Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity

R. Preston McAfee; Marius Schwartz


Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28199403%2994%3A1%3C210%3AOVMCN%3E2.0.CO%3B2-C

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*The American Economic Review* is published by American Economic Association. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/aea.html.

*The American Economic Review*
©1994 American Economic Association

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR
Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity

By R. Preston McAfee and Marius Schwartz *

An input supplier selling to competing downstream firms would benefit from publicly committing at the outset to all contracts. Efficient commitment, however, would require complete contracts. We study instead bilateral contracting, without commitment regarding others’ terms. Each firm then fears that the supplier might opportunistically renegotiate another’s contract to increase bilateral profit at the firm’s expense. We show that nondiscrimination clauses generally cannot curb such third-party opportunism, even with symmetric firms. To reassure firms, crude forms of commitment may be adopted. This could explain the pervasiveness of exclusivity arrangements and the striking uniformity and intertemporal rigidity of franchise contracts. (JEL C72, D43, D45, L14, L42)

“Once he has made a deal, Mr. Lorenzo can’t sleep on it. He wants to renegotiate.”
—David Shapiro, court-appointed examiner in Eastern Airlines bankruptcy
(Washington Post, 22 April 1990)

This paper studies the contracting problem of an input supplier dealing with several firms that compete in an output market. Examples of this relationship include a manufacturer selling to distributors, a patent-holder licensing several producers, and a franchisor with several franchisees. To focus on the vertical-contracting aspect, we abstract from supplier competition by considering an input monopolist. Since the monopolist’s customers compete with one another, their input demands will be interdependent: each firm cares about the terms that the monopolist offers to all.

With a few recent exceptions discussed shortly, the literature addressing such a setup has assumed “public commitment.” The monopolist publicly makes committed offers to all firms; therefore each firm can predict rivals’ marginal costs when evaluating the offer it received (see e.g., Frank Mathewson and Ralph A. Winter, 1984; Martin K. Perry and Robert H. Porter, 1989).1 The commitment assumption is significant: once a firm incurs certain relationship-specific investments—such as buying the monopolist’s inputs, paying a franchise fee, or engaging in promotion—the monopolist’s incentives typically are altered. Since the initial firm is somewhat locked in, the monopolist might gain by recontracting with another firm, for instance, by reducing the latter’s input price in exchange for a higher fixed fee.

There is abundant evidence of businessmen’s concern with opportunism. Allegations have included various attempts to drive out incumbent franchisees or dealers from profitable locations, and encroachment on

* McAfee: Department of Economics, University of Texas, Austin, TX 78712; Schwartz: Department of Economics, Georgetown University, Washington, DC 20057-1045, and Antitrust Division, U.S. Department of Justice. The views expressed in this paper do not purport to represent those of the U.S. Department of Justice. For helpful comments and discussions we thank Tim Brennan, Maxim Engers, Drew Fudenberg, Francine Lafontaine, George Mailath, Dave Maluie, Greg Shaffer, Jean Tirole, Greg Vistnes, Greg Werden, Robert Willig, and Ralph Winter. Two anonymous referees suggested the “wary beliefs” and other changes that greatly improved the paper.

1Further references can be found in Jean Tirole (1988) and Michael L. Katz (1989).
incumbents’ market areas by the addition of new outlets (Gillian K. Hadfield, 1990; Business Week, 1992; Wall Street Journal, 1991, 1993). In two of the oldest distribution systems, automobiles and gasoline retailing, federal statutes govern termination and nonrenewal of dealers. Since the 1970’s, numerous states have passed similar termination and nonrenewal laws covering these and other industries (American Bar Association Antitrust Section [henceforth ABA], 1990; Hadfield, 1990); in the case of automobiles, since 1963, 37 states have adopted “relevant market area laws” restricting encroachment by manufacturers into areas of existing dealers (ABA, 1990). While some have characterized these initiatives as special-interest legislation, various commentators and the Supreme Court have seen them as at least partly designed to protect franchisees against opportunistic abuses of bargaining power (ABA, 1990).2

Opportunism, of course, is an issue also in bilateral relationships (see e.g., Benjamin Klein et al., 1976; Oliver E. Williamson, 1985; Paul L. Joskow, 1987). Efficient commitment against opportunism would require complete state-contingent contracts, but actual contracts are highly incomplete, due to difficulties of anticipating all contingencies and spelling out appropriate performance, of verifying performance or the state, and of enforcement (Perry, 1989; Hadfield, 1990; ABA, 1991; Paul R. Milgrom and John Roberts, 1992). When a supplier sells to competing firms, the difficulties of guarding contractually against opportunism are compounded, as opportunism can take the additional form of changing the terms offered to a firm’s rivals. Committing efficiently to customers about one’s dealings with third parties can be especially difficult.

A difficulty unique to such multilateral contracting is that some of the monopolist’s relevant contracts can be unobservable even to insiders: with secret discounts, a firm simply may not know the true input prices charged to its rivals (inferring prices ex post from rivals’ market behavior also can be problematic if there are exogenous, firm-specific shocks). Even if insiders can observe all that is relevant, the usual costs of writing and enforcing complete contracts can be higher in multilateral than in bilateral contracting, because the universe of items to be specified and verified rises with the number of parties.

Although we do not formally model the above obstacles to efficient commitment, these obstacles, which are comprehensively discussed by Milgrom and Roberts (1992), motivate our inquiry. In this paper we explore the consequences for the monopolist of selling to competing firms under no commitment to any one firm regarding the terms offered to others. Our unifying theme is that fear of multilateral opportunism harms the monopolist in a variety of settings. The monopolist therefore may accept distortions arising from crude and seemingly inefficient forms of commitments in order to reassure its prospective trading partners.

The paper is organized as follows. Section I presents the model, defines the benchmark commitment solution, and shows why this joint-maximizing (commitment) arrangement generally is susceptible to opportunism. This is shown in a simple perfect-information game, to make the argument transparent and defer the delicate issue of firms’ beliefs about rivals’ contracts. We assume throughout that the monopolist can share in downstream rent via a two-part tariff: a constant marginal price and a fixed

---

2 The absence of comparable legislation in business format franchising (as in fast foods) is probably due to its being a relatively young industry, dating only to the 1960’s; with franchise disputes showing no sign of abating, pressure has mounted to extend legislation to this industry. Recently Iowa has enacted a law that would restrict termination and encroachment in all franchise relationships, and several states are contemplating following suit (Wall Street Journal, 1992a, 1993).

Even reputable companies are not immune to charges of opportunism. In 1987 General Motors launched a plan to designate some 2,000 independent repair shops nationwide as Delco-Tech Service Centers stocking GM parts, in an attempt to capture some of the business in service and minor repairs that was being lost to independents (such as Jiffy Lube). GM suspended the program in 1990 under stiff opposition from its dealers, who complained vociferously that the proposed program would divert even more repair business away from them, after they had invested heavily in repair facilities (Automotive News, 1990).
fee. The two-part tariff is one of the most basic pricing schemes, and it has been emphasized in the vertical-control literature as a powerful instrument for aligning incentives in bilateral relations (e.g., by alleviating double-marginalization and moral hazard). To simplify the exposition we focus on two-part tariffs, but we will indicate how results change with more general pricing schemes. 3

Section II studies the possibility of curbing opportunism by offering nondiscrimination clauses. Since contracts are incomplete, parties in a bilateral relation will often renegotiate terms as new information emerges, in order to move toward ex post efficiency. In a multilateral setting, however, recontracting that is bilaterally profitable can harm others. 4 Those other parties cannot be protected if they cannot observe such bilateral recontracting; where firms (eventually) can observe all contracts, however, the challenge is to preserve flexibility for legitimate recontracting while providing assurance to third parties. Nondiscrimination or "most-favored-customer" clauses naturally spring to mind, as they have been widely touted as a way to prevent selective price cuts. (Shortcomings of some other approaches are discussed later.) In the Coase-conjecture literature, the durable-good monopolist profitably eliminates its urge to offer future price cuts if it commits to make past customers retroactively eligible for any future discounts (David Butz, 1990) (see also Ronald H. Coase, 1972; Tirole, 1988). Thomas E. Cooper and Timothy L. Fries (1991) show that a monopolist bargaining sequentially with noncompeting buyers can stiffen its posture against the second buyer by adopting a policy of nondiscrimination. Patrick J. DeGraba and Andrew Postlewaite (1992) study an input monopolist selling to competing buyers and still find that, for contracts that include only fixed fees, nondiscrimination clauses prevent opportunism.

We show that for two-part tariffs (or more complicated contracts) the use of nondiscrimination clauses to curb opportunism against competing firms can be ineffective. Once a firm accepts a lower marginal price in exchange for a higher fixed fee, a second firm may well prefer to stay with its higher input price and lower fee even if the deviation contract is available to it ex post ("non-discrimination"). Thus, nondiscrimination clauses can have little bite—even when firms are ex ante identical. This is our most novel finding.

Section III studies opportunism in a more natural setting than the perfect-information game of Section I, by letting every firm potentially fear that the monopolist might recontract with rivals. We consider the monopolist offering contracts simultaneously and secretly; thus, each firm is ignorant of offers made to others when it decides whether to accept. A firm either never learns others' contracts (and hence, marginal costs) ("unobservability game") or learns them after accepting its contract but before competing downstream ("ex post observability game"). The former is motivated by cases in which price commitments are impossible because of secret discounts; the latter, by cases in which verifiability problems (or other contracting costs) preclude complete

---

3 In practice, simple affine pricing schemes are common. Francine Lafontaine (1992a) surveyed franchisors in different industries and found that, of the 127 respondents, 122 charge an up-front franchise fee and 123 charge a royalty rate; of the 123, 93 state that the royalty rate is constant (20 state that it is piecewise linear: 18 increasing and 2 decreasing). In such franchise contracts, the royalty rate is typically a percentage of revenue ("sales") rather than based on quantity. However, since the royalty acts as a revenue tax, raising the royalty rate has qualitatively similar effects to raising the marginal price in a two-part tariff: both will decrease a firm's optimal output.

4 Several authors have noted the advantages in a bilateral setting of committing to a default option (the status quo) and readjusting terms whenever this is mutually beneficial, both in models without initial investment (Steven Shavell, 1984; Milton Harris and Bengt R. Holmstrom, 1987) and in ones where expected division of surplus does influence initial investments (Oliver D. Hart and John Moore, 1988; Gur Huberman and Charles M. Kahn, 1988). There is also a literature showing that foreseen renegotiation can be inefficient because it undermines commitment (e.g., Hart and Tirole, 1988; Drew Fudenberg and Tirole, 1990). Our interest is not in the net impact of renegotiation in a bilateral setting, but in how bilateral renegotiation affects third parties. In the formal model we consider "renegotiation" designed solely to exploit third parties.
contracts and committing to uncontingent fixed-price contracts would be inefficient.

In both games, contract offers are secret; hence firms' beliefs are crucial. We consider three sets of beliefs: "symmetry beliefs" (each firm believes all receive the same contract), "passive beliefs" (each firm does not revise its beliefs about other offers based on its offer), and "wary beliefs" (explained later). Except when firms hold symmetry beliefs (and are identical), the monopolist suffers from fear of opportunism in both games. Under unobservability, a fairly general result emerges: for both passive beliefs and wary beliefs, the equilibrium outcome is for the monopolist to set price equal to marginal cost to all firms. This outcome is independent of the details of the downstream competition (strategic substitutes or strategic complements).\(^5\)

Since lack of commitment harms the monopolist in various settings, some vertical practices that are seemingly crude and inefficient may be adopted because they help curb opportunism. Section IV discusses some such practices. Short of integration, one approach is to promise to deal with a single firm. This does entail some commitment; but it is typically much easier to determine whether a new competitor has been brought in than to monitor and verify a court the precise terms offered to rivals. Our analysis thus suggests an explanation for practices such as exclusive distribution territories and exclusive dealing even absent the usual free-rider problems due to the provision of "public" services or assets (Lester G. Telser, 1960; Howard P. Marvel, 1982). Indeed, committing to exclusivity might be profitable despite an efficiency loss.

If the efficiency loss from dealing with a single firm in a given market is too high (e.g., due to sharply increasing costs) or if such an arrangement is impractical (e.g., due to customer mobility), the monopolist would deal with several firms. In such cases we would expect contract terms to be transparent and quite uniform across firms, to reduce the scope for camouflaging selective discounts. This could help explain the remarkable and seemingly inefficient degree of uniformity observed in franchise contracts, and the rigidity of contract terms over time (Shumeet Banerji and Carol Simon, 1992; LaFontaine, 1992a, 1993; LaFontaine and Kathryn L. Shaw, 1992).

1. Commitment Benchmark and Opportunism Incentive

Consider an input monopolist facing \( n \geq 2 \) potential downstream firms that could use the input to produce substitute products (perfect or imperfect). The monopolist has no fixed cost and has constant marginal cost \( z > 0 \). A two-part tariff offered to firm \( i \) is a pair \((r_i, f_i)\) where \( f_i \) is a fixed fee and \( r_i \) is the marginal price per unit of the input. A firm's marginal cost increases with \( r \), and a sufficiently high \( r \) would make marginal cost prohibitive. In our environment, downstream efficiency may require having one or more firms active; the latter could arise if the firms offer imperfect substitutes or experience increasing marginal costs for reasons other than the monopolist's input price.

For simplicity, suppose the monopolist can make take-it-or-leave-it offers. Our benchmark commitment solution is the (subgame-perfect) equilibrium to the following "commitment game":

stage 1 (offers): The monopolist publicly announces a set of offers, one for each firm: \( \{r_i, f_i\}, i = 1, \ldots, n \).

stage 2 (acceptances): Firms accept or reject offers simultaneously. Accepting means paying the fixed fee.

stage 3 (learning): Accepted contracts are announced; firms learn others' marginal costs.

stage 4 (competition): Firms simultaneously

(i) set their downstream instruments, prices, or outputs; and

\(^5\) Variants of this result have been established by Hart and Tirole (1990), DeGraba and Postlewaite (1992), and Daniel P. O'Brien and Greg Shaffer (1992) (all independently, to our knowledge). Their contributions are more conveniently compared in Section III. These authors, like us, abstract from issues of risk-sharing; Mathias Dewatripont and Khalid Sekkat (1991) present a model in which threat of opportunism against an established dealer can alter the renegotiation equilibrium in a way that improves risk-sharing.
(ii) purchase the necessary amounts of the monopolist’s input.

Consider stage 4. We assume throughout the paper that for any vector of n input prices \( r = (r_1, \ldots, r_n) \) accepted in stage 2 and learned by all in stage 3, there is a unique noncooperative equilibrium to the competition in stage 4, with firm \( i \)'s indirect equilibrium-profit function denoted \( \pi_i(r) \). (Idle firms are allowed in this formulation; an idle firm is offered \( f = 0 \), \( R = \infty \), accepts, and produces zero.) Equilibrium-profit functions are assumed to display the normal properties: an active firm’s gross profit decreases in its own input price and increases in the price to an active rival: \( \partial \pi_i(r)/\partial r_i < 0 \), \( \partial \pi_i(r)/\partial r_k > 0 \), \( k \neq i \). A firm’s input-demand function, incorporating the downstream competition whether in prices or outputs, is denoted \( q_i(r) \).

In stage 2 of the game, if \( f_i \leq \pi_i(r) \) then it is a (weakly) dominant strategy for firm \( i \) to accept its contract: if a rival that might have been active rejects its contract then, by assumption, firm \( i \)'s profit will be at least \( \pi_i(r) \).

In stage 1, the monopolist therefore sets \( f_i = \pi_i(r) \); hence, its objective in the commitment game is to choose \( r \) so as to maximize overall profit,

\[
\text{(1) } G(r) = \sum_{i=1}^{n} (r_i - z)q_i(r) + \sum_{i=1}^{n} \pi_i(r).
\]

Let \( G* \) denote the maximum profit and \( r* \) any maximizing vector. Thus, in the commitment game the monopolist’s profit is \( V = G* \). Since firms know rivals’ proposed costs before accepting their own contracts, cutting price to one firm reduces the maximum fees that others will pay; hence, the monopolist chooses prices to induce the maximum overall profit, \( G* \). \( G* \) will serve as our benchmark against which to compare outcomes under various no-commitment regimes.

To see clearly the opportunism incentive, consider an alternative sequential game in which the monopolist does not commit at the outset to all contracts. In stage 1 (offers), the monopolist approaches firms 1, \ldots, n sequentially. Each firm accepts or rejects its offer having observed all prior offers and decisions. Once all firms have chosen, downstream competition occurs.\(^6\) It will prove useful to define the term \( u_i \), the monopolist’s net revenue from input sales plus the profit of firm \( i \):

\[
\text{(2) } u_i(r) = \sum_{j=1}^{n} (r_j - z)q_j(r) + \pi_i(r)
= G(r) - \sum_{k \neq i} \pi_k(r).
\]

Observe that \( u_i \) only depends on the marginal prices \( r \) and not on the fixed fees.

**PROPOSITION 1**: If attaining \( G* \) in the commitment game requires more than one firm to be active, then in the sequential game the monopolist’s profit is \( V < G* \).

(Proofs of all propositions can be found in the Appendix.)

\(^6\) This game is chosen primarily for transparency. Stepping beyond the model, however, the failure to commit initially to all prices can be reconciled with the implicit assumption of complete information. When contracts are multidimensional (specifying delivery terms, promotion allowances, etc.), different contracts can yield the same implicit price. The monopolist may wish to preserve the flexibility of tailoring the details of a second firm’s contract to fit that firm’s attributes (e.g., whether it is proficient at doing its own delivery), attributes which may be unknown when contracting with the first firm. If courts are unable to value different dimensions of a contract, however, and therefore are unable to infer the implicit “total” price in a given contract, the monopolist would have to specify the details of the second firm’s contract prematurely if it wished to commit to the first firm regarding the price that will be offered to the second firm. Thus, although “insiders” may have a good idea of the total implicit prices that will be offered to later movers, the monopolist may nevertheless find it costly to commit to these prices.

That courts indeed have serious difficulty evaluating different contracts is vividly illustrated by the U.S. experience with its main price-discrimination law, the Robinson-Patman Act. It has spawned voluminous litigation over whether a given price cut is a selective discount or is “cost-justified” because the buyer undertakes some functions otherwise performed by the supplier (ABA, 1980; Schwartz, 1986).
The intuition for Proposition 1 is straightforward. Having collected fixed fees from all but the last mover, the monopolist sets the last price to maximize joint profit with the last mover. It ignores the reduction in others’ profits from cutting price to the last firm, an effect internalized when computing \( r^* \). This externality will be present whenever firms earn quasi-rents in equilibrium; thus, under quite general assumptions about the downstream competition, the monopolist wants to cut price to the last firm in exchange for a higher fixed fee. Overall profit thus will be less than \( G^* \); given foresight by firms when computing their acceptable fixed fees, the monopolist’s profit is less than \( G^* \).\(^7\)

II. Nondiscrimination Clauses

As noted previously, an important drawback of commitment stems from loss of flexibility to cope efficiently with exogenous changes, given that complete contracts are infeasible. To fix ideas, suppose that firms can observe all contracts and that incompleteness of contracts is due to the cost of specifying contingencies and difficulties with verifiability.\(^8\) How might early buyers be assured that flexibility will be used only to make efficient changes and not to recontract opportunistically with future buyers?

Giving each firm veto power over changes to any other contracts would create serious holdup problems; it may also raise antitrust issues (Philip F. Zeidman, 1991). A less drastic option is to let a firm “back out” of its contract if it disapproves of any new contract offered by the monopolist. This option, too, is problematic. A firm could threaten to back out simply as blackmail; perhaps more importantly, the firm’s ability to back out profitably may be limited by its other investments in relationship-specific assets. A seemingly natural device for guarding against opportunism while preserving some flexibility is nondiscrimination (or most-favored-customer) clauses, entitling each firm to replace its initially accepted contract with any other contract later accepted by a rival.\(^9\)

Consider the following nondiscrimination game. Initial contracting is the same as in the sequential game, but before downstream competition occurs, there now is a recontracting stage. All contracts accepted in stage 1 become part of the new menu, and the monopolist approaches all firms that have accepted contracts, sequentially and in reverse order, letting each firm exchange its previously accepted contract for any new one. Once recontracting is com-

\(^7\)If one downstream firm suffices to attain \( G^* \) under commitment (e.g., homogeneous products and natural monopoly downstream) then the monopolist could attain \( G^* \) also in the sequential game, by offering \( r = \infty, f = 0 \) to all but the designated firm and \( r = z \) to it, the last mover. This possibility, however, is an artifact of the sequential game and is relaxed in Section III, where there is no last mover.

\(^8\)For example, the monopolist may wish to preserve the option of reducing price in the future if its marginal cost falls, but it cannot guarantee contractually that price will be lowered only in such circumstances, because courts cannot verify the reason for a price decrease.

\(^9\)To enforce such nondiscrimination, courts must be able to verify that a given contract is offered to others. Nondiscrimination clauses therefore are informationally infeasible if courts cannot verify parties’ performances (as in Sanford J. Grossman and Hart [1986]). However, such clauses potentially do serve a role if courts can verify performance, but efficient price commitments are precluded because contracts are incomplete for other reasons (e.g., courts’ inability to verify the state of nature—why was price cut, etc.).
pleted, downstream competition occurs as in the sequential game.\textsuperscript{10}

The following property will prove useful both here and in Section III.

**Definition:** A set of contracts \( \{r_k, f_k\}, k = 1, \ldots, n \), is pairwise-proof if \( u_i(r_j, r_{-i}) \geq u_i(r'_j, r_{-i}) \) for any \( r'_j \), and every firm \( i, i = 1, \ldots, n \), where the first entry always denotes the input price to firm \( i \), \( r_{-i} \) denotes the vector of input prices charged to other firms, and \( u_i \) is defined in equation (2).

That is, a set of contracts is pairwise-proof if, holding all other contracts fixed, any firm \( i \) and the monopolist cannot increase their combined profit (ignoring all fixed fees already extracted from other firms \( j \neq i \)) by changing the marginal price to firm \( i \). It is not obvious that pairwise-proofness is relevant in the nondiscrimination game, since offering a new contract to one firm entitles all firms to change their contract. Proposition 2 below, therefore, has some bite. We make the following assumptions.

**ASSUMPTION 1:** Firms are symmetric: \( \pi_i(r_i, r_{-i}) = \pi(r_i, r_{-i}), i = 1, \ldots, n \).

**ASSUMPTION 2:** In the commitment game, attaining \( G^* \) requires two or more firms to be active and to face the same input price.

\textsuperscript{10}The assumption of reverse-order recontracting by all is only for simplicity. As will be discussed shortly, the key feature for Proposition 2 is that the last firm to accept a contract in the contracting stage moves first in the recontracting stage and thus can commit to stay at the contract it accepts in the first round.

While the nondiscrimination game is chosen mainly for its simplicity, there do exist informational conditions corresponding to this game, under which nondiscrimination is feasible yet efficient commitment at the outset is not. When contracts are multidimensional (see footnote 6), committing to a “price” requires specifying in advance all contract dimensions. If courts cannot value such attributes, committing to a total price would require specifying inefficiently at the outset all the details of these multidimensional contracts. Enforcing nondiscrimination requires only that courts be able to verify that a particular contract was offered to all.

**ASSUMPTION 3:** \( \frac{\partial^2 \pi_i}{\partial r_k \partial r_j} < 0 \) for all active firms \( j, k, j \neq k \).

Assumption 1, symmetry, focuses attention on the more interesting case. It would not be surprising if nondiscrimination clauses had limited effectiveness when firms were asymmetric; with asymmetries, a common menu of contracts could be used to induce agents to self-select. Given the assumed symmetry, Assumption 2 also is natural when the downstream industry is not a natural monopoly (e.g., because firms have increasing marginal costs or produce symmetrically differentiated products). Attaining \( G^* \) need not require all \( n \) firms to be active if there are fixed costs.

Assumption 3 says that a decrease in a firm’s marginal cost is less valuable to it the lower a rival’s marginal cost. This property is satisfied in many standard models. Roughly speaking, a lower marginal cost to a rival makes the rival more aggressive in any equilibrium (produces a higher output or charges a lower price). The initial firm can thus be expected to produce a lower output; hence the value to it of a given reduction in marginal cost also can be expected to be lower.\textsuperscript{11} Under Assumptions 1–3, we have the following proposition:

**PROPOSITION 2:** If there exists a symmetric equilibrium to the nondiscrimination game in which two or more firms are active, each receiving the contract \( (r_0, f_0) \), then \( r_0 \) must be pairwise-proof. Therefore, the monopo-

\textsuperscript{11}We cannot be definitive, because demand might be such that (i) a reduction in the firm’s own cost would lead to a larger contraction in the rival’s planned output when a rival’s cost is low rather than high, and (ii) this strategic effect might be sufficiently strong to outweigh the fact that the direct benefit from a cost reduction is smaller when a rival’s cost is low (because the firm plans to produce less). The property must therefore be verified for specific models. It holds, for example, if firms have constant marginal costs additive in \( r \) (as under fixed-proportions production) and compete Cournot in homogeneous products with inverse demand \( p = a - bQ \), \( p = a - b \ln Q \), or \( p = a + b / Q \) (Jennifer F. Reinganum, 1983). For the same costs, it also holds if firms compete Bertrand with differentiated products and a linear demand system.
list’s profit in the nondiscrimination game is $V < G^*$.

This result is understood as follows. Suppose $r_i$ is not pairwise-proof, for concreteness, suppose the joint profit of the monopolist and any firm would increase if that firm alone accepted a deviation $r_i < r_0$. Consider firm $k$, $k \leq n$, the last firm that accepts $(r_0, f_0)$ in stage 1 along the candidate-equilibrium path. Suppose the monopolist offered firm $k$ instead $(r_1, f_1)$, where the fixed fee $f_1 > f_0$ is chosen such that firm $k$ finds it just profitable to switch from $(r_0, f_0)$ to $(r_1, f_1)$, provided nobody else switches in the recontracting stage. Given $\frac{\partial^2 \pi_i}{\partial r_i \partial r_i} < 0$, once firm $k$ accepts the deviation, firms $j < k$ indeed would not switch; although they lose from the price reduction to firm $k$, they would lose even more by paying $f_1 > f_0$ in exchange for the same price reduction. Anticipating this, firm $k$ would accept the deviation, breaking the candidate equilibrium. Thus, a symmetric equilibrium contract $(r_0, f_0)$ must be pairwise-proof.

With two or more firms active, however, pairwise-proofness of a vector $r$ implies that $r$ does not attain $G^*$. This is simply the flip side of the logic used in Proposition 1 to show that the commitment solution is not pairwise-proof with respect to the last mover. Given Assumption 2, it follows that the monopolist cannot attain $G^*$, whatever is the equilibrium to the nondiscrimination game.  

Proposition 2 contrasts with the findings of DeGraba and Postlewaite (1992), who show that a commitment not to discriminate restores the optimum when the monopolist essentially charges only fixed fees (in their model, each firm buys at most one unit of the input). There is a basic difference between using nondiscrimination clauses for one-dimensional contracts (linear prices or fixed fees) and for other contracts.

Typically, a firm unconditionally prefers a lower price or lower fixed fee; the monopolist thus recognizes that, with nondiscrimination clauses, any discount offered to later buyers would be demanded by all previous buyers. This dominance property generally does not hold for two-part tariffs or for other contracts with a fixed-fee component (e.g., agreeing to buy a minimum level of inputs at a given price). The value to a firm of a cut in its marginal cost (hence, the maximum additional fixed fee it would pay to obtain this cut) is typically less than the lower are rivals’ marginal costs. Nondiscrimination clauses thus may have no bite—because only one firm will accept a deviation contract that offers a low marginal cost, provided the average cost is sufficiently high; other firms will elect not to exercise the option of exchanging their contracts for this new contract.

Note that we interpret nondiscrimination as allowing the seller to offer contracts with different marginal prices, provided the same menu of contracts is offered to all. U.S. price-discrimination law is somewhat ambiguous on this point, but it is tilted toward our interpretation. Interestingly, courts recognize that a menu of universally “available” contracts can be designed to induce

---

12If firms had the option of backing out of their contracts altogether (with the fixed fees refunded), the monopolist would be deterred from offering the deviation. As noted earlier, there are serious problems (not modeled explicitly here) with that approach. Note also that, since firm $k$ moves first in the recontracting stage, other firms take $k$’s choice as given. If firm $k$ could not “commit” to staying at this deviation contract (e.g., because it, too, could later recontract), it may be reluctant to accept a deviation initially for fear that, in the ensuing subgame, another firm (and not it) would ultimately end up with this contract. Such a possibility can be ruled out in various ways (e.g., by having firm $k$ voluntarily accept a contract without a nondiscrimination clause and containing a penalty for breach by the firm).

13An equilibrium, whether symmetric or not, will exist, since the game is one of perfect information.

14The ensuing discussion of U.S. law draws on ABA (1980, 1991) and McAfee and Schwartz (1992). The Robinson-Patman Act (the chief federal statute governing price discrimination) was interpreted in a 1960 Supreme Court decision (FTC v. Anheuser Busch, Inc.) to find price discrimination whenever there is “merely a difference in price”; since then, however, courts and enforcement agencies (primarily the Federal Trade Commission) have generally been willing to consider as a defense the fact that the challenged offer, while entailing a different marginal price, was made available to all competing buyers.
self-selection when buyers are heterogeneous, and the courts have conditioned the availability defense on offers being "functionally" (i.e., practically) available to all.\textsuperscript{15} This functional-availability test, however, seems directed at preventing sellers from exploiting inherent asymmetries among customers; price discounts have been upheld when tied to "reasonable" conditions (e.g., to stocking inventory or purchasing acceptable minimum quantities). Our analysis suggests that the meaning of "practical availability" and "nondiscrimination" in input markets is rather subtle. Even where buyers are symmetric, offering the same menu of contracts to all can entail discrimination, in that a second firm may rationally choose to reject an offer once a competitor has accepted.\textsuperscript{16}

III. Secret Offers

The sequential game in Section I illustrates the opportunism incentive but is somewhat artificial. Firms take as given the monopolist’s contracts with earlier movers and fear only its dealings with later movers; for example, the last mover has nothing to fear. More plausibly, all firms will be leery that the monopolist might react with their rivals. In this section we consider other representations of contracting without commitment, to establish some robustness for the result that the monopolist is harmed by its anticipated opportunism. The models presented also facilitate discussion of related work by other authors.

In the commitment game, when accepting its contract, a firm knows that rivals’ future costs will not be lowered. Rather than modeling recommitting explicitly, we represent lack of commitment by treating offers as made simultaneously and secretly. This captures the key idea that each firm is not confident about rivals’ future costs when it must invest the fixed fee. We have studied two simple games with secret offers.

Here we focus on the unobservability game: offers are simultaneous and secret, and firms never learn others’ contracts. Accepting firms then compete as in stage 4 of the commitment game. (In pure-strategy equilibrium, of course, each firm holds correct beliefs about others’ costs when competing downstream.) This game reflects cases in which commitment is difficult because of the possibility of secret discounts.

Our 1992 paper analyzes a second game: ex post observability. That game is identical to unobservability, except that firms learn rivals’ contracts after paying their fixed fees, but before competing downstream.\textsuperscript{17} That game reflects situations where commitment is problematic even if firms can eventually observe others’ contracts, because of the usual obstacles to complete (efficient) contracts. To retain flexibility, the monopolist thus may assure each firm only about the terms of its contract (subject to bilaterally acceptable renegotiation). Initial offers, even if public, would then convey little about what rivals’ costs ultimately will be. We represent this by having firms learn rivals’ contracts only after paying the fixed fees.\textsuperscript{18}

\textsuperscript{15}In a 1948 landmark case (FTC v. Morton Salt Co.) the Supreme Court struck down a quantity discount schedule in which only the large chain buyers could realistically qualify for the top discounts.

\textsuperscript{16}This observation is timely, as enforcement agencies have been increasingly sympathetic to a "practical availability" defense. In 1977 the U.S. Department of Justice proposed amending the Robinson-Patman Act to incorporate such a defense explicitly. More recently, the Canadian Bureau of Competition Policy (1992) issued enforcement guidelines on price discrimination, stressing the role of practical availability as a defense.

\textsuperscript{17}The strategies in the two games, therefore, are as follows. In both games, the monopolist makes a set of n offers, one to each firm. A firm (i) accepts (pays the fixed fee) or rejects, knowing only the offer it receives and (ii) chooses its input order and the level of its downstream variable (price or output), knowing only its own costs under unobservability, and all costs under ex post observability.

\textsuperscript{18}Taken literally, this deterministic game suppresses an obvious remedy to opportunism: to specify all prices in each firm’s contract and collect the fixed fees later only if the monopolist has adhered to these contracts.
We focus here on unobservability and report briefly our results for \textit{ex post} observability.

Throughout, we consider only \textit{pure}-strategy, perfect Bayesian-Nash equilibria. In both games, there are multiple equilibria. When a firm receives an off-equilibrium offer, there is considerable latitude in how it might revise its beliefs about the secret offers made to others, and therefore in how it will react to its offer. We study three sets of beliefs: symmetry, passive, and wary. To concentrate on beliefs, we simplify by assuming that firms are symmetric and that under commitment the maximum profit $G^*$ could be attained in a symmetric equilibrium by offering all firms the same price $r^*$. We begin with symmetry beliefs.

\textbf{A. Symmetry Beliefs}

\textit{Remark:} Given symmetry beliefs, the equilibrium outcome to both the unobservability game and the \textit{ex post} observability game is $V = G^*$. This equivalence is easily seen. By assumption, in the commitment game if all firms were offered $(r^*, \pi(r^*))$ they would accept, and the monopolist would earn $V(r^*) = G^*$. Under symmetry beliefs, each firm believes that all others receive the same offer as it receives; hence all would accept $(r^*, \pi(r^*))$ under either unobservability or \textit{ex post} observability as well.\footnote{The monopolist cannot gain by deviating and offering $r < r^*$ to any firm. Under symmetry beliefs (and identical firms), the highest fixed fee a firm would pay equals $1/n$ of the profit that the industry would earn if all firms accepted the same price offered to that firm. A firm offered $r < r^*$ would believe that all others also received this offer and hence would not be willing to pay as high a fixed fee as it would if it thought the discount was to it alone (as the last mover knows in the sequential game).} Symmetry beliefs thus illustrate that the optimum might be attainable even without commitment. Such beliefs, however, are not very compelling: the monopolist's preferred contract with other firms generally differs from the contract that the first firm has accepted.

One rationale for symmetry beliefs is that firms interpret unexpected offers as trembles (mistakes) by the monopolist and assume that trembles are perfectly correlated (e.g., because the monopolist has miscalculated $r^*$). A polar, arguably more plausible, assumption is that firms believe trembles to be uncorrelated (e.g., because the monopolist appoints a different agent to deal with each firm). This leads us to passive beliefs.

\textbf{B. Passive Beliefs}

Under passive beliefs, when a firm receives an offer different from what it expects in the candidate equilibrium, it does not revise its beliefs about the offers made to others. Passive beliefs have been invoked, explicitly or implicitly, by other authors dealing with a single-principal and multiple-agents framework, including Jacques Cremer and Michael H. Riordan (1987), Henrick Horn and Asher Wolinsky (1988), Hart and Tirole (1990), and O'Brien and Shaffer (1992). (See our 1992 paper for details.)
Given passive beliefs, equilibrium contracts in either game must be pairwise-proof (see Section II): \( u_i(r_i, r_{-i}) \geq u_i(r'_i, r_{-i}) \) for any \( r'_i, i = 1, \ldots, n \).

If the inequality failed for any firm, the monopolist could offer a bilaterally profitable deviation at the offers stage, and the firm would accept given passive beliefs. We now trace some implications of this requirement that the equilibrium be immune to bilateral recontracting.

**Proposition 3:** Given passive beliefs, in the unobservability game, the equilibrium outcome is for the monopolist to charge all firms a price equal to its marginal cost: \( r_i = z, i = 1, \ldots, n \). All firms are active, and the monopolist earns \( V = n\pi(z) \).

The underlying intuition is as follows. Since a firm’s decisions (input purchases and choice of downstream price or output) are unaffected by the unobserved changes in input prices to rivals, in its dealings with any firm the monopolist acts as if the two are integrated and face a given residual downstream demand. Pairwise maximization then involves setting input price equal to the monopolist’s marginal cost. This result is quite general: it does not hinge on the nature of downstream production (fixed versus variable proportions) or of downstream competition (strategic substitutes or strategic complements).

Hart and Tirole (1990) show, in a restricted environment, that the same downstream outcome emerges when the monopolist can employ more general contracts than two-part tariffs. Their result is driven by

\[ A \text{ remark on notation: in the ex post observability game, } u_i(r_i, r_{-i}) \text{ is evaluated assuming that all firms observe the input prices } (r'_i, r_{-i}) \text{ before making downstream choices. In the unobservability game, } u_i(r_i, r_{-i}) \text{ is evaluated assuming that firm } i \text{ continues to believe that others accept their candidate equilibrium offers } r_{-i} \text{ (by passive beliefs); firms } j \neq i \text{, which do not observe the deviation offer to firm } i, \text{ continue believing } (r_i, r_{-i}). \text{ With this slight abuse of notation, we use the same expression to describe the pairwise-proof conditions in both games.} \]

\[ B \text{ The monopolist’s instruments then are no longer the marginal prices, but the targets are set at the same levels; hence profit is the same. Hart and Tirole (1990) are only tangentially concerned with opportunism; their main interest is in foreclosure possibilities when potentially there are competing parties at both levels. However, in a setup similar to our unobservability game with passive beliefs—and under what amounts to Cournot duopoly downstream, and assuming fixed-proportions technology downstream—their proposition 1 implies that a monopolist free to use any pricing scheme earns only the same profit as if it charged all firms a two-part tariff with the marginal price equal to its marginal cost.} \]

\[ C \text{ DeGraba and Postlewaite (1992) consider an input monopolist selling sequentially to identical competing firms, each demanding at most one unit of the input (e.g., a machine). They show that input price is driven to marginal cost if the monopolist cannot commit to restrict the number of buyers to whom it sells.} \]

\[ D \text{ O’Brien and Shaffer (1992) assume Bertrand competition downstream with differentiated products (possibly asymmetric) and allow the monopolist to charge general fee schedules as a function of a firm’s realized level of sales. (This is similar to inputs being purchased after sales in our unobservability game, since granting a discount to one firm would affect the monopolist’s revenue from another.) They show that in “contract equilibrium” (one that is pairwise-proof, as is also mandated by imposing our passive beliefs) the marginal charges to all firms are set equal to the monopolist’s marginal cost.} \]
A corollary of Proposition 3 is that under unobservability the monopolist earns \( V < G^* \), since to attain \( G^* \) the monopolist would have to set input prices above marginal cost, in order to counteract the negative competitive externality existing whenever there are \( n \geq 2 \) firms downstream. Indeed, under unobservability the monopolist’s profit can be driven to zero. Since it prices inputs at (constant) marginal cost, the monopolist’s profit accrues entirely from the fixed fees that collect downstream profits. Given sufficient competition downstream (e.g., many homogeneous-product Cournot firms or two Bertrand firms offering close substitutes), downstream profit will be arbitrarily small, and with it the monopolist’s profit.

Ex post observability adds a new element: by changing the price to any firm \( i \), the monopolist can now alter other firms’ subsequent downstream choices to raise firm \( i \)’s expected profit and thus its fixed fee. Pairwise-proof input prices thus will no longer equal marginal cost \( z \). Rather, the incentive is to set \( r > z \) if firms’ downstream instruments are strategic complements, and \( r < z \) if strategic substitutes, by standard reasoning. (See Tirole [1988] or Carl Shapiro [1989] for discussions of incentives in two-stage games.) In either case, if a symmetric equilibrium exists under passive beliefs, then its outcome will be \( V < G^* \). This is so because when maximizing bilateral profit \( u_i \), with any firm \( i \), the monopolist still ignores the direct profit reduction to rivals from cutting price to firm \( i \) (even under ex post observability, it cares only about rivals’ responses to \( i \)’s change in cost, not the reduction in their profits per se).

In the case of strategic substitutes, however, an equilibrium in two-part tariffs need not exist. In our 1992 paper we consider homogeneous-products Cournot competition downstream among \( n \) identical firms with cost functions \( C_i(q_i) = r_i q_i \) (a firm’s marginal cost is its input price). There exists no (pure-strategy) equilibrium in “normal” environments (e.g., \( n = 2 \) and linear demand). Starting at the unobservability solution of \( r = z \), it now pays to set \( r < z \) to any firm \( i \) in order to induce firms \( j \neq i \) to contract outputs and thus raise firm \( i \)’s profit and fixed fee. Once \( r < z \), there is a new source of gain from reducing \( r \) to any firm: this induces contraction of below-cost input sales to rivals.\(^{24}\) The last effect can be strong enough to push the input prices that satisfy pairwise-proofness (\( \partial u_i / \partial r = 0 \)) so far below the monopolist’s cost that the resulting output price would be below combined costs, implying negative overall profit and hence no equilibrium.\(^{25}\)

C. Wary Beliefs

Symmetry beliefs and passive beliefs are based on the notion that firms view unexpected offers as trembles by the monopolist. Here we treat firms instead as somewhat suspicious and assume that they interpret any offer as a deliberate choice by the monopolist. Specifically, under our wary beliefs each firm thinks that others received offers that are the monopolist’s optimal choices given the offer it made to that firm.

We illustrate these beliefs for the simple case of two symmetric downstream firms. Let \( \pi(r, s) \) and \( q(r, s) \) denote a firm’s profit and input demand, respectively, when the firm gets input price \( r \) and the rival gets \( s \). A firm’s equilibrium strategy will be summa-

\(^{24}\) It is this effect that differentiates our model from those of authors who found similar incentives for \textit{oligopolist} manufacturers to signal toughness by, in effect, reducing downstream marginal costs when downstream competition is Cournot (e.g., James A. Brander and Barbara J. Spencer, 1985; John S. Vickers, 1985; Chaim Fershtman and Kenneth L. Judd, 1987). In those models, an equilibrium typically exists. Here the monopolist contracts with multiple downstream parties, and its profit depends on input sales to all. This multilateral contracting, coupled with passive beliefs, produces the stronger incentive to cut input prices, leading to nonexistence.

\(^{25}\) As pointed out to us by the referees, the nonexistence problem is even worse than suggested by this pairwise-proofness argument, since an equilibrium must also be immune to \textit{multilateral} deviations—offers which the monopolist can profitably make to several downstream firms and which the latter would accept.
rized by a set A denoting all contracts that a firm will accept. The set A will be optimal given the firm's equilibrium belief function \( R(r) \) about the price that the monopolist offers the rival if the firm is offered \( r \). If offered a contract \((r, f) \in A\), firm \( i \) believes that:

(a) the monopolist expected it to accept the offer;
(b) the monopolist offers firm \( k \) the contract \((R(r), F(r))\) that is best for the monopolist, given that firm \( i \) accepts \((r, f)\), from among all contracts acceptable to firm \( k \);
(c) firm \( k \) reasons the same way.

Given these beliefs, a firm accepts a contract if and only if it is in the set \( A = \{(r, f) : f \leq \pi(r, R(r))\} \), a strategy that clearly is optimal given its beliefs. The monopolist's equilibrium strategy is to choose from the set A its profit-maximizing contracts, one for each firm. In a symmetric equilibrium, firms get the same contract \((r^e, f^e)\), \( f^e = \pi(r^e, r^e) \) with \( R(r^e) = r^e \). We obtain the following results (see Appendix).

Symmetry beliefs, which were shown to support \( G^* \) under unobservability or \( \text{ex post} \) observability, are not wary beliefs in either game. Under unobservability, the equilibrium outcome with wary beliefs is the same as with passive beliefs: the monopolist offers all firms \( r = z \). The intuition is that changing the contract to one firm does not affect how much the monopolist collects from another under unobservability. Wary beliefs then dictate that each firm must expect the monopolist to behave as if integrated with the other firm and therefore must expect marginal-cost pricing to the rival regardless of the offer it receives. Moreover, the expectation of "integrated behavior" (and thus the downstream outcome) generalizes to arbitrary nonlinear contracts. Under \( \text{ex post} \) observability, we only succeeded in analyzing a Cournot example. Whereas no equilibrium exists for that environment under passive beliefs, under wary beliefs an equilibrium exists and displays opportunism: the monopolist sets price below the commitment level.

IV. Conclusion

We began by illustrating the opportunism problem in a game with sequential contracting. We then modified that game to incorporate nondiscrimination (most-favored-customer) clauses, and we showed that they do not generally restore the commitment solution, even in symmetric environments. This inability of nondiscrimination clauses to curb opportunism in multilateral vertical contracting is our most novel result.

To relax the artificial last-mover feature of the sequential game, we considered symmetric firms receiving simultaneous and secret offers. We examined three alternative beliefs of firms about others’ offers, and two games (unobservability and \( \text{ex post} \) observability) distinguished by whether each firm either never learns others' contracts or does so after accepting its own contract. Only when firms hold symmetry beliefs (each firm believes all others always get the same offer as it) does the commitment solution emerge as the outcome.

For passive beliefs (firms receiving off-equilibrium offers do not revise their beliefs about offers made to others), under unobservability the monopolist sets price equal to marginal cost to all firms, whatever the downstream competition. This result has been independently shown by others. Under \( \text{ex post} \) observability, we identified a frequent nonexistence of equilibrium when the downstream competition is Cournot.

Finally, we introduced a new restriction on beliefs, termed wary beliefs. The assumptions underlying these beliefs seem natural: a firm expects that the monopolist offers to the other firm the monopolist’s

---

26 We consider point beliefs because they are simple, and because we focus on pure strategies.
27 If offered a contract not in \( A \), the firm believes that the monopolist offers to the rival the optimal contract conditional on only the rival being in the market. Without loss of generality we can confine attention to contracts that are acceptable [since instead of a contract that will be rejected, the monopolist could offer \((\infty, 0)\), which is trivially accepted].
best response to the initial offer. Wary beliefs yield the same outcome as passive beliefs under unobservability. That outcome also seems natural for that environment: since other firms’ choices are unaffected by the monopolist’s secret discounts, the monopolist acts with each firm as though vertically integrated with that firm alone. This exemplifies the monopolist’s desire to renegotiate with each firm individually, undercutting the commitment solution. Under \textit{ex post} observability, the analysis of wary-beliefs equilibria is quite difficult and is an interesting topic for future research.

The general thrust of our analysis is that fears of opportunistic recontracting harm the monopolist if it fails to commit. One resolution, also noted by Hart and Tirole (1990), DeGraba and Postlewaite (1992), and O’Brien and Shaffer (1992), is vertical integration. Integration, however, can entail its own inefficiencies; thus, the issue remains of how to assure independent firms against opportunism. We noted in the Introduction that commitment to multilateral contracts is costly, because of the difficulties of writing and verifying efficient state-contingent contracts. The prospective losses from opportunism, however, suggest that some crude commitment will be undertaken despite its attendant inefficiencies, to reassure customers against opportunism. We now turn to some vertical practices that may play this role. Most available data are on franchise contracts, from which our evidence below is primarily drawn.\textsuperscript{29}

An approach we find informationally plausible is for the monopolist to deal with only one firm in a given market. A commitment still is required; as noted below, this commitment problem is nontrivial. Nevertheless, it is surely easier to convince a court that a significant new downstream competitor has been brought in than to delve into the details of whether and why a particular price cut was offered to an existing competitor.\textsuperscript{30} Our analysis therefore may explain the adoption of practices such as exclusive territories even when there is no problem of free-riding on downstream services, and even when efficiency would dictate against a downstream monopoly (e.g., because downstream firms face increasing costs or offer differentiated products).\textsuperscript{31} In fact, exclusive territories are widely used. \textit{Entrepreneur Magazine}’s (1987) survey, the latest that offers data about exclusive territories, reveals that of the 139 top franchisors in various businesses, 93 granted exclusive territory franchises.

Exclusive territories, however, are not a panacea. Granting long-term exclusivity can be inefficient, especially if demand is expected to grow. Moreover, as demonstrated

\textsuperscript{28}O’Brien and Shaffer (1992) also suggest another solution: eliminating downstream margins through maximum resale-price maintenance (RPM), thus leaving no downstream profit vulnerable to opportunism. However, when there is moral hazard downstream, it is important to preserve marginal-profit incentives. Also, RPM is per se illegal in the United States.

\textsuperscript{29}Hadfield (1990) reports estimates by the U.S. Department of Commerce that franchise sales in 1987 were $591 billion, or about one-third of all retail sales in the United States. The Commerce Department distinguishes between “traditional franchising” or product franchising, involving franchised dealers as in automobile and gasoline retailing, and “business-format franchising,” as in fast foods. The latter accounted for about 32 percent (or $191 billion) of all franchise sales. Lafontaine (1993) notes that in traditional franchising, franchisors derive revenue primarily from input markups that are unobservable to outsiders, while business-format franchising involves royalties and franchise fees for which data are more available (e.g., from disclosure statements). Consequently, most empirical work has concentrated on business-format franchising.

\textsuperscript{30}Patrick Rey and Tirole (1986) suggest a similar observability advantage of uniform two-part tariffs over more sophisticated contracts. It is easier to convince a court that a dealer is carrying a manufacturer’s product (beyond some minimum threshold, to be labeled a “dealer”) than to monitor the exact sales of the dealer. Thus, a manufacturer might be able to extract a uniform franchise fee from all dealers but not be able to implement more sophisticated pricing schemes.

\textsuperscript{31}In a different setting, Andrea Shepherd (1987) and Joseph Farrell and Nancy T. Gallini (1988) invoke a similar contracting-cost argument to explain why an initial monopolist may choose to license competing suppliers when seemingly this reduces its ability to extract maximal downstream rents. They consider a monopolist facing prospective buyers that must make product-specific investments and argue that the monopolist would find it difficult to commit contractually to low future prices or high future quality. To assure buyers against opportunism, the monopolist commits by licensing additional suppliers.
by numerous suits charging violation of exclusive territories, the protection offered by territories is far from ironclad. It is not easy to determine who is a competitor when customers are mobile, or when the franchisor can offer products slightly differentiated from the franchisee's or even the same product through different types of outlets (ABA, 1990).

When exclusivity arrangements fail to insulate firms from competition, our analysis suggests that a supplier's ability to alter the set of contracts it offers opens the door for opportunistic behavior. Thus, franchisors may seek to reassure franchisees by following a consistent policy of making contracts simple, transparent, and uniform across franchisees, and rigid over time. Moreover, uniformity and rigidity need not be substitutes for reputation as commitment devices; they can be complements. For a reputation mechanism to work, outsiders must be able to discern whether a change was efficient or opportunistic. This will be more difficult if terms are highly complex and volatile.32

The simplicity and uniformity of franchise contracts is striking.33 A given franchisor offers remarkably similar terms across franchisees. Lafontaine's (1992a) survey shows that all 126 franchisors responding to the question said they offered the same franchise contract to all potential franchisees at a given point in time. Moreover, 104 of them said that the offer was "take-it-or-leave-it" or that they might negotiate only on nonmonetary clauses. Significantly, variation in royalty rates was particularly low. Of the 55 firms that also submitted franchise disclosure statements, 22 used a single franchise fee, while 45 employed a constant royalty rate (uniform across franchisees) at all levels of sales.34 This conforms with our theory, since a rival's franchise fee is a fixed cost, whereas the royalty rate affects the rival's aggressiveness as a competitor. Thus, a firm will be more concerned about the rival's royalty rate than about the fixed fee.35

Such uniformity in contract terms seems inefficient given the likely heterogeneity of potential franchisees and market circumstances. As Lafontaine (1992b) notes, various agency theories predict that the franchisor should offer a menu of contracts featuring different royalty rates. Indeed, when asked in Lafontaine's (1992a) survey what might be the disadvantages of using the same contract to all franchisees, only 23 of 88 respondents claimed no disadvantage; the rest noted disadvantages from the loss of flexibility.36

This uniformity in franchise contracts does not seem entirely or even primarily

32 Milgrom and Roberts (1992) discuss similar themes in the context of internal firm organization, stressing the value of simple and rigid procedures given bounded rationality and imperfect monitoring.

33 Across different franchisors, there is considerable variation in contract terms. In a sample of fast-food firms alone, Banerji and Simon (1992) find that the franchise fees (in thousands of 1980 dollars) range between 4 and 37 with a mean of 22.15 and a standard deviation of 8.32. Royalty rates range from 0 to 6 percent of sales (i.e., revenue), with a mean of 4 percent and a standard deviation of 1.22 percent.

34 This information, not reported in Lafontaine (1992a), is from Sugato Bhattacharyya and Lafontaine (1992). Lafontaine (1993) notes that the use of different franchise fees by 22 of the 55 respondents is consistent with the above-mentioned preponderance of franchisors offering a uniform "contract" to all franchisees: the contract specifies formulas for computing the franchise fee (e.g., depending on territory size). She further notes that even the low reported variability in royalty rates overstates the different treatment of different franchisees, as it partly reflects differences in the services being offered by the franchisor. Moreover, Bhattacharyya and Lafontaine (1992) report that requirements to buy inputs at markup from the franchisor or approved suppliers, requirements that can substitute for royalties, are completely uniform across all franchisees. This further supports the finding that a franchisor does not discriminate across franchisees in its marginal charges (Bhattacharyya and Lafontaine [1992] offer an alternative explanation for this relative constancy of marginal charges as compared to fixed fees.)

35 The fixed fee could matter insofar as a lower fee could allow the franchisor to bring in additional franchisees. This danger, however, may be more easily addressed by the exclusive-territory provisions.

36 Multiple responses were allowed in the survey. Of the 88 respondents, 35 cited loss of flexibility to cope with special economic or geographic conditions, 28 cited loss of flexibility to prevent losing potential franchisees (presumably due to differences in their outside opportunities), and seven cited other reasons.
attributable to legal constraints.\textsuperscript{37} In Lafontaine's (1992a) survey, when asked to identify the advantages of using the same contract with all franchisees, only 29 of the 120 respondents cited compliance with legal, FTC, or disclosure requirements, and 48 cited transaction costs (e.g., negotiation and ease of administration). In contrast, 62 cited uniformity and consistency in methods of dealing, and 33 cited a desire for fairness and equity (multiple responses were allowed). The latter replies can be construed as attempts to reassure against opportunism.

The rigidity of franchise terms over time is also striking. Banerji and Simon (1992) track 31 fast-food franchisors over a minimum of three years (and an average of over six years), finding that 71 never changed their royalty rate and 13 changed it only once. Interestingly, the franchise fee—which does not affect marginal incentives—changed more frequently (the fee, in real 1980 dollars, was changed at least once by 20 respondents). Similarly, Lafontaine (1992a) finds that 59 out of 124 respondents had never changed the royalty rate, while only 36 out of 125 had never changed the franchise fee (though this understates the rigidity, since the fees are in nominal dollars). For a different sample of 125 franchisors, each observed twice at least five years apart, Lafontaine (1993) reports that both the average royalty rate and franchise fee changed insignificantly (see also Lafontaine and Shaw, 1992). This intertemporal rigidity in contract terms seems even less plausibly explained by legal constraints.

To conclude, our discussion may provide clues to how the number of trading partners is chosen. There can be a trade-off between static efficiency and the need to reduce opportunistic renegotiation incentives: dealing with several firms might be statically more efficient but could increase scope for opportunism. The choice might depend on how important it is to preserve contracting flexibility. If the environment is relatively stable, long-term contracts become more feasible, so opportunism might be curbed while admitting multiple downstream firms. If the environment is changing rapidly and unpredictably, requiring ongoing adjustments in contract terms, or if secret discounts are easy, exclusivity becomes more attractive.

In order to make serious headway on these contracting problems it will be necessary to introduce asymmetric information explicitly and to specify more fully the particular environment. However, the issue of opportunistic recontracting in multilateral vertical relations offers a fruitful perspective for advancing our understanding of a wide range of business practices.

\textbf{Appendix}

\textbf{Proof of Proposition 1:}

Consider the last firm that the monopolist intends to be active, firm $k \leq n$. Suppose that, in the commitment game, $G^*$ is attainable by offering firms 1 through $k$ the vector $\mathbf{r}^*$, and that in the sequential game the monopolist has offered to each firm $i < k$ its price from $\mathbf{r}^*$ (otherwise necessarily $V < G^*$). When contracting with firm $k$ the monopolist, having collected fixed fees from other firms, does not maximize overall profit $G$, but rather maximizes $u_k = G - \sum_{i \neq k} \pi_i$. Observe that

$$\frac{\partial u_k(\mathbf{r}^*)}{\partial r_k} = \frac{\partial G(\mathbf{r}^*)}{\partial r_k} - \sum_{i \neq k} \frac{\partial \pi_i(\mathbf{r}^*)}{\partial r_k} < 0$$

since $\frac{\partial G(\mathbf{r}^*)}{\partial r_k} = 0$ by definition of $\mathbf{r}^*$ and since $\frac{\partial \pi_i(\mathbf{r}^*)}{\partial r_k} > 0$ for at least some firm $i$ (as at least one firm must be active in

\textsuperscript{37}The Robinson-Patman Act does not apply to services, and courts have consistently held that grants of trademark and franchise licenses do not constitute sales of commodities (Zeidman, 1991); only eight states have price-discrimination laws that do cover services (ABA, 1991). State statutes governing franchising specifically apply only to automobile and gasoline dealers (Zeidman, 1991). Some states have "little FTC Acts" and administrative policies governing franchising, and conceivably these may account for uniformity in those states. However, state treatments differ widely and are probably insufficient to explain the widespread uniformity observed. The administrative convenience of having a single disclosure statement may also play a role; but, as noted in the text, replies to Lafontaine's survey suggest that this is not paramount.
addition to \( k \) to attain \( G^* \), by hypothesis). So if the monopolist has offered to all firms \( i < k \) their prices from \( r^* \), to firm \( k \) it will offer \( r_k < r^*_k \), implying \( G < G^* \). In equilibrium the monopolist's behavior is foreseen; hence the monopolist earns \( V = G < G^* \).

**PROOF OF PROPOSITION 2:**

Consider a symmetric candidate equilibrium in which \( m \) firms are active, \( 2 \leq m \leq n \), with the same contract \((r_0, f_0)\), \( f_0 = \pi(r_0) \). Suppose \( r_0 \) is not pairwise-proof:

\[
\frac{\partial \pi_i(r_0)}{\partial r_i} \neq 0.
\]

Thus, there is a deviation price \( r_i \neq r_0 \) that would increase the joint profit of the monopolist and any firm, if that firm alone accepted \( r_i \) in stage 1 (contracting stage) and all \( m - 1 \) others stayed at \((r_0, f_0)\) after the reconstructing stage. Let \( f_1 = \pi(r_1, r_0) \) denote a firm's profit if it gets \( r_1 \) and all \( m - 1 \) others do stay with \( r_0 \). Consider the last firm accepting \((r_0, f_0)\) in stage 1, firm \( m \leq n \). Suppose the monopolist instead offers \((r_1, f_1)\) to firm \( m \). Firm \( m \) will accept this deviation offer if it expects all others to stay with \( r_0 \).

To see whether others would stay, consider the reconstructing stage. Let firm \( k \), \( k < m \), be the first mover after \( m \) in the reconstructing stage; firm \( k \) is the next-to-last firm that initially accepted \((r_0, f_0)\). Firm \( k \) knows that firm \( m \) has accepted \( r_i \). Suppose it expects all later movers in the reconstructing stage to remain with \( r_0 \); we relax this assumption below. Let \( v_1 \) denote firm \( k \)'s expected profit if it switches to \((r_1, f_1)\) and let \( v_0 \) denote the profit if it stays with \((r_0, f_0)\). Suppressing input prices to firms \( j \neq k, m \) we have

\[
v_1 = \left[ \pi(r_1, r_1) - f_1 \right] = \left[ \pi(r_1, r_1) - \pi(r_1, r_0) \right]
\]

\[
v_0 = \left[ \pi(r_0, r_1) - f_0 \right] = \left[ \pi(r_0, r_1) - \pi(r_0, r_0) \right]
\]

\[
\Rightarrow v_1 - v_0 = \left[ \pi(r_1, r_1) - \pi(r_0, r_1) \right] - \left[ \pi(r_1, r_0) - \pi(r_0, r_0) \right]
\]

\[
= \int_{r_0}^{r_1} \frac{\partial^2 \pi_k}{\partial r_k} (r_k, r_m) dr_m dr_k < 0.
\]

If firm \( k \) expected others to switch to \((r_1, f_1)\), then by the same argument, its loss from switching would be even greater. Thus, neither firm \( k \) nor any other would follow \( m \)'s deviation; foreseeing this, firm \( m \) would accept such a contract in stage 1, breaking the hypothesized equilibrium. Therefore, a symmetric equilibrium must be pairwise-proof with respect to firm \( m \): \( \frac{\partial u_m(r_0)}{\partial r_m} = 0 \).

However, \( \frac{\partial u_m(r_0)}{\partial r_m} = 0 \) implies

\[
\frac{\partial G(r_0)}{\partial r_m} = \frac{\partial u_m(r_0)}{\partial r_m} + \sum_{i \neq m} \frac{\partial \pi_i(r_0)}{\partial r_m} > 0
\]

since \( \frac{\partial \pi_i(r)}{\partial r_m} > 0 \) for any \( r \) and any active firm \( i \neq m \). Therefore, in symmetric equilibrium, total profit is less than \( G^* \). By assumption, \( G^* \) cannot be attained in an asymmetric equilibrium. Thus, the monopolist earns \( V < G^* \).

**PROOF OF PROPOSITION 3:**

When considering a deviation from equilibrium with an active firm \( i \), the monopolist maximizes

\[
u_i = (r_i - z) q_i + \pi_i + \sum_{j \neq i} (r_j - z) q_j.
\]

By assumption, firms \( j \neq i \) cannot observe changes in \( r_i \); hence the last term is invariant to \( r_i \). By passive beliefs, firm \( i \) assumes that prices \( r_j \) remain constant; hence we need consider only the effect of \( r_i \) on the first two terms on the right. The first-order effect on \( \pi_i \) from changing firm \( i \)'s own downstream variable (price or quantity) in response to \( r_i \) is zero by the envelope theorem; the direct effect of raising \( r_i \) (namely, \( q_i \)) is a pure transfer and hence cancels. Therefore, for each active firm \( i \), equilibrium prices \( r^* \) must satisfy

\[
\frac{\partial u_i(r^*)}{\partial r_i} = (r^* - z) \frac{\partial q_i(r^*)}{\partial r_i} = 0.
\]

Here \( \frac{\partial q_i(r^*)}{\partial r_i} \) denotes the change in firm \( i \)'s optimal input purchases assuming (by passive beliefs) that other firms' choices
are fixed (at the candidate equilibrium levels). As $\partial q_i / \partial r_i < 0$, $r^e_i = z$ for each active firm $i$. With symmetric firms, each would enter if offered price $z$ (assuming its active rivals also will be charged $z$). Since bringing in another firm lets the monopolist capture its profit and leaves unchanged the profit collected from other firms, the monopolist brings in all $n$ firms and earns $n\pi(z)$.

**Wary Beliefs**

Given the postulated beliefs, a firm accepts a contract only from the set $A = \{(r, f) : f \leq \pi(r, R(r))\}$. Consistency of the firm's beliefs with the rival's strategy requires

$$\begin{align} (A1) \quad (R(r), F(r)) &\in A. \end{align}$$

Otherwise the rival would reject its offer, contradicting the postulated beliefs. Consistency of beliefs with the monopolist’s strategy requires

$$\begin{align} (A2) \quad (R(r), F(r)) &= \arg\max_{(s, g) \in A} \left[V(r, f, s, g)\right] \\
&\text{where } V \text{ denotes the monopolist's profit function. (The function } V \text{ is different in the two games, but we use the same symbol to simplify notation.) Thus,} \\
&F(r) = \pi(R(r), R(R(r))) \\
&\text{implying} \\
&\begin{align} (A3) \quad R(r) &= \arg\max_{s} (r - z)q(r, s') + (s - z)q(s, r') + f + \pi(s, R(s)) \\
&\text{where} \\
&(A4) \quad s' = R(r) \quad r' = R(s) \end{align} \\
&\text{in the unobservability game, and where} \\
&(A5) \quad s' = s \quad r' = r \end{align}$$

in the ex post observability game. Equation (A3) represents the imposition of wary beliefs: the monopolist offers a best response $R(r)$ to the other firm, given the observed offer $r$. The monopolist's equilibrium strategy is to choose from the set $A$ its two profit-maximizing contracts, one for each firm. In a symmetric equilibrium, firms get the same contract $(r^e_i, f^e_i)$, where $f^e_i = \pi(r^e_i, r^e_i)$ and $R(r^e_i) = r^e_i$.

**Remark:** Symmetry beliefs are not wary beliefs under unobservability ($s' = r, r' = s$, below) or ex post observability ($s' = s, r' = r$). To see this, note that under symmetry beliefs, the monopolist solves

$$\max_{r, s} \left[V(r, s)\right] = \left[(r - z)q(r, s') + (s - z)q(s, r') + \pi(r, r) + \pi(s, s)\right]$$

and $s = r$ generally is not best given $r$ (although it will be when $r = r^*$).

Substituting (A4) into (A3) shows that in the unobservability game

$$\begin{align} (A6) \quad R(r) &= \arg\max_{s} [(r - z)q(r, R(r)) + f] \\
&\quad + (s - z)q(s, R(s)) + \pi(s, R(s)). \end{align}$$

Since $s$ is absent from the term in square brackets (the profit from the first firm), $R(r)$ is independent of $r$. Similarly, $R(s)$ is independent of $s$. Thus, $R(r)$ is the value of $s$ that maximizes an expression of the form $A + [(s - z)q(s, x) + \pi(s, x)]$ where $A$ and $x$ are independent of $s$. Therefore, $R(r) = z$, as claimed in the text.

This result generalizes to arbitrary nonlinear contracts. Let a contract $c$, which specifies the charge $c(q)$ as a function of the input quantity ordered $q$, induce belief $Y(\cdot;c)$. That is, if a firm is offered $c$, it expects the other firm to be offered any quantity $q'$ at cost $Y(q'; c)$. For any two contracts $c$ and $s$, let $Q(c, s)$ be the equilibrium quantity ordered by the firm with the contract $c$ if it expects the rival to accept $s$. If the seller offers accepted contracts $c$ and $s$ to the firms, it therefore earns

$$\begin{align} (A7) \quad V &= c(Q(c, Y(\cdot;c))) - zQ(c, Y(\cdot;c)) \\
&\quad + s(Q(s, Y(\cdot;s))) - zQ(s, Y(\cdot;s)). \end{align}$$
Wary beliefs impose the restriction that $Y(\cdot; c)$ maximizes $V$ over $s$, and thus $Y$ is independent of $c$. Let $Y(\cdot; c) = T(\cdot)$. Then,

$$V = c(Q(c, T)) - zQ(c, T)$$

$$+ s(Q(s, T)) - zQ(s, T).$$

For any $c$, holding constant the belief $T$, there is a two-part tariff that implements the quantity $Q(c, T)$ with the same total charge. Thus, the maximization problem over $c$ is equivalent to maximizing over a two-part tariff which, from our previous result, produces marginal-cost pricing ($r = z$) by the monopolist.

For \textit{ex post} observability, our 1992 paper analyzes Cournot duopoly downstream with inverse demand $p = 1 - (q_1 + q_2)$, constant costs $r_1, r_2$, and $z = 0$. An equilibrium in two-part tariffs exists. Beliefs are: $R(r) = a + br$, where $a \approx 0.69$, $b \approx 0.035 \Rightarrow r^e = b/(1 - a) \approx 0.116$. Under commitment, $r^* = 0.25$. Thus, $z < r^e < r^*$. 

REFERENCES


\textbf{Automotive News.} “GM, Dealers, Call Truce on Delco Outlets.” 28 March 1990, pp. 1, 49.


Reingen, Jennifer F. "Technology Adoption under Imperfect Information." *Bell


