies*

$$\Pi^* = (1 - \delta)(2\Pi^* - \Delta c D(p^*(\Pi^*))) + \delta \underline{\Pi}(\delta).$$

Further, $\Pi^* > \Pi^C$ iff $\underline{\Pi}(\delta) < \Pi^C$, and Π^* is strictly increasing in δ whenever $\underline{\Pi}(\cdot)$ is non-increasing in δ.

The essence of this theorem is that optimal collusion awards each firm its entire stage profits in its home market; profit-sharing cannot be optimal. An optimal collusive arrangement then divides the market along cost lines with a floor price $p^*(\Pi^*) < p^*_H$, subject to occasional overcharging and business stealing. In these circumstances each firm can capture extra profits $\Pi^* - \Delta c D(p^*(\Pi^*))$ by undercutting the target price in the away market, which must be deterred with a punishment continuation of $\underline{\Pi}(\delta)$. At the optimum, Π^* is set to just saturate the resulting incentive constraint. Note in particular that the Pareto frontier is degenerate, and both firms agree on the optimal collusive scheme. There are no asymmetric optimal stationary equilibria.

The previous result holds no matter the specific punishments available to the firms. The next theorem characterizes the harshest sustainable punishment for sufficiently large δ.

Theorem 3. *Suppose $\delta \geq \underline{\delta} \equiv \frac{1}{2+\Delta c D(p^*_H)/c}$. Then there exists an SPNE yielding lifetime profits of zero to both firms. Hence $\underline{\Pi}(\delta) = 0$.*

The punishment equilibrium supporting zero profits has a bang-bang structure, with a short-term price war followed by optimal collusion afterward. In the first period, firms drive prices below the break-even level. Both firms enter each market, the home firm prices at some $p^{PW} \leq \underline{p}_H$, and the away firm mixes over a distribution with support on $(p^{PW}, \underline{p}_H]$. The home firm is therefore forced to serve the market at an unprofitable price, while the away firm loses money on his fixed costs. If both firms participate in the price war for one period, they transition to a continuation game with lifetime profits Π^* characterized by theorem 2. Otherwise the price war continues. p^{PW} is chosen so that lifetime profits $(1-\delta)(D(p^{PW})(p^{PW} - c_H) - 2c) + \delta\Pi^*$ are set equal to zero. The restriction on δ in the theorem ensures that $p^{PW} \leq \underline{p}_H$ and thus that neither firm has a deviation worth more than 0 (achieved by exiting both markets).

The next theorem synthesizes the results of theorems 2 and 3. Under a mild sufficiency condition on Π^M, it shows that $\underline{\delta} < \delta^M$. Thus there exists a range of discount factors below δ^M for which optimal collusion can be completely characterized.

Theorem 4. *Suppose $\Pi^M \geq c + \Delta c D(p^*_H)$. Then there exists a $\underline{\delta} < \delta^M$ charactered by lemma 3 such that, for all $\delta \in [\underline{\delta}, \delta^M]$, $\underline{\Pi}(\delta) = 0$ and the unique Pareto-optimal stationary equilibrium profit vector (Π^*, Π^*) satisfies*

$$\Pi^* = (1-\delta)(2\Pi^* - \Delta c D(p^*(\Pi^*))).$$

Further, Π^* is continuous, strictly greater than Π^C, and strictly increasing in δ. Finally, δ^M is characterized by

$$\delta^M = \frac{\Pi^M - \Delta c D(p^*_H)}{2\Pi^M - \Delta c D(p^*_H)},$$

and perfect collusion is sustainable for all $\delta \geq \delta^M$.

For $\delta < \underline{\delta}$ we can't guarantee that a continuation payoff of 0 is sustainable as an SPNE. Whenever $\underline{\Pi}(\delta) > 0$, maximum sustainable collusive payoffs fall below the level of theorem 2. However, the form of optimal collusion will be the same even for discount factors below $\underline{\delta}$.

Finally, we describe a stationary equilibrium supporting profits Π^*. This construction holds regardless of $\underline{\Pi}(\delta)$.

Theorem 5. *When $\delta < \delta^M$, lifetime profits (Π^*, Π^*) are supported by a stationary equilibrium with the following on-path properties:*

1. *The home firm's strategy is the same in each market, as is the away firm's.*
2. *The home firm always enters, while the away firm enters with probability strictly between 0 and 1.*
3. *The home firm makes stage profits Π^*, while the away firm makes zero profits.*
4. *Each firm's price distribution has full support on $[p^*(\Pi^*), p_H^*]$ and is continuous on $(p^*(\Pi^*), p_H^*)$. At $p^*(\Pi^*)$ and p_H^* the home firm places atoms while the away firm's price distribution is continuous.*
5. *Business stealing occurs with strictly positive probability and is decreasing in Π^*.*
6. *Deviations by the away firm to prices at or below $p^*(\Pi^*)$ are punished by a continuation payoff of $\underline{\Pi}(\delta)$ to that firm.*

This equilibrium is unique among those whose price distributions have full support on $[p^*(\Pi^*), p_H^*]$ in each market.

This equilibrium features a "target price" $p^*(\Pi^*)$ with occasional overcharging, but no undercutting, by each firm. The home firm always participates in his market and makes expected profits Π^* for any price he charges, while the away firm only occasionally enters and just makes up his fixed costs at all prices charged. Business stealing occurs with positive probability, which decreases as the equilibrium becomes more collusive (Π^* rises) and goes to zero as $\Pi^* \uparrow \Pi^M$. However, business stealing occurs only at prices strictly above the target price. The home firm would always capture the market if he charged exactly the target price! As a result, business stealing actually makes the customer worse off than a clean division of the market at the target price due to the accompanying overcharging.

4.3 The role of business stealing in collusion

We have shown that optimal collusive profits can be supported by an optimal stationary equilibrium exhibiting business stealing. Is this equilibrium the only possible one? It turns out that there exist a multiplicity of optimality equilibria in which firms' price distributions don't have full support on $[p^*(\Pi^*), p^*_H]$. Such equilibria can be constructed by "drilling holes" in the full-support distribution of theorem 5. The away firm will have a profitable deviation into each hole; but so long as the holes are sufficiently small, this deviation will be less profitable than undercutting $p^*(\Pi^*)$ and so will not tighten the IC constraints.

Despite this multiplicity, we can characterize several important features of any optimal collusive agreement. In particular, we can show that business stealing is inevitable in an optimal stationary equilibrium, and in fact the equilibrium of theorem 5 places a lower bound on the amount of business stealing in an optimal stationary equilibrium. This result, along with other properties possessed by all optimal equilibria, are stated formally in the following theorem:

Theorem 6. *When $\delta < \delta^M$, all stationary equilibria supporting profits (Π^*, Π^*) feature the same target price, maximum price, and entry probabilities by each firm in each market. In particular, the home firm always enters while the away firm enters with a fixed probability between 0 and 1. Further, the equilibrium characterized in theorem 5 uniquely minimizes business stealing among all stationary equilibria supporting profits (Π^*, Π^*).*

The necessity of business stealing is robust to the consideration of non-optimal equilibria. For collusive schemes yielding profits sufficiently close to (Π^*, Π^*), any stationary equilibrium supporting these profits must involve business stealing. To see this, take a stationary equilibrium yielding lifetime profits $(\Pi, \Pi) > (\Pi^C, \Pi^C)$ to each firm which exhibits no business stealing. For simplicity, consider the case where market i is monopolized by firm i, who charges price $p^*(\Pi)$. The most profitable deviation by each firm is to overcharge at price p^*_H in his home market, and to just undercut price $p^*(\Pi)$ in his away market. The IC constraint implied by this deviation is

$$\Pi \geq (1-\delta)(\Pi - \Delta c D(p^*(\Pi)) + \Pi^M) + \delta \underline{\Pi}(\delta).$$

If we raise Π to Π^*, this constraint must be violated given $\Pi^* < \Pi^M$. Then by continuity of $D(\cdot)$ and $p^*(\cdot)$, there exists a $\widetilde{\Pi} < \Pi^*$ at which the constraint is just saturated, and above which it is always violated.

Consider in particular the special case where the customer's reservation value is known, i.e. $D(p) = 1$ below some cutoff v. In this case the rhs of the IC constraint actually rises at a lower rate than does the lhs as we raise Π. Therefore if any collusion is possible without business stealing, so is perfect collusion! We collect these results in the following theorem:

Theorem 7. *Fix $\delta < \delta^M$. Then there exists a $\widetilde{\Pi} < \Pi^*$ such that all stationary equilibria supporting profits $(\Pi_1, \Pi_2) > (\widetilde{\Pi}, \widetilde{\Pi})$ exhibit business stealing. When the purchaser's reservation value is known, $\widetilde{\Pi} = \Pi^C$.*

The essence of this result is that in the absence of business stealing, each firm has an incentive to overcharge in his home market in addition to undercutting his competitor. To sustain substantial collusion, this overcharging incentive must be damped, which can be accomplished via competition from the away firm at high prices. The compensation demanded by the away firm for this competition is a share of the market large enough to cover his entry costs, generated through occasional overcharging by the home firm. Thus in equilibrium we must have both overcharging by the home firm and business stealing by the away firm to eliminate profitable overcharging deviations. (Overcharging still occurs in equilibrium, but it is no longer more profitable than following the target price.)

5 Generalization to many competitors

One implication of our results is that $\delta^M < 1/2$ (by theorem 4), and thus a pair of firms can sustain perfect collusion for discount factors bounded well away from 1. If δ is taken literally as reflecting the market interest rate and the length between periods, then firms that interact with any frequency should have no difficulties achieving perfect collusion. In particular, the form of collusion below δ^M would be mostly irrelevant to understanding cartel behavior.

However, δ is better interpreted more expansively as a reduced-form stand-in for a variety of factors cutting against forward-looking behavior. For instance, firms may face above-market internal rates of return; agency problems or leadership turnover; potential changes in market structure; and so on. Firms might also face more than one competitor in the market, which intuition suggests would make coordination and effective collusion more difficult.

In this section we explore the possibility of multiple competitors in detail. We show that collusion indeed becomes more difficult to sustain as the cartel size increases, and that with even a moderate number of firms imperfect collusion is inevitable even at high discount factors. We also generalize the theorems of section 4 and characterize optimal collusion below δ^M, which looks broadly similar to the two-firm case. Our results thus serve as a robustness check on the applicability of our insights about optimal collusion to more general market structures.

5.1 Setup

We extend the two-firm market structure in a symmetric way. There are now $N + 1$ firms and $N + 1$ markets, where $N \geq 2$. In market i firm i is the home firm and produces at marginal cost c_H, while all other firms are away firms and produce at marginal cost $c_A > c_H$. All other aspects of the model are identical to the two-firm case.

5.2 Analyzing the stage game

Consider one-shot competition in a single market with 1 home firm and N away firms. Let $\{H\} \cup \mathscr{I}$ be the set of firms, with $\mathscr{I} = \{1, ..., N\}$ the set of away firms. An immediate corollary of the existence of a two-firm equilibrium is that there exist a multiplicity of Nash equilibria when $N \geq 2$. For if H and any firm $i \in \mathscr{I}$ play the two-firm equilibrium, no away firm will want to enter (as he would receive strictly less than i's profits of zero). Hence there exists at least N Nash equilibria involving competition among two firms.

In fact, there exist Nash equilibria involving every possible non-empty subset of away firms competing with the home firm. Our main result is that there is no other multiplicity: once a subset of away firms is chosen, there exists a unique Nash equilibrium involving those firms. The form of this equilibrium looks broadly similar to the two-firm case, with mixing from $\underline{p}_A$ up to p_H^*, sure entry by the home firm, and occasional entry by the away firms. Further, the equilibrium is symmetric in that all away firms play identical strategies. These results are summarized in the following theorem:

Theorem 8. *Fix any non-empty subset $\mathscr{J} \subset \mathscr{I}$ of away firms. Then there exists a unique Nash equilibrium of the stage game in which every firm in $\mathscr{J}$ enters with positive probability and no firm in $\mathscr{I} \setminus \mathscr{J}$ ever enters. In this equilibrium:*

1. *The home firm always enters and makes profits $\Pi_H = \Delta c D(\underline{p}_A)$.*
2. *Each away firm $i \in \mathscr{J}$ enters w.p. strictly less than 1 and makes profits $\Pi_i = 0$.*
3. *Each entering firm's price distribution has full support on $[\underline{p}_A, p_H^*]$.*
4. *All entering away firms play the same strategy.*

There are no other Nash equilibria.

5.3 Market-symmetric equilibrium

When many firms are competing, the set of possible collusive structures is much richer than in the two-firm case. In the two-firm case, we saw that it is always optimal to allocate each firm all of their profits in their home market. By contrast, with at least three firms it can be worthwhile for firms to spread their profits across multiple markets. This reduces the profitability of undercutting in each market, and can loosen incentive constraints in some cases.

In the Appendix we provide a fully worked example of an equilibrium which, for particular choices of $c, \Delta c$, and δ, Pareto-dominates the best equilibrium in which firms earn profits in only their home market. Table 1 displays the division of profits among firms for this example when $N = 2$. Firms are index by row, markets are indexed by column, and a + indicates positive profits while 0 indicates zero profits.

	M1	M2	M3
F1	+	0	0
F2	+	+	0
F3	0	0	+

Table 1: A non-market-symmetric division of profits

With no further restrictions on the form of the equilibrium, it is therefore difficult to fully characterize optimal collusion. We propose a further mild condition on equilibrium which disciplines possible cartel structures. Note that the distribution of profits in Table 1 displays two anomalous features - first, firms earn positive profits in their away markets; and second, different (ex ante identical) away firms play different strategies, and earn different profits, in a given market. Directly ruling out positive away firm profits would immediately recover the optimality results of section 4, but this simply assumes the answer and is not an obviously reasonable restriction. On the other hand, we think it reasonable that ex ante identical firms should play identical strategies in a given market. Any other outcome would require coordination between the away firms and bargaining over profit-sharing in the market, which might induce cartel breakdown.

We therefore restrict attention to equilibria with a symmetry requirement within markets:

Definition 2. *Fix a stationary equilibrium σ with associated on-path stage-game strategy profile τ. Then σ is* market-symmetric *if, for each market m and all away firms i and j, $\tau_m^i = \tau_m^j$.*

Market symmetric equilibria are a subset of stationary equilibria which require all away firms to play identical strategies in a given market.* Thus divisions of profits as in Table 1 are ruled out. Note that firms are *not* required to play the same strategy across markets. Also, all stationary equilibria are trivially market-symmetric when $N = 1$, so market symmetry disturbs none of our results from the two-firm case. Market symmetry will turn out to provide the structure we need to characterize optimal collusion.

*Alternatively, we could have required them to receive identical profits. This alternative definition would yield identical results.

5.4 Optimal market-symmetric equilibria

A bit of notation: let $\delta^M(N)$ be the minimal discount factor sustaining perfect collusion with $N+1$ firms. Also let $\overline{\delta}(N) \equiv \left(1 - \frac{1}{N}\right)\left(1 + \frac{c}{\Pi^M}\right)$, an important bound in our theorems which is sometimes larger and sometimes smaller than $\delta^M(N)$. We defer our discussion of its meaning for the moment. Finally, let $\underline{\Pi}(\delta; N)$ be the minimum SPNE-sustainable lifetime profits under discount factor δ with $N+1$ firms.

Our first theorem generalizes theorem 2:

Theorem 9. *Suppose $\delta < \delta^M(N)$ and either $N \geq \sqrt{1 + \Pi^M/c}$ or $\delta < \overline{\delta}(N)$. Then there exists a unique Pareto-optimal market-symmetric equilibrium payoff vector* $(\Pi^*, ..., \Pi^*)$, *where* Π^* *satisfies*

$$\Pi^* = (1-\delta)((N+1)\Pi^* - \Delta c N D(p^*(\Pi^*))) + \delta \underline{\Pi}(\delta; N).$$

Further, $\Pi^* > \Pi^C$ iff $\underline{\Pi}(\delta; N) < \Pi^C$, and Π^* is strictly increasing in δ whenever $\underline{\Pi}(\cdot; N)$ is nonincreasing in δ. Finally, Π^* is strictly decreasing in N whenever $\underline{\Pi}(\delta; \cdot)$ is nonincreasing in N.

The content of this theorem is identical to that of theorem 2 - provided δ is not too high, optimal collusion involves splitting profits along cost lines. The maximum sustainable level of collusion is then characterizable and depends on the optimal punishment $\underline{\Pi}(\delta; N)$ which can be meted out following a deviation.

In contrast to the two-firm case, the optimality of dividing profits along cost lines isn't guaranteed for all $\delta < \delta^M(N)$. We additionally need $\delta < \overline{\delta}(N)$, a condition which ensures that allocating profits *only* to away firms in a market-symmetric way can't do even better. In the Appendix we provide a fully worked example showing that an arrangement with away firms receiving all the profit from every market can sometimes Pareto-dominate the profits characterized in Theorem 9 when $\delta \geq \overline{\delta}(N)$. Table 2 reports the division of profits in this example, in which $N = 2$.

	M1	M2	M3
F1	0	+	+
F2	+	0	+
F3	+	+	0

Table 2: A cartel with all profits awarded to away firms

If $\delta^M(N) < \overline{\delta}(N)$, then this additional restriction is trivially satisfied. A sufficient condition for this inequality to hold is $N \geq \sqrt{1 + \Pi^M/c}$, a mild lower bound on the number of competitors relative to market size. Because this bound grows very slowly in

Π^M/c, it is likely to be satisfied in practice. For instance, $\Pi^M/c = 3$ implies $N \geq 2$ (which is true by assumption), while $\Pi^M/c = 15$ implies $N \geq 4$. Even if fixed costs were a trivial portion of monopoly profits, say 1%, the implied bound on N would be only $N \geq 10$, so we feel this restriction is relatively innocuous.

Alternatively, we could impose no restriction on N and simply require that $\delta < \overline{\delta}(N)$. The following proposition says that the range of discount factors excluded by this restriction is small:

Proposition 1. $[\overline{\delta}(N), \delta^M(N)] \subset (1-1/N, 1-1/(N+1))$. *Therefore if* $\delta \in [\overline{\delta}(N), \delta^M(N))$, *then* $\delta < \overline{\delta}(N+1)$ *and* $\delta > \delta^M(N-1)$.

This proposition tells us that $\delta \in [\overline{\delta}(N), \delta^M(N))$ is a knife-edge case: add one more firm, and alternative collusive structures collapse because $\delta < \overline{\delta}(N+1)$. Subtract one firm, and the resulting cartel can sustain perfect collusion. The size of the problematic interval $[\overline{\delta}(N), \delta^M(N))$ is also of size at most $1/(N(N+1))$, and so collapses rapidly with N. We therefore consider the possibility of alternative collusive structures an edge case that can be safely ignored.

The next theorem generalizes Theorem 3, using a punishment equilibrium construction similar to that theorem:

Theorem 10. *Suppose* $\delta \geq \underline{\delta}(N) \equiv \frac{c}{\Delta c D(p_H^*) + (1+1/N)c}$. *Then there exists an SPNE supporting lifetime profits of 0 for each firm, hence* $\underline{\Pi}(\delta; N) = 0$.

The punishment equilibrium is similar to the two-firm case, but is asymmetric: if firm i is to be punished, he is matched with some other firm, say $i+1$, and those two firms engage in a price war for one period in their respective markets. All other firms stay out of those markets and play the stage-game NE in the remaining markets. After one round of a successful price war, firms revert to the cooperative phase. Note that $\underline{\delta}(N)$ is increasing in N, but is bounded strictly away from 1. Thus even with a large number of competitors, minmax punishments are feasible without requiring arbitrarily patient firms.

Finally, we generalize Theorem 4, and fully characterize optimal collusive payoffs for a range of discount factors below $\delta^M(N)$:

Theorem 11. *Suppose* $\Pi^M > \Delta c D(p_H^*) + \frac{c}{N}$ *and* $N \geq \sqrt{1 + \Pi^M/c}$. *Then* $\underline{\delta}(N) < \delta^M(N) < \overline{\delta}(N)$, *and for all* $\delta \in [\underline{\delta}, \delta^M(N)]$ *the unique Pareto-optimal market-symmetric equilibrium profit vector* $(\Pi^*, ..., \Pi^*)$ *satisfies*

$$\Pi^* = (1-\delta)((N+1)\Pi^* - \Delta c N D(p^*(\Pi^*))).$$

Further, Π^* *is continuous, strictly greater than* Π^C, *and strictly increasing in* δ.

As usual, we require a mild sufficiency condition on Π^M to ensure $\underline{\delta}(N) < \delta^M(N)$. This condition grows weaker as N grows, and is trivially satisfied for sufficiently large N. Theorem 11 does not address the case $N < \sqrt{1 + \Pi^M/c}$. In the Appendix, we derive another lower bound on N which ensures that $\underline{\delta}(N) < \overline{\delta}(N)$, in which case Theorem 11 continues to characterize optimal profits for a range of discount factors.

Finally, we characterize an equilibrium supporting profits Π^* for each firm. As in the two-firm case, this construction holds regardless of the value of $\underline{\Pi}(\delta; N)$.

Theorem 12. *Suppose $\delta < \delta^M(N)$. Then lifetime profits $(\Pi^*, ..., \Pi^*)$ are supported by a market-symmetric equilibrium with the following properties:*

1) *The home firm's strategy is the same in all markets, and all away firms play the same strategy in all markets.*

2) *The home firm enters w.p. 1, while all away firms enter with probability strictly between zero and 1, which is decreasing in Π^*.*

3) *The home firm earns profits Π^*, while all away firms make zero profits.*

4) *Each firm's price distribution has full support on $(p^*(\Pi^*), p^*_H)$.*

5) *If $\Pi^* > \Delta c D(\underline{p}_A)$, the home firm places an atom at $p^*(\Pi^*)$, whose size is increasing in Π^*.*

6) *Business-stealing occurs with strictly positive probability, which is strictly decreasing in Π^* when $\Pi^* \geq \frac{1}{2}\Pi^M$.*

7) *Any unilateral deviation by an away firm to a price at or below $p^*(\Pi^*)$ results in a continuation payoff of $\underline{\Pi}(\delta; N)$ to that firm.*

This result mirrors the optimal collusive structure of the two-firm case, and features business-stealing for the same basic reason.

5.5 Imperfect collusion in large cartels

The following theorem explores how the range of discount factors for which we have characterized optimal collusion varies with cartel size.

Theorem 13. *$\underline{\delta}(N)$ and $\delta^M(N)$ are strictly increasing in N, and $\lim_{N\to\infty} \underline{\delta}(N) < 1$ while $\lim_{N\to\infty} \delta^M(N) = 1$. Further, $\delta^M(N) - \underline{\delta}(N)$ is increasing in N.*

Because $\delta^M(N)$ goes to 1 as N grows large, imperfect collusion is inevitable even with patient firms when the cartel is large. The minimal discount factor required to sustain a price war yielding zero profits also grows with N, but more slowly. Thus the range of discount factors for which we completely characterize optimal collusion expands with cartel size. This theorem demonstrates the robustness of our results in the many-firm case. It also illustrates that the structure of imperfectly collusive arrangements remains relevant even when firms are relatively patient.

6 Conclusion

An important empirical feature of many real-world cartels is the presence of apparently deliberate, unpunished cheating. Existing theories of collusion in repeated games struggle to explain this observation, as they predict that cheating does not occuring equilibrium, and if it does it must lead to a price war or other punishment as a deterrent. We have attempted to bridge this gap by constructing a theory of equilibrium business stealing in imperfect cartels. Our theory considers a model in which posting prices is costly, and focuses on discount factors for which some, but not perfect, collusion is sustainable.

We find that, within a natural class of equilibria, optimal collusion necessarily entails business stealing. Moreover, equilibria sustaining optimal collusion look very much like imperfectly enforced cartel agreements fixing prices and dividing markets. Cartel members attempt to divide the market according to firms' cost advantages at a fixed price, but occasionally overcharge in their own market and post prices above the target price in their competitors' markets. Business stealing above the target price is met with forbearance, yielding a prediction of equilibrium cheating.

A primary feature of collusion in our model is that not all business stealing is created equal. Business stealing below a floor price is punished harshly and never occurs in equilibrium, while business stealing above the price floor is both accepted and critical to cartel functioning. The stark divide between acceptable and unacceptable business stealing in this model would be qualified in an extension with imperfect monitoring, say due to random price cuts by sales staff. Nonetheless, a robust testable implication of our theory is that business stealing at high prices will trigger punishment less frequently than at low prices.