Bilateral Relationships Governed by Incomplete Contracts

Author(s): Mehmet Bac

Source: *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift fÃ¼r die gesamte Staatswissenschaft*, Vol. 153, No. 2 (June 1997), pp. 320-333
Published by: Mohr Siebeck GmbH & Co. KG
Stable URL: http://www.jstor.org/stable/40751977
Accessed: 30-10-2017 10:51 UTC

Bilateral Relationships Governed by Incomplete Contracts

by

MEHMET BAC

This paper explores implications of interactions between noncontractibility of quality, multidimensional hidden information, switching costs and the frequency of trade on the terms of contracts in a buyer-seller setup. Optimal contractual arrangements are shown to consist of a sequence of two contracts with nondecreasing prices and nonincreasing quality and volume of exchange. In the absence of switching costs, an increase in the frequency of trade is absorbed by the first contract. For high frequencies of trade, switching costs may enhance welfare by improving the efficiency of screening through a better allocation of time between contracts. (JEL: L 14)

1. Introduction

A growing literature focusing on institutional arrangements when complete contracts are not feasible has emerged from the insights of ALCHIAN and DEMSETZ [1972] and WILLIAMSON [1975]. Examples are TIROLE [1986] on procurement contracts, HART and MOORE [1988] on renegotiations of incomplete contracts, FARRELL and SHAPIRO [1989] on switching costs, and MACLEOD and MALCOMSON [1988] and HOSIOS and PETERS [1993] in the labor context. The incomplete contracts approach of GROSSMAN and HART [1986] and HART and MOORE [1990] has also contributed substantially to our understanding of the nature of the firm.

This paper considers a discrete-time variant of the incomplete contracting model introduced in BAC [1993] to study the evolution of prices, qualities and the volume of trade, and to derive implications of introducing exogenous switching costs. It also investigates how the contractual arrangement responds to an increase in the frequency of trade.

The model comprises a buyer with noncontractible idiosyncratic needs and a pool of potential sellers with private information about their quality-relevant characteristics. The type of a seller is decomposed into two components. Sellers differ not only in their *innate values* (a quality shifting parameter), but also in their *cost of effort*. The innate value is a substitute for quality, as for example job-specific abilities are for performance in the labor context. But quality can

Journal of Institutional and Theoretical Economics (*JITE*), Vol. 153 (1997)

also be improved through effort. Thanks to this decomposition of the sellers' types, the buyer's learning process is nontrivial; it is complicated by the possibility that a "bad" seller imitates the performance of a "good" seller by exerting a sufficiently high effort.

In a bilateral monopoly framework, BAC [1993] showed the importance of learning about innate values for the buyer (the uninformed party) in choosing contract terms, and that the number of governing contracts is limited to two. Prices and, as we show here, the volume of trade, cannot be made contingent on quality but rather serve screening purposes. We derive further implications concerning the evolution of the parties' relative bargaining powers. Though the buyer faces many potential trading partners, hence retains considerable bargaining power at the outset, this power is partly captured by the selected seller as he transmits type-related information, possibly by exerting effort. A bilateral monopoly develops through the contractual arrangement where, after a probationary first contract, the seller who proves a high innate value in equilibrium is "tenured" with a second contract.

These featues of the model accord well with those observed in real-life contracting practices where the transactions involved are rather idiosyncratic. *First*, the emergence of a bilateral monopoly through the relationship is a salient feature of many actual settings such as labor, lending-borrowing and vertical buyer-seller relationships. The standard explanation for this phenomenon is the building up of reciprocal relationship-specific investments. In this model it is learning: Continuation dominates switching partners when a good matching prevails. *Second*, the evolution of the parties' bargaining powers is associated with an increasing profile of prices, which finds ample support in the labor context.[1] We therefore show that these features may be generated by the interaction between contract incompleteness and workers' private, performance-relevant information. *Third*, the analysis provides insights about the role of frictions in dynamic relationships. Frictions, or in their most simple form, exogenous switching costs, are natural aspects of many long-term relationships. It is widely believed that such switching costs enlarge the scope for opportunism (e.g., KLEMPERER [1987], FARRELL and SHAPIRO [1988]), but the analysis identifies an environment in which the opposite happens. Especially if the parties transact quite frequently, increasing the switching cost of one of the parties can be mutually beneficial even in the absence of direct mechanisms to check the hazard of opportunism. The key to this result is the potential conflict between

[1] Among the theories that explain why wages grow with seniority, the most closely linked to this paper are MALCOMSON's [1984] rank order tournament whereby the best performing young workers can win the prize of a wage greater than marginal product when old, and HOLMSTROM's [1982] analysis of provision of intertemporal managerial incentives. An important feature of these models and the present one is the reversal of the ratchet problem: dynamics has a positive influence on effort. Other explanations of wage patterns based on human capital theory using "delayed payment contracts" include CARMICHAEL [1983] and HARRIS and HOLMSTROM [1982].

the buyer's ex-ante screening motive and ex-post switching incentives. We show that endogenous switching costs in the form of cancellation fees may eliminate this conflict and even increase the efficiency of screening through a better allocation of time among the governing contracts.

The paper is organized as follows. The basic model is presented in section 2. Section 3 investigates the response of contractual arrangements to increasing the frequency of trade, and section 4 looks at the role of switching costs. Section 5 contains a summary and concluding remarks.

2. The Model

A buyer with idiosyncratic needs (called quality and denoted q) faces many potential sellers. Court ordering with respect to other aspects of trade, such as the unit price p, the quantity Q to be delivered to the buyer, cancellation fees, and durations of contracts, is assumed efficacious. There are two periods.[2] The per-period utility function of the buyer is $U(Q, q, p) = u(Q, q) - pQ$, where $u(Q, q)$ is concave and increasing in both Q and q. Quality is affected by effort x exerted per unit of output, and by the seller's innate value θ through the following simple technology (assumption 1):

$$q(\theta, x) = \theta + x. \quad \text{(A1)}$$

The buyer observes quality but not the effort. Besides θ, the seller's type has a second component β, a parameter measuring the convexity of his effort-disutility function $d(\beta; x)$. The type $\{\theta, \beta\}$ of a seller is his private knowledge, where $\theta \in \{\theta_H, \theta_L\}$ and $\beta \in [\underline{\beta}, \overline{\beta}]$. We assume that θ and β are independently distributed, with $\pi = prob(\theta = \theta_H)$ and β is distributed on $[\underline{\beta}, \overline{\beta}]$ with a continuous, strictly positive density function. We have $\theta_H > \theta_L$, so a θ_H-seller has an absolute advantage in providing quality.

The per-period utility of a seller with effort-disutility parameter β is $\upsilon = [p - c - d(\beta; x)]Q$ if he is on contract with the buyer. The constant marginal production cost net of quality effort is denoted c, therefore total marginal cost is $c + d(\beta; x)$ for an output of quality $\theta_i + x$. We make the following assumptions about $d(\beta; x)$:

(A2) *(i)* $d(\beta; x)$ is convex and increasing in x, and $d(\beta; 0) = 0$ for all β; *(ii)* there exists a finite level of effort, $\bar{x}(\beta)$ such that $d(\beta; x) \to \infty$ as $x \to \bar{x}(\beta)$ for all β; *(iii)* given $\beta_1 > \beta_2$, we have $\bar{x}(\beta_1) > \bar{x}(\beta_2)$. Furthermore, $d(\beta_1; x) < d(\beta_2; x)$ and $\partial d(\beta_1; x)/\partial x < \partial d(\beta_2; x)/\partial x$ for all x.

A seller with a higher β has a higher effort capacity and a lower absolute and marginal effort disutility at all effort levels. To exemplify this decomposition of

[2] The implications of increasing the number of periods are studied in section 3.

types, in the context of employment relationships β would be a measure of laboriousness and a high innate value would represent a talented, creative type.

Two implications of this setting should be mentioned at this stage. First, noncontractibility of quality severely limits the potential types of contractual arrangements. Second, noncontractibility of quality and hidden type-information together generate the potential for the use of effort to conceal differential innate values because the two are substitutes. For example, this phenomenon manifests itself in the labor context where a less talented but hardworking employee occasionally performs better than a highly talented employee, and in vertical relationships where "bad" sellers invest to provide a temporary quality improvement and satisfy the buyer's needs, thereby to obtain an increase in price/volume of future orders.

Assuming no discounting for simplicity, the buyer's expected utility is

$$U = E\{u(Q_1, q_1) - p_1 Q_1 + [u(Q_2, q_2) - p_2 Q_2]\}, \tag{1}$$

where the subscripts 1 and 2 denote the corresponding periods. Note that Q_1 and Q_2 are not necessarily supplied by the same seller because the buyer may switch at the interim date. The possibility of changing one's trading partner can obviously depend on the contents of the initial contract. The decision on the form and contents of the initial contract is left to the buyer. The justification for this bargaining power of the buyer is competition between potential sellers who are all initially identical from the buyer's viewpoint. The sellers have an outside option which yields them a per-period utility normalized to zero.

Given a contract proposal, sellers make acceptance choices, the buyer chooses one among those who accept, and the relationship begins. The expected utility of the selected seller as viewed from the outset is:

$$V = E\{[p_1 - c - d(\beta; x)] Q_1 + v_2\}. \tag{2}$$

In (2), v_2 represents the seller's second-period utility which, in general, may depend on his actions during the first period indirectly through the buyer's beliefs about the seller's type. As shown in Bac [1993], a long-term contractual arrangement that locks the buyer with a seller for the entire planning horizon can do no screening and is dominated by short-term contracts. Let $\mu(\theta_H \mid q_1)$ denote the buyer's updated beliefs about her trading partner's innate value as the one-period contract $\{p_1, Q_1\}$ expires. At this stage the buyer makes a binary continuation decision denoted $b \in \{0, 1\}$. If she opts for keeping the seller ($b = 1$), then the parties negotiate the terms of the second contract, a price p_2 and a quantity Q_2. The buyer's initial contract proposal $\{p_1, Q_1\}$ maximizes (1) subject to

$$(p_1 - c) Q_1 \leq 0, \tag{3}$$

and

$$E\{[p_1 - c - d(\beta; x^*)] Q_1 + b v_2(p_2, Q_2)\} \geq 0. \tag{4}$$

The initial contract must eliminate all "bad" seller types who, withholding effort, would derive a net benefit from the relationship. This is stated in (3), and can be termed the "no free lunch for unwanted seller types" condition. It implies

$$(3') \qquad p_1 \leq c.$$

On the other hand, the relationship must provide the "good" seller types (the θ_H-sellers) a nonnegative utility. This is stated in (4). Combining (3′) and (4) yields the dynamics of prices: $p_1 \leq c \leq p_2$.

The crucial parameter determining the parties' bargaining positions at the negotiation stage is the buyer's learning about her trading partner (in fact, $\mu(\theta_H; q)$ provides the only link between the two periods). For instance, the incumbent seller who proves an innate value of θ_H enjoys a great advantage over outsiders because the buyer will strictly prefer keeping the seller who she knows to be of type θ_H rather than contracting with another seller of unknown type. The expected surplus from continuation of the relationship is strictly positive whenever $\mu > \pi$.

Once the parties agree on p_2, we let the buyer choose Q_2. Using the parties' disagreement utilities, we can determine the support of the set including the negotiated price p_2^* as a function of μ. Given the commonly known negotiation procedure and rational expectations about the outcome of negotiations, the price $p_2^*(\mu) \in [c, \bar{p}_2(\mu)]$ that will arise from the negotiation process will be anticipated in advance. Instead of adopting a specific negotiation model, we impose an intuitive condition on the bargaining outcome. We assume that $p_2^*(\mu)$ is an increasing function of μ. The size of the expected surplus is increasing in μ, hence it is natural that the seller captures some of this incremental surplus through a higher price. The more the seller convinces the buyer that he is of type θ_H, the higher is his bargaining power, thus the higher the price he can obtain for the second contract.[3] Given $p_2^*(\mu)$ and the buyer's optimal quantity choice Q_2^*, we can write the seller's second period welfare as $v_2^*(\mu) = (p_2^*(\mu) - c)Q_2^*$.

Consider now the design of the first contract $\{p_1, Q_1\}$ and corresponding post-contract strategies. Let S represent the set of first contracts satisfying (3′) and (4), and let $\bar{S}$ denote the boundary of S. For each price $p_1 \leq c$, the largest

[3] Notice, however, that the seller who accepts the first contract incurs a welfare loss with respect to his outside option during the corresponding period. The seller incurs this loss because he foresees that he will improve his bargaining power sufficiently to capture some of the surplus in the second period. The seller's relative welfare loss during the first period can be interpreted as a relationship-specific investment producting type-related information. Since this "output" is useless outside the relationship, the seller will undertake the investment only if he expects a quasi-rent from continuation. Expected shifts in the parties' bargaining powers are therefore essential in this relationship. They are also among major concerns of buyers in real-life vertical relationships. As noted by MONTEVERDE and TEECE [1982], General Motors and Ford have been likely to integrate components requiring special engineering to avoid the increase in the suppliers' bargaining powers.

quantity that makes (4) binding (which obviously corresponds to $x = 0$) yields the $\bar{S}$ locus:

$$\bar{S} = \{p_1, Q_1 \mid (p_1 - c) Q_1 + (p_2^*(\mu) - c) Q_2^* = 0\}.$$

During the first contract, the selected seller chooses an effort to maximize (2), where $v_2 = (p_2^*(\mu) - c) Q_2^*$ and expectations are taken over the effort strategies of other seller types, the buyer's beliefs and continuation decision. The combined strategies of the parties must form a *perfect Bayesian equilibrium* (PBE). A *continuation equilibrium* can be defined as a strategy profile and an updating rule, excluding the buyer's first-contract offer from the strategies defining PBE.

As in BAC [1993], the buyer interprets a quality higher than the equilibrium quality as coming from a θ_H-seller because the latter has an absolute innate quality advantage over a θ_L-seller. We define below $x_{\theta_L}^M(\bar{\beta})$, the maximum effort that makes the $\{\theta_L, \bar{\beta}\}$-seller indifferent between accepting and rejecting the first contract, by

$$d(\bar{\beta}; x_{\theta_L}^M(\bar{\beta})) = (p_1 - c) + \frac{(p_2^*(\mu) - c) Q_2^*}{Q_1}. \tag{5}$$

Note that $x_{\theta_L}^M(\bar{\beta})$ is increasing in p_1 but decreasing in Q_1. The behavior of the $\{\theta_L, \bar{\beta}\}$-seller will be critical in determining the equilibrium effort strategies of other seller types. Proposition 1 characterizes continuation equilibria, and the properties of a PBE.[4]

Proposition 1: (A) Given $p_2^*(\mu)$, a unique continuation equilibrium exists. It has the following properties: *(i)* $\mu^* = 1$ and the buyer sets $b^* = 1$ as the first contract expires; *(ii)* θ_H-sellers accept and choose the effort $x_{\theta_H}^* = \max\{0, x_{\theta_L}^M(\bar{\beta}) - (\theta_H - \theta_L)\}$; *(iii)* all θ_L-sellers reject the first contract, except the $\{\theta_L, \bar{\beta}\}$-seller who is indifferent between rejecting and accepting with exerting the effort $x_{\theta_L}^M(\bar{\beta})$.

(B) Continuation equilibrium quality is at least θ_H. To obtain a higher quality, however, the buyer must decrease the quantity and/or increase the price specified in the first contract. As for the pattern of prices and qualities: $p_1^* \le p_2^*(1)$ and $q_1^* \ge q_2^* = \theta_H$. Furthermore, assuming $\partial^2 u(Q, q)/\partial Q \partial q \ge 0$, in an interior solution where the seller's effort is not too high we have $Q_1^* > Q_2^*$.

(C) Information about seller's innate values has nonpositive value.

Two potential classes of PBE arise according to whether θ_H-sellers exert effort, that is, whether or not the $\{\theta_L, \bar{\beta}\}$-seller is able to provide a quality higher than

[4] The proof follows those of propositions 2 and 3 in BAC [1993]. It is omitted here.

$q = \theta_H$.[5] If the equilibrium exhibits no effort, we have $\{p_1, Q_1\} \in \bar{S}$, and along $\bar{S}$ all θ_H-sellers' participation constraints are identical and binding. The second potential class of continuation equilibria exhibits positive effort. In that case we have $x^M_{\theta_L}(\bar{\beta}) > \theta_H - \theta_L$. A θ_H-seller expends effort to improve quality thanks to the impact of quality on the buyer's posterior beliefs, thereby on the expected price that will be negotiated for the second contract.

The buyer would obviously like to obtain the highest quantity and quality at the lowest price, but these goals are in conflict as mentioned in part (B) of proposition 1. Given the surplus that the θ_H-seller expects from the second contract, varying p_1 and/or Q_1 affects directly his first-period utility, so the seller adjusts his equilibrium effort (hence, quality) accordingly to eliminate the possibility of being imitated by a θ_L-seller. The behavior of the least effort-averse θ_L-seller, i.e., the $\{\theta_L, \bar{\beta}\}$-seller, plays the key role in this process. A sequence of two contracts is a net improvement over a single long-term contract that covers both periods.[6] Besides directly affecting the buyer's welfare, the low price and high volume of exchange specified in the first contract serve screening purposes, and effort, if any, comes about as a by-product of this process. Not surprisingly, efforts exerted in equilibrium are inefficient except for the $\{\theta_L, \bar{\beta}\}$-seller who is indifferent between accepting and rejecting the contract. Since all θ_H-sellers exert the same effort (provide the same quality) but differ in their costs of effort, their efforts must be less than efficient.

The intuition behind part (C) of proposition 1 lies in the intertemporal welfare transfer induced by the contractual arrangement. The low price and high quality during the first period more than compensates the buyer for the high price she pays during the second period. Similarly, the price for the second contract is sufficiently high to offset the seller's cost of effort and the low price specified in the first contract. Noncontractible quality is improved indirectly by the rat race for the second contract, during which type-related information is conveyed.

---

[5] One can construct other continuation equilibria by modifying the buyer's belief system, but most of these beliefs fail an intuitive equilibrium refinement such as CHO and KREPS' [1987] criterion. For instance, there is a continuation equilibrium in which no agent accepts a given contract satisfying (3′) and (4), supported by the buyer's belief (and sellers' anticipation of this belief) that she faces a θ_H-seller if and only if first-period quality is above some sufficiently high threshold. But these beliefs are not sensible becaue there are quality signals that cannot be conveyed by a θ_L-seller, which should therefore convince the buyer that she faces a θ_H-seller.

[6] An example is academic contracts, where assistant professors provide both a considerable flow of services and invest in the quality of their research and teaching in order to signal their innate values, thereby avoid being denied tenure. The department (buyer) enjoys both type-related information and good performance. Another example is General Motors' selecting suppliers on the basis of short-term, fixed-price contracting. A supplier's winning this year's contract is no guarantee of getting any business next year. See MONTEVERDE and TEECE [1982] for details.

Would the buyer benefit from splitting up his purchase and buying from several sources?[7] The potential advantage of multiple sourcing is to sample quality from several sellers, as the buyer would benefit from being able to pick the seller who is the most likely to be of high innate value. Another advantage of multiple sourcing may be that it increases the buyer's bargaining power later, at the contract renegotiation stage. For instance, if the incumbent sellers compete à la Bertrand, the buyer would appropriate the whole continuation surplus, leaving the incumbent sellers with a nonpositive expected surplus from the relationship. However, rational sellers who foresee this, will reject the buyer's initial multiple sourcing scheme. The possibility of ex-post competition reduces (in the limit, eliminates) the effectiveness of the screening process.

3. The Frequency of Trade

In the two-period version of our model, the opportunity cost of switching to another seller at the interim date is an increasing function of the buyer's posterior beliefs. But because learning about the innate value of the incumbent seller with noncontractible quality requires at least two periods, the result that the buyer does not switch at the interim date whenever $\mu > \pi$ may be an artefact of the two-period assumption. We now investigate the impact of increasing the frequency of trade, i.e. the number of periods, on the optimal number and length of governing contracts. FARRELL and SHAPIRO's [1989] principle of negative protection applies here: whatever their number, contracts should not include commitments for contractible variables beyond their termination dates.

Proposition 2: For any finite number n of periods, the optimal number of contracts is two. The first contract $C_1^* = \{p_1^*, Q_1^*\}$ covers the first n-1 periods, satisfies the "no free lunch" constraint $(n-1)(p_1^* - c)Q_1^* \leq 0$ and the participation constraint $(n-1)(p_1^* - c - d(\beta; x^*))Q_1^* + (p_n^*(1) - c)Q_n^* \geq 0$ where $p_n^*(1)$ and Q_n^* are respectively the price to be negotiated and the quantity determined by the buyer for the second contract which covers trade in period n.

Proof: To see that the first contract must cover the first n-1 periods, consider any other contract $C = \{(p_1, Q_1), \ldots, (p_m, Q_m)\}$ satisfying the "no free lunch" constraint for θ_L-sellers and the participation constraint of θ_H-sellers. This contract governs the first m periods where $1 \leq m < n-1$. Assume that it is accepted by only θ_H-sellers, so that $\mu = 1$ at the renegotiation stage. The negotiated price $p_j^*(1)$ for periods $j = m+1, \ldots, n$ would yield the buyer a share of the surplus. The incumbent θ_H-seller would also obtain a share that would at least offset his welfare loss during the first m periods. However, this is not an

[7] We are grateful to a referee for raising this question.

equilibrium outcome because the buyer will enhance her welfare by switching to another seller with the following contract offer: The price and quantity for $n - m - 1$ periods are such that only θ_H-sellers accept to obtain a zero net surplus, and this contract covers all remaining periods except the last. Notice that quality is again $q = \theta_H$. Since the buyer obtains the same quality in both options but her switching option makes the new seller's participation constraint binding, we conclude that the buyer is better off with the switching option. Given the buyer's incentive to switch, no seller will accept the initial contract $C = \{(p_1, Q_1), \ldots, (p_m, Q_m)\}$ for all $1 \leq m < n - 1$. This upsets the proposed equilibrium, leaving us with the equilibrium contract described in the proposition. Given $p_n^*(1) > c$, one can easily verify that there exist prices $\{p_1^*, \ldots, p_{n-1}^*\}$ and quantities $\{Q_1^*, \ldots, Q_{n-1}^*\}$ satisfying the two constraints stated in the proposition. The parties' continuation equilibrium strategies under this contractual arrangement are as described in proposition 1 for the two-period case. *Q.E.D.*

Proposition 2 shows that the optimal number of contracts is two for any finite number of periods, but in the absence of exogenous switching costs the second contract covers only the last period. Increasing the number of periods is entirely absorbed by the first contract. The intuition is quite simple: to induce screening, the buyer must choose an initial contract that results in some relationship-specific surplus at the end of the initial contract. If there is more than one period left at the end of the initial contract the buyer can always go back to the market, offer another screening contract which (as in the two-period case) attracts only high types, induces zero effort and gives all of the resulting surplus to the buyer. The buyer cannot do any better by continuing with the seller whom she believes to be a high type. Hence, screening is possible only if the initial contract covers all but the last period. This result is valid in the absence of switching costs.

4. *The Role of Switching Costs*

We study the impact of increasing the buyer's switching cost beyond the foregone continuation benefit in a three-period version of this relationship. The analysis extends to more than three periods in a natural way. Now there are four possible contractual arrangements to govern three periods. First, a single long-term contract *C1* can be used. The arrangements $C2 = \{C2_1\}$, $\{C2_2, C2_3\}$ and $C3 = \{C3_1, C3_2\}$, $\{C3_3\}$ consist of a sequence of two contracts with differential durations; under $C2$, the first contract $\{C2_1\}$, governs only the first period, whereas $\{C3_1, C3_2\}$ covers the first two periods under $C3$. The fourth possible arrangement is $C4 = \{C4_1\}$, $\{C4_2\}$, $\{C4_3\}$, where each period is governed by one contract; $\{C4_2\}$ and $\{C4_3\}$ will be determined through negotiations, except that the buyer specifies $\{C4_1\}$. In the *absence of switching costs*, proposition 2 stipulates that $C2$ and $C4$ are not feasible because the buyer will switch. Leaving aside the dominated arrangement *C1*, the only feasible arrangement in the absence of switching costs is therefore *C3*. However, below we show that an appropriate switching cost can make *C2* feasible and even generate a larger total surplus.

We simplify the model to show most clearly the impact of introducing switching costs. Assume that the volume of trade is restricted to one in each period, and the parties' per-period utilities are given by $u = q - p$ for the buyer, and $v = p - c - d(\beta; x)$ for the seller.[8]

Consider first the arrangement *C3*. The probationary contract $\{C3_1, C3_2\} = \{p_1^{C3}, p_2^{C3}\}$, where $p_1^{C3} < c$ and $p_2^{C3} < c$, selects the θ_H-seller. The total surplus from continuation through $\{C3_3\} = \{p_3^{C3}(1)\}$ is $\theta_H - c > 0$, and the price $p_3^{C3}(1)$ will be determined through negotiations. Letting Z_3^B and Z_3^S denote respectively the buyer's and the seller's share of the continuation surplus, it follows that $Z_3^B + Z_3^S = \theta_H - c$. To facilitate the comparison between *C3* and *C2*, let us assume that in the continuation equilibrium generated by the contract $\{C3_1, C3_2\}$ the θ_H-seller exerts no effort. In this case the seller's overall participation constraint is binding, which means $Z_1^S + Z_2^S + Z_3^S = 0$. Recall that $Z_1^S + Z_2^S$ must be negative from the "no free lunch" constraint. The neagtive surplus the seller incurs goes to the buyer, but this loss is exactly offset by Z_3^S, the surplus the seller captures later in period 3. Thus the buyer's total welfare under arrangement *C3* is $U_{C3} = 3(\theta_H - c)$, which is also the total surplus under *C3*.

The arrangement *C2* is not feasible in the absence of switching costs because, as $\{C2_1\}$ expires, the buyer can switch and obtain $U^d \geq 2(\theta_H - c)$, which is strictly higher than any continuation welfare she may obtain. We show below that if the buyer's switching incentives could somehow be eliminated, *C2* would become feasible and perhaps dominate *C3*. A switching cost k such that $U^d - k < 2(\theta_H - c)$ achieves this goal by making continuation in the mutual interest of the parties. To confirm our claim, let us choose a switching cost k^* such that, given the bargaining scheme employed, the incumbent θ_H-seller's share from the total surplus corresponding to the second contract $\{C2_2, C2_3\}$ is $2Z_3^S$ (twice the surplus resulting from the contract $\{C3_3\}$). Since this surplus is strictly higher than Z_3^S, it will generate ex-ante competition, and equilibrium effort during $\{C2_1\}$ may become positive so that quality $q_1^{C2} > \theta_H$. The overall participation constraint of the seller under arrangement *C2* with a switching cost of k^* is thus $p_1^{C2} - c - d(\bar{\beta}; x^*) + 2Z_3^S \geq 0$, and the buyer's total utility is $U_{C2} = q_1^{C2} - p_1^{C2} + 2(\theta_H - c) - 2Z_3^S$ where we used the fact that $Z_3^B = (\theta_H - c) - Z_3^S$. We are now ready to compare *C2* and *C3*. Let $\phi(\bar{\beta}; d)$ denote the inverse of $d(\bar{\beta}; x)$.

[8] With this simplification the parameter β of the effort disutility function has no role to play. Without loss of generality, we shall accordingly assume the quality-effort disutility function of all sellers is $d(\bar{\beta}; x)$.

Proposition 3: Assume $\phi(\bar{\beta}; p_1^{C3} + p_2^{C3} - 2c + (Z_3^S/2)) < \theta_H - \theta_L < \phi(\bar{\beta}; p_1^{C2} - c + 2Z_3^S)$. The total surplus under $C2$ (TS_{C2}) with a switching cost of k^* is strictly higher than under $C3$ (TS_{C3}).

The assumption stated in proposition 3 is derived from the equilibrium condition (5) defining $x_{\theta_L}^M(\bar{\beta})$, the maximum effort that a θ_L-seller can exert to imitate the equilibrium quality provided by a θ_H-seller. The assumption states that equilibrium effort is positive under $C2$ with the switching cost k^* but zero under $C3$. This, of course, depends on the curvature of $d(\bar{\beta}; x)$ and hence also on its inverse, $\phi(\bar{\beta}; d)$.

Proof: We know that $TS_{C3} = 3(\theta_H - c)$. Under $C2$ the total surplus is $TS_{C2} = q_1^{C2} - c - d(\bar{\beta}; x_{\theta_H}^*) + 2(\theta_H - c)$. By comparing the two surpluses, we obtain $TS_{C3} < TS_{C2}$ if an only if $q_1^{C2} - d(\bar{\beta}; x_{\theta_H}^*) > 0_H$, which simplifies to $x_{\theta_H}^* > d(\bar{\beta}; x_{\theta_H}^*)$. This condition holds in equilibrium. To see this, note that if p_1 is optimal the first-order condition $\phi' = 1$ must hold. Since $\phi(\bar{\beta}; .)$ is the inverse of $d(\bar{\beta}; .)$, $\phi' = 1$ can be expressed as $\partial d(\bar{\beta}; x_{\theta_L}^M(\bar{\beta}))/\partial x = 1$. But since $x_{\theta_H}^* = x_{\theta_L}^M(\bar{\beta}) + (\theta_H - \theta_L)$, we must have $\partial d(\bar{\beta}; x_{\theta_H}^*)/\partial x < 1$. Strict convexity of $d(\bar{\beta}; x)$ in x implies that $x_{\theta_H}^* > d(\bar{\beta}; x_{\theta_H}^*)$. *Q.E.D.*

Note that if the seller's surplus $2Z_3^S$ under $C2$ is not sufficiently high to induce effort, then the buyer would obtain the same surplus from the two arrangements. The *form* of the intertemporal welfare transfer induced through the arrangement $C2$ plays a key role in inducing effort when the buyer is almost locked-in. Ex-ante the buyer has the power to manipulate p_1^{C2} and capture most of the seller's overall surplus. Part of the surplus transferred to the buyer may be in the form of a better quality. This may generate a Pareto improvement. Such intertemporal welfare transfers allow the parties to capture part of the surplus otherwise left unexplored due to contract incompleteness.[9] This brings us to a perhaps obvious but important point. Higher switching costs can enhance welfare only if another contractible variable (duration of contracts) is appropriately modified; a higher switching cost per se cannot increase the total surplus. This is why we compared $C2$ and $C3$, two arrangements that differ in contract lengths. Finally, it should be noted that the level of the switching cost underlying this result is a critical magnitude. A low value of k may not be sufficient to eliminate the buyer's switching incentives even when the incumbent seller proves to be of type θ_H, so $C2$ may not even be feasible. Very large switching costs on the other hand may lock-in the buyer completely; they may shift all the bargaining power away to the seller, reducing the buyer's surplus to zero.

The discussion above brings in the notion of *optimal friction* in relationships when contracts are necessarily incomplete and information is asymmetric.

[9] Similar arguments can be used to show that the arrangement *C4* becomes feasible *if* an appropriate switching cost is introduced.

Some friction exists naturally in every relationship in the form of search and selection costs, but the parties can also artificially create friction through cancellation fees very similar to SHAVELL's [1980], [1984] damage payments and WILLIAMSON's [1983] hostages. Because potential gains from intertemporal welfare transfers are larger, a cancellation fee will most likely enhance the buyer's welfare if the frequency of trade (the number of periods) is high.[10] The buyer's gain will be an increased efficiency of screening due to a better allocation of time among the two contracts, and this gain may well outbalance the cost of losing some bargaining power to the seller at the interim negotiation stage. Several aspects of relationships considered as generating inefficiencies may have potentially beneficial roles in a world of incomplete contracts.

5. *Summary and Conclusions*

The basic buyer-seller framework presented in this paper is essentially a shirking model, as it is familiar from the efficiency wage literature. The buyer is in need of a good or a service that can be produced by a seller. The number of potential sellers is large but quality is not contractible so that the only contract that screens the seller types includes a price-quantity pair. In the one-period version of the model, the seller will find it optimal to exert no effort to improve quality, hence the buyer will offer a price that just covers the seller's cost of producing the lowest quality. We extend the framework in BAC [1993] and show that in the equilibrium of a multi-period relationship the buyer can use an initial contract (price and quantity) to screen seller types. This multi-period extension also provides insights about the number of contracts in a discrete-time framework, the role of switching costs, and the impact of the frequency of trade on contract terms.

The analysis shows that the unit price is increasing from the first contract to the second, whereas quality and volume of trade are nonincreasing. As mentioned in section 1, these dynamics accord well with the empirical evidence on the dynamics of wages and performance in the labor context, thus suggesting that learning about innate values plays an important role in determining the terms of labor contracts. An interesting example where contract incompleteness is plain is academics. Assistant professors provide extra effort in order to signal their teaching and research abilities, and thereby avoid being denied tenure.

[10] Recall that in the absence of switching costs the optimum arrangement consists of two contracts where the second contract covers the last period. Since in equilibrium the total surplus in the last period is limited by $\theta_H - c$ and the first contract is too long with $p_1 < c$, the seller does not have much incentive to exert effort. Hence a shorter duration for the first contract is likely to generate higher effort, which improves quality early in the relationship at the cost of letting the seller capture a higher surplus later on. This intertemporal welfare transfer may improve the buyer's welfare.

When contracts are incomplete, switching costs and hidden information – generally thought of as sources of inefficiencies – have potentially beneficial effects. In the buyer-seller relationship considered in this paper, increasing the buyer's switching cost can enhance her own welfare by allowing for a better allocation of time among the two contracts. Since a sufficiently high switching cost checks the hazard of termination, the duration of the first contract can be made shorter if this improves the efficiency of the screening process. As a result, the seller's "expected return" on effort (which produces type-information specific to the relationship) becomes nonnegative. The equilibrium will exhibit higher effort because there is ex-ante competition among seller types and the second-contract price depends on the first-contract quality through the buyer's posterior beliefs. The size of the switching cost is critical. Too much of it will harm the buyer by reducing her bargaining power while a switching cost too small may not be sufficient to generate the mentioned desirable effect.

We close the paper with a remark. Viewed as a screening model, this paper can be classified in the literature on rat races. Comparing the excessive effort outcomes of rat races with the first-best outcomes derived in standard principal-agent models where effort is priced on a piece rate basis has led many authors to conclude that rat races are likely to generate inefficient outcomes (see for instance Akerlof [1976] and Lazear and Rosen [1981]). This conclusion is clearly true when one adopts complete piece-rate contracts as a benchmark. Since many real-life contracting practices involve incompleteness and have little in common with sophisticated contracts, an incomplete contracts framework seems to be more appropriate as a benchmark. That is to say, contrasting the hidden-information (rat race) case with the perfect information case in incomplete contracts frameworks can contribute better to our understanding of screening properties of real-life contractual arrangements. In a simplified version of our model we have shown that rat races may generate welfare-enhancing outcomes; they provide both information and quality.

References

Akerlof, G. [1976], "The Economics of Caste and of the Rat Race and Other Woeful Tales," *Quarterly Journal of Economics*, 90(4), 599–617.

Alchian, A. and Demsetz, H. [1972], "Production, Information Costs, and Economic Organization," *American Economic Review*, 62(4), 777–795.

Bac, M. [1993], "Opportunism and the Dynamics of Incomplete Contracts," *International Economic Review*, 34(3), 663–683.

Carmichael, L. H. [1983], "Firm Specific Human Capital and Promotion Ladders," *Bell Journal of Economics*, 14(2), 251–258.

Cho, I. K. and Kreps, D. M. [1987], "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102(2), 179–221.

Farrell, J. and Shapiro, M. [1988], "Dynamic Competition With Switching Costs," *Rand Journal of Economics*, 19(1), 123–137.

-- and -- [1989], "Optimal Contracts With Look-In," *American Economic Review*, 79(1), 51–68.

GROSSMAN, S. and HART, O. [1986], "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94(4), 691–719.

HARRIS, M. and HOLMSTROM, B. [1982], "A Theory of Wage Dynamics," *Review of Economic Studies*, 49(3), 315–333.

HART, O. and MOORE, J. [1988], "Incomplete Contracts and Renegotiations," *Econometrica*, 56(4), 755–785.

–– and –– [1990], "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98(6), 1119–1158.

HOLMSTROM, B. [1982], "Managerial Incentive Problems – A Dynamic Perspective," pp. 156–177 in: B. Walross (ed.), *Essays in Economics and Management in Honor of Lars Wahlbeck*, Swedish School of Economics: Helsinki.

HOSIOS, A. J. and PETERS, M. [1993], "Self-Selection and Monitoring in Dynamic Incentive Problems With Incomplete Contracts," *Review of Economic Studies*, 60(1), 149–174.

KLEMPERER, P. [1987], "The Competitiveness of Markets With Consumer Switching Costs," *Rand Journal of Economics*, 18(2), 138–150.

LAZEAR, E. and ROSEN, S. [1981], "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89(5), 841–864.

MACLEOD, B. and MALCOMSON, J. [1988], "Reputation and Hierarchy in Dynamic Models of Employment," *Journal of Political Economy*, 96(4), 832–854.

MALCOMSON, J. [1984], "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy*, 92(3), 486–507.

MONTEVERDE, K. and TEECE, D. [1982], "Supplier Switching Costs and Vertical Integration in the Automobile Industry," *Bell Journal of Economics*, 13(2), 206–213.

SHAVELL, S. [1980], "Damage Measures for Breach of Contract," *Bell Journal of Economics*, 11(2), 466–490.

–– [1984], "The Design of Contracts and Remedies for Breach," *Quarterly Journal of Economics*, 99(1), 121–148.

TIROLE, J. [1986], "Procurement and Renegotiation," *Journal of Political Economy*, 94(2), 235–259.

WILLIAMSON, O. [1975], *Markets and Hierarchies: Analysis and Antitrust Implications*, The Free Press: New York.

–– [1983], "Credible Commitments: Using Hostages to Support Exchange," *American Economic Review*, 73(3), 519–540.

Professor Mehmet Bac
Department of Economics
Bilkent University
06533 Bilkent, Ankara
Turkey