Renegotiation and Conflict Resolution in Relational Contracting

Rui R. Zhao
Department of Economics
University at Albany - SUNY
Albany, NY 12222, USA
Tel: 518-442-4760
Fax: 518-442-4736
E-mail: rzhao@albany.edu

April 7, 2011

Abstract

Renegotiation and conflict resolution are studied in relational contracting under subjective evaluation. Renegotiation has three effects. First, it makes the incentive pay scheme low powered: the maximum variation of compensation across performance levels is compressed and the contract is less extreme compared to the case without renegotiation. This effect is stronger when the players are less patient. Second, renegotiation typically renders termination impossible; the contract relies on a “low morale" mechanism to enforce mutual cooperation. Finally, renegotiation compels the players to resolve their conflicts by selecting a contract that maximizes the lowest possible surplus along the path of the contract.

Key words: repeated principal-agent, self-enforcing contracts, renegotiation proof, private evaluation
JEL Classification: C73, D82, L14, J30

1 Introduction

Performance measurement lies at the heart of an effective incentive system. Traditional incentive theory has mainly looked at situations where performance can be measured objectively and hence compensation can be explicitly dependent on objective measures. For many jobs, however, performance is hard to measure let alone to verify objectively and agents are often compensated with discretionary payments such as bonuses that are based on subjective assessments of performances (Baker et al., 1994; Prendergast, 1999). Several studies have shown that self-enforcing contracts can mitigate or even completely resolve the moral hazard problems that come with nonverifiable performance measures, as long as the measures are mutually observable to the principal and agent.[1]

Matters are harder however if there are asymmetric information in each party's own subjective evaluation. This line of research has recently been pursued in Levin (2003), MacLeod(2003), and Fuchs (2007), by introducing *private* performance evaluation, which captures the potential differences in agents' opinions and reflects the often unobservable nature of the principal's valuation of the agent's contribution in many practical applications. For self-enforcing contracts to work, repeated interactions are needed to ensure that the principal and agent can credibly punish each other if either of them deviates from their implicit agreement. As shown by these authors, when output is privately observed *mutual* punishments actually are needed to support self-enforcing contracts. These punishments, in the form of terminating the relationship or players' carrying out inefficient actions for several periods, are broadly consistent with the observation that conflicts are commonplace in organizations and general long-term relationships.

This "favorable" view of conflict, however, does not seem to be shared by many organizations as they devote considerable resources to minimizing if not completely eliminating conflicts. Institutions such as arbitration and mediation are also established to resolve disputes. Such a sentiment towards conflict is also generally shared in the management literature (e.g. Milkovich and Newman, 1996).

In this paper I argue that such efforts toward conflict resolution can be reconciled with the functional role of conflict in enforcing relational contracts, when the possibility of contract renegotiation is taken into consideration. The point of departure is to recognize that ex post inefficient continuation contracts are open to renegotiation when the mechanism for enforcing the contracts relies on none other than the agents' self interests. Specifically, I begin with a

[1] For example, Bull (1987) and MacLeod and Malcomson (1989) show that when the principal and agent have the same beliefs regarding a subjective evaluation there exists first-best efficient self-enforcing contracts. Pearce and Stacchetti (1998) (also see Baker et al., 1994) show that the existence of contractible measures of performance can enhance the effectives of implicit contracts.

model similar to that of Levin (2003) in which the principal privately monitors the agent's performance. Although the principal has private information about the agent's performance, the principal's wage payment as well as any performance report are public information and are shared by the agent. This particular formulation allows me to focus on a class of perfect public equilibria of the repeated game (Fudenberg et al., 1994) (see Section 2). I then introduce a new ingredient in the model: to allow the two parties to renegotiate their agreements in every period. I study relational contracts that are renegotiation proof in the sense of Pearce (1987).

The renegotiation-proof concept has been previously applied to symmetric repeated games (see Abreu et al., 1993). Although the game in the present model is not symmetric, the concept turns out to be particularly suitable thanks to transferable utilities. In the present model collective welfare is measured by the total surplus generated in the relationship, and punishment for deviation takes the form of surplus destruction. As such agents can focus on what level of total surplus is acceptable when (re)negotiating agreements.

The idea is that when renegotiating current agreements the principal and agent realize that what they deem as an unacceptably low surplus today will be treated as such in the future as well. This will give them pause when renegotiating away from some "low" surpluses because such actions will forfeit *future* punishments and hence may actually hurt their current welfare as such punishments may be necessary for achieving good outcomes today.

The resulting renegotiation-proof contracts display several features that are consistent with some important characteristics of relational contracts.

First, renegotiation-proof contracts turns the group preference into a type of Rawlsian preference in the sense that the principal and agent select the contracts that maximize the lowest continuation surplus across all histories. This result, following from the renegotiation-proof concept, provides a rigorous formulation of conflict resolution: conflicts are minimized to achieve the best worst-case surplus. I show that shifting the focus to the *maximin* surplus indeed raises the minimum continuation surplus compared to optimal contracts without renegotiation. The result is a compression of the overall variation of continuation surpluses across all histories. It helps explain the apparent contradiction between the usefulness of invoking conflicts at enforcing cooperation and the vigorous efforts directed at minimizing conflicts in many organizations.

Second, the need to compress continuation surpluses also translates into *low-powered* incentive contracts *in every period*: the maximum variation of pay-to-performance ratio is more compressed and the incentive scheme is less extreme compared to optimal contracts without renegotiation. For instance, previous studies have shown that under private evaluation optimal contracts would punish the agent only when the worst performance is observed but with a big stick (MacLeod, 2003). In contrast, the compensation scale is less extreme in renegotiation-proof contracts: punishment is reduced in size but spread beyond the worst performance level, or equivalently, the size of reward and the chances of getting it (controlling for effort) are both reduced. This result provides an explanation for the observation that many real world incentive contracts are not as powerful as what existing theory predicts they should be.

Third, renegotiation-proof contracts generally render termination impossible; instead, the contracts rely on the “low morale" mechanism, episodes in the relationship during which the agent exerts low effort and receives low pay, as punishments for perceived deviations. This result again is consistent with stylized facts. For instance, union-firm contracting is frequently subject to renegotiation and the parties rarely got stuck in the worst conflict forever: they may go through periodic conflicts such as strikes or lockouts but they generally get back to better outcomes after a series of bad ones, consistent with what the theory predicts. The result that renegotiation-proof contracts select recurrent conflicts over termination as the mechanism to destruct surpluses for the purpose of enforcing cooperative behaviors also throw light on how incentives are generally provided under subjective evaluation, given that termination is infrequent relative to the wide use of subjective evaluation in the real world.

In closing the introduction, note that although renegotiation is treated as a technological issue, i.e. agents do not have access to a (legal) technology that enables them to commit to ex ante agreements, it is also useful to link it to preferences over long-term outcomes. One could view the desire of reducing conflicts as part of a preference trait. Then the renegotiation-proof contracts merely depict what is optimal under the maximin Rawlsian social preference.

In the rest of the paper I first introduce the model, then discuss the motivation and definition of renegotiation, and finally study the properties of renegotiation-proof contracts.

2 The Model

Two risk-neutral parties, a principal and an agent, have the opportunity to enter into an employment relationship in periods $t = 0, 1, 2, ...$ At the beginning of each period t, the principal offers the agent a compensation package, the details of which will be specified below. If the agent accepts the offer, he chooses an effort level e_t from a compact set $\mathbb{E} \subset \Re$; the effort level e_t is known to the agent but is not observed by the principal. Effort e_t produces some stochastic output y_t for the principal, which I assume is drawn from some interval $Y = [\underline{y}, \bar{y}] \subset \Re^+$ according to a distribution function $F(\cdot|e)$ and a density function $f(\cdot|e)$, independent of time t. At the end of the period, realized output is secretly observed and enjoyed by the principal; she then compensates the agent with a payment $w_t \in \Re$, which may be based on her observations of past and current performances of the agent. The agent's net payoff in the period equals $w_t - g(e_t)$, where $g(e_t)$ is the effort cost.

The model departs from the standard principal-agent model by allowing for private performance monitoring, which captures the inherent differences between the two parties' perspectives when an objective performance measure is unavailable.

I make the following assumption about technology.

A 1. Density $f(y|e) > 0$ on Y for all $e \in \mathbb{E}$. For all $e, e' \in \mathbb{E}$, $e \neq e' \Rightarrow g(e) \neq g(e')$, and $g(e) > g(e') \Rightarrow E(y|e) > E(y|e')$, where $E(y|e) \equiv \int_{\underline{y}}^{\overline{y}} y f(y|e) dy$ is the expected output given e.

I now describe the principal's offer in more details. At the beginning of period t, the principal offers the agent a one-period compensation package. The offer specifies the agent's wage as a function of his performance, namely a function $w_t : Y \longrightarrow \Re$. Since performance y is only observed by the principal, not by the agent, this offer needs some explanation. Let $k_t = \min_y w_t(y)$ and $b_t(y) = w_t(y) - k_t \geq 0$. The total wage w_t in fact consists of two components: a fixed component, k_t, and a contingent component, $b_t(y)$. The idea is that if the agent accepts the offer then the principal is obligated to make the fixed payment k_t regardless of output, but she is under no obligation to make any specific contingent payment $b_t(y)$ to the agent – this part is completely at her discretion. In other words, the contingent component b_t is only a good-faith offer. The fixed payment k_t can be negative, in which case the agent is required to pay the fee to get the job.

If the agent rejects the principal's offer then both parties take their outside options for the period. The two parties can also take their outside options at the beginning of every period before the principal makes any offer. I assume that the agent gets a fixed payoff u_0 per period from his outside option and the principal gets a fixed payoff v_0 per period from her outside option. Let $\underline{e} = \arg\min_e g(e)$ be the agent's minimum effort level. Assume that taking outside options is weakly more efficient than exerting minimum effort:

$$s_0 \equiv u_0 + v_0 \geqslant s(\underline{e}) \equiv E(y|\underline{e}) - g(\underline{e}). \tag{1}$$

So the parties weakly prefers not to form a relationship if the agent is not expected to perform.

In the long-term relationship, the principal's current offer in period t depends on the information she has when she makes the offer. For the purpose of the present analysis, the information available at the beginning of period t, denoted by h^t, includes the principal's past reports $\hat{y}^t = (\hat{y}_0, ..., \hat{y}_{t-1})$, her past offers, past wage payments $w^t = (w_0, ..., w_{t-1})$, and whether the relationship has ended due to the agent's refusal of past offers. In other words, we restrict attention to public information (the concluding section briefly discusses private histories). It is convenient and customary to denote by h^0 the null history at the beginning of period 0.

A *relational contract* σ specifies, for each period t and for each history h^t, the principal's wage offer $w_t(h^t, y_t)$ and her report $\hat{y}_t(h^t, y_t)$ if the offer is accepted, both functions of the principal's privately observed output y_t (there is no need to let wage depend on report: in that case we can think of $w(y) = \tilde{w}(y, \hat{y}(y))$ as the composite function); in addition, the contract also specifies the agent's decision as to whether accept or reject the principal's current offer and the agent's effort $e_t(h^t)$ if he accepts it. In sum, the contract specifies a pair of strategies for the two players as functions of histories. In subsequent analysis random wage payments by the principal will also be permitted.

Both the principal and the agent maximize the expected value of sum of discounted future utilities, using a common discount factor $\delta \in (0,1)$. Specifically, at the beginning of period t the payoffs of the principal and agent are respectively given by

$$v_t = (1-\delta)E_t\left\{\sum_{s=t}^{\infty}\delta^{s-t}[y_s - w_s]\right\},$$

$$u_t = (1-\delta)E_t\left\{\sum_{s=t}^{\infty}\delta^{s-t}[w_s - g(e_s)]\right\},$$

where the expectations are taken with respect to information available at the beginning of period t; also, payoffs are normalized by the factor $1-\delta$ to make them comparable to one-period payoffs. Risk neutrality takes intertemporal risk sharing out of the question and allows me to focus on dynamics of the relationship that come solely from incentive reasons.

Let σ be a given relational contract. Given any history h^t in period t, let $v(\sigma|h^t)$ be the *continuation payoff* the principal receives if the two parties follow through on the terms of the *continuation contract* $\sigma|h^t$. Similarly, let $u(\sigma|h^t)$ be the agent's continuation payoff.

At each h^t, since the principal is free to make reports and pay discretionary wages, for her to adhere to the terms of the contract she must be indifferent between the alternative choices intended for different y. That is, the following incentive constraints need be satisfied: for all $y, y' \in Y$,

$$-(1-\delta)w(y) + \delta v(h^t, w(y), \hat{y}(y)) = -(1-\delta)w(y') + \delta v(h^t, w(y'), \hat{y}(y')), \tag{PIC}$$

where $v(h^t, w(y), \hat{y}(y))$ is the principal's continuation payoff in period $t+1$. On the other hand, the agent's incentive constraint dictates that effort $e(h^t)$ should maximize his expected payoff:

$$e(h^t) \in \arg\max_e \int_{\underline{y}}^{\overline{y}} \left[(1-\delta)\big(w(y) - g(e)\big) + \delta u(h^t, w(y), \hat{y}(y))\right] f(y|e)dy. \tag{AIC}$$

In addition to these on-equilibrium-path (i.e. the principal always makes and the agent always accepts the "correct" offers) incentive constraints, there are also off-equilibrium-path incentive constraints that ensure the principal will make the correct offer at each h^t and will not pay a wage or make a report that is not designed for any y, and the agent will choose to accept or reject the principal's offer optimally.

The relational contract is *self-enforcing* if after every history h^t the contract satisfies the incentive constraints (PIC) and (AIC) as well as the off-equilibrium-path incentive constraints. In other words, the contract is self-enforcing if the players' strategies constitute a *perfect public equilibrium* (Fudenberg et al., 1994). Note that the principal's report and wage payment do depend on her private information y but her *strategy*, i.e. wage and report as functions of y, $w(y_t), \hat{y}(y_t)$, only depend on public history h^t.

A special class of contracts are the ones that do not ask the principal to make reports; this includes the trivial self-enforcing contract in which the two parties always take their outside options in every period. I keep the reports in the initial formulation because even though they do not affect the principal's payoff they may affect the agent's payoff (e.g. see problem in (2)). However they turn out to be irrelevant for the analysis of renegotiation-proof contracts.

3 Optimal Contracting and Renegotiation

To motivate the problem of renegotiation, first consider optimal contracting without renegotiation; in particular, start with the one-period contracting problem. In this one-period problem, the principal reports realized output y and the agent is paid according to a wage schedule $w(y)$; based on this incentive scheme, the agent optimally chooses some effort e. For the principal to reveal output y truthfully, the principal's out-of-pocket payment cannot vary with her report; so the principal makes available a fixed amount of money $\bar{w}$ to defray the agent's wage $w(y)$. An optimal contract is a tuple $(\bar{w}, w(y), e)$ that maximizes total surplus, that is, to solve the following problem:

$$\max_{e,w(\cdot),\bar{w}} \int_{\underline{y}}^{\overline{y}} (y - \bar{w} + w(y)) f(y|e) dy - g(e) \tag{2}$$

subject to:

$$e \in \arg\max_{e' \in \mathbb{E}} \int_{\underline{y}}^{\overline{y}} w(y) f(y|e') dy - g(e'),$$

$$w(y) \leqslant \bar{w}, \forall y.$$

Note that $\bar{w}$ only affects how surplus is divided between the two parties and generally is indeterminate. More importantly, in order to motivate effort, wage $w(y)$ should depend on output y, so money need be burnt whenever $w(y) < \bar{w}$. Such joint punishments are not unusual when all contracting parties face incentive constraints, as in the problem of moral hazard in teams.

Let s^f be the maximum surplus in the above one-period full contracting problem. To have a non-trivial problem, I assume this surplus is greater than the default surplus:

$$s^f > s_0. \tag{3}$$

In the repeated model the principal and agent face similar incentive constraints and joint punishments are also needed for incentive provision. I assume that explicit money burning is impossible in this context. Instead joint punishments are achieved through surplus destruction. To see this more clearly, suppose an optimal self-enforcing contract achieves the maximum surplus s^f.

Now by adding the principal's *invariant* payoff in (PIC) to the agent's payoff in (AIC) for each y, we can rewrite the agent's incentive constraint (AIC) as (I drop the h^t everywhere)

$$e_t \in \arg\max_e \int_{\underline{y}}^{\overline{y}} \delta s_{t+1}\big(w(y), \hat{y}(y)\big) f(y|e)dy - (1-\delta)g(e), \tag{4}$$

where $s_{t+1}(w(y), \hat{y}(y)) = u(w(y), \hat{y}(y)) + v(w(y), \hat{y}(y))$ is the *continuation surplus*. Incentive constraint (4) makes it clear that in order to motivate above-minimum-level effort it is necessary to vary continuation surplus across output y, as pointed out by Levin (2003) and MacLeod (2003). Intuitively, since the principal's payoff does not vary with y due to incentive compatibility, variations in the agent's payoff must correspond to variations in surplus. Then for some y, continuation surplus must fall below the maximum level s^f, i.e surplus must be destroyed. Destruction of surplus may be viewed as conflicts in the relationship and they are an integral part of optimal intertemporal incentive provision. But their effectiveness rests on the parties' commitment to carry out inefficient actions ex post. The goal of the present paper is to examine the problem when the two parties lack such commitments and always try to resolve their conflicts by renegotiating their agreements.

Before I proceed with the subsequent analysis, it is instructive to point out that surplus destruction need not occur in optimal contracts so renegotiation poses no additional challenge if output, though non-verifiable, is *mutually observable*. In this case optimal relational contracts can be structured so that every continuation contract is itself optimal (Levin, 2003). For instance, let $w^{FB}(y)$, e^{FB} be first-best wage schedule and effort that maximize total surplus when output is verifiable. Then when the two parties are sufficiently patient there is a stationary optimal relational contract as follows: the principal offers the same wage schedule $w^{FB}(y)$ and the agent chooses e^{FB} in every period as long as no one deviates; if either party deviates then in the next period the deviator needs to make a fixed transfer so that the deviator's payoff is reduced to the default level but the incentive pay schedule and effort still remain the same. This way, continuation equilibria from the start of every period, on or off equilibrium path, are all optimal. This sequential optimality ensures that the contract is immune from renegotiation.

In contrast, as explained in the above when output is only privately observed optimal contracts are not sequentially optimal. Specifically, Levin (2003) describes a class of optimal contracts as follows: the principal continues to offer the same wage schedule to the agent and the agent chooses the same effort as long as the relationship has not terminated; if output falls below a threshold then the principal would pay a low discretionary wage and the relationship terminates. In contrast, Zhao (2009) describes another class of optimal contracts that do not resort to termination but rather rely on carrot and stick regimes to enforce cooperation. In either case, some continuation contracts are necessarily suboptimal, making renegotiation a nontrivial matter. The latter class of contracts will play a crucial role in the subsequent analysis.

To study the effects of renegotiation, following the literature I focus on contracts that are renegotiation proof in certain sense. For starters, sequential optimality inspires a natural notion of renegotiation-proofness. The idea is easiest to describe with finite horizon. In the final period T, the agreement should be Pareto optimal, for otherwise it will be renegotiated; then in period $T-1$, the remaining two-period contract should be Pareto optimal subject to the condition that continuation contracts in period T are Pareto optimal. This *Pareto perfection* can be applied backwards period by period to obtain renegotiation-proof contract in period $t=0$. The idea can be adapted to infinite horizon, giving rise to related renegotiation-proof concepts as proposed by Farrell and Maskin (1989) (FM), Bernheim and Ray (1989) (BR), and Ray (1994).

Pareto perfection however turns out to be too demanding in the current context. For one thing, a basic requirement of the concept is that continuation payoffs of a renegotiation-proof contract must not be Pareto ranked; this is referred to as weak renegotiation-proofness (FM), internal consistency (BR), or internal renegotiation-proofness (Ray). The reasoning is quite intuitive: if some continuation contract is Pareto dominated by another continuation contract then agents would renegotiate to replace the former with the latter. Since utility is transferable in the current context, this requirement implies that continuation surplus must be constant across all histories. But we already know that without variation in continuation surpluses the agent cannot be given any incentive to exert effort. So the only possible renegotiation-proof contract is the repetition of static Nash equilibrium, which is for the two parties to take their outside options in every period.

Partly to avoid such a total collapse of cooperation and partly to explore the effects of renegotiation from a different angle, in this paper I model renegotiation-proofness differently from the Pareto perfection idea. I ask the following question: When the players renegotiate an ongoing contract at some point t, how will this action affect their renegotiation behavior in the future? The view I shall take is that renegotiation at time t effectively declares what payoffs are unacceptable. These "norms" will then be used by future "generations:" If generation t finds an agreement unacceptable by renegotiating it then future generations will find it unacceptable as well. This approach is originated in Pearce (1987) and has since been applied to symmetric repeated games (Abreu, Pearce, and Stacchetti, 1993). It turns out to be a fruitful approach in the current setting with transferable utilities: as will be seen below, it permits an acknowledgment of the undesirability of surplus destruction and at the same time avoids a total collapse of cooperation between the two parties.

Formally, suppose the two parties propose an alternative contract σ' to replace an ongoing continuation contract $\sigma|h^t$. The fact that σ' is being proposed would mean that any surplus below $s(\sigma')$ is *unacceptable* (presumably $s(\sigma|h^t) < s(\sigma')$!). Now, if the alternative contract σ' is to be believable in the first place its continuation surpluses should not be unacceptable themselves, i.e. every continuation surplus $s(\sigma'|h^t) \geq s(\sigma')$; otherwise, the parties would want to replace $\sigma'|h^t$ with σ'. A contract becomes renegotiation proof if it survives the challenge by such alternative σ'. This leads to the following concept by Pearce (1987).[2]

Definition 1. A self-enforcing contract σ is *renegotiation proof* if for every h^t there does not exist a self-enforcing contract σ' satisfying $s(\sigma'|\hat{h}^t) \geq s(\sigma')$, $\forall \hat{h}^t$, that improves the continuation contract $\sigma|h^t$, i.e. $s(\sigma') > s(\sigma|h^t)$. (Note that σ' can be the contract that lets the parties take their outside options in every period.)

The following lemma, which generalizes a previous result for symmetric games (Pearce, 1987; Abreu et al., 1993), shows that renegotiation-proof contracts can be characterized by the lowest surplus generated across histories.

Given a self-enforcing contract σ, let $\mathscr{S}(\sigma) = \{s \in \Re \,|\, s(\sigma|h^t) = s \text{ for some } h^t\}$ be the set of continuation surpluses attained by σ at all histories. The greatest lower bound of this set is denoted by $\inf \mathscr{S}(\sigma)$.

Lemma 1. *A self-enforcing contract σ is renegotiation proof if and only if*

$$\inf \mathscr{S}(\sigma) \geqslant \inf \mathscr{S}(\sigma')$$

for every self-enforcing contract σ'.

Proof. By Definition 1, if σ is not renegotiation proof then there exists some history h^t and some self-enforcing contract σ' such that $s(\sigma|h^t) < \inf \mathscr{S}(\sigma')$. Then $\inf \mathscr{S}(\sigma) < \inf \mathscr{S}(\sigma')$.

Suppose $\inf \mathscr{S}(\sigma) < \inf \mathscr{S}(\sigma')$ for some self-enforcing contract σ'. Then there exists some history h^t with $s(\sigma|h^t) < \inf \mathscr{S}(\sigma')$, so σ is not renegotiation proof. □

[2] In Section 6 I discuss further the relation between this concept and other renegotiation-proof concepts in the literature.

In conclusion, renegotiation-proofness makes the players' collective preference *Rawlsian* in the following sense. If we treat the two parties as a single entity or a generation at every history h^t, then the society is populated by all these generations, one at each h^t; the Rawlsian objective is to maximize the welfare of the least-well-off of all these entities. This idea, by Lemma 1, is precisely embodied in Definition 1. Thus, here renegotiation is a justification for the Rawlsian-type social preference when the "individuals" of a society are the different generations of its citizens. Renegotiation-proofness also lends a specific meaning to conflict resolution: minimize conflict to the point of achieving the best surplus in the worst conflict.

4 Renegotiation-Proof Contracts and Low Morale Mechanism

Thanks to Lemma 1, renegotiation-proof contracts can be characterized by the *maximin* level of continuation surplus. To find the maximin surplus, I first identify an upper bound of the minimum continuation surplus, $\inf \mathscr{S}(\sigma)$, of any self-enforcing contract σ. The result can be derived heuristically as follows. Let $\bar{s}$ be the maximum continuation surplus along the path of contract σ. Then $\bar{s}$ can be decomposed into the following sum of current net output and next-period surpluses:

$$\bar{s} = \int_{\underline{y}}^{\bar{y}} [(1-\delta)(y-g(e)) + \delta s(y)] f(y|e) dy.$$

Subtracting $\delta\bar{s}$ from both sides yields

$$(1-\delta)\bar{s} = \int_{\underline{y}}^{\bar{y}} [(1-\delta)(y-g(e)) - \delta(\bar{s}-s(y))] f(y|e) dy.$$

Here $(1-\delta)\varphi(y) \equiv \delta(\bar{s}-s(y))$ is the (normalized) surplus destruction when output equals y. Note that $-\varphi(y)$ is effectively the agent's (non-normalized) incentive pay schedule, differing from it only by a constant term. If we let $\underline{s}$ be the minimum continuation surplus along the path of σ, then surplus destruction $(1-\delta)\varphi(y)$ cannot exceed the maximum possible amount, given by

$$(1-\delta)M \equiv \delta(\bar{s}-\underline{s}).$$

Rearrange this equation to express the minimum surplus $\underline{s}$ as

$$\underline{s} = \bar{s} - \frac{1-\delta}{\delta}M = \int_{\underline{y}}^{\bar{y}} (y-g(e)-\varphi(y)) f(y|e) dy - \frac{1-\delta}{\delta}M.$$

To obtain the bound on maximin surplus, we maximize the last term subject to the agent's incentive constraint and constraints on surplus destruction. Specifically, define the following optimization program and let s_* be its optimal value.

Program (P):

$$s_* = \max_{e,\varphi(\cdot),M} \int_{\underline{y}}^{\overline{y}} (y - \varphi(y)) f(y|e) dy - g(e) - \frac{1-\delta}{\delta} M$$

subject to:

$$e \in \arg\max_{e' \in \mathbb{E}} \int_{\underline{y}}^{\overline{y}} -\varphi(y) f(y|e') dy - g(e'), \tag{5}$$

$$0 \leqslant \varphi(y) \leqslant M, \ \ \forall y \in Y.$$

The following theorem proves the validity of s_* as an upper bound of minimum surpluses. The proof expands on the above intuition to take care of the possibility that $\bar{s}$ may not be attained.

Theorem 1. *If σ is self-enforcing then*

$$\inf \mathscr{S}(\sigma) \leqslant s_*.$$

Proof. See the appendix. □

Program (P) makes it clear that when selecting contracts the trade-off is between providing incentive for the agent and limiting the size of the worst surplus destruction.

Theorem 1 and Lemma 1 immediately imply the following characterization of renegotiation-proof contracts.

Theorem 2. *If σ is self-enforcing and $\inf \mathscr{S}(\sigma) = s_*$ then σ is renegotiation proof and every renegotiation-proof contract σ' satisfies $\inf \mathscr{S}(\sigma') = s_*$.*

Thus to establish s_* as the maximin level of surplus, we need only exhibit a contract that attains it. Indeed the logical question next is precisely how renegotiation-proof contracts look like beyond this maximin level of surplus. In the remainder of this section I construct a class of renegotiation-proof contracts that feature what shall be called low-morale mechanism.

To proceed, I assume that the discount factor δ is sufficiently close to one so that

$$s_* \geqslant s_0. \tag{6}$$

This is possible because at $\delta = 1$, Program (P) and the one-period full contracting problem in (2) are identical, so by assumption (3), $s_* = s^f > s_0$. The condition ensures that the relationship can generate enough surplus to support nontrivial self-enforcing contracts.

The renegotiation-proof contracts that I will construct have the stick-and-carrot flavor as in Abreu (1988). In each period the principal proposes the same menu that consists of just two wage levels: a guaranteed wage w_1 and a wage $w_2 = w_1 + b$ that includes a bonus $b > 0$. The agent decides whether to accept the menu. A promised conflict phase upon deviation from the agreement will ensure that the principal will propose the correct menu and the agent will accept it. Then after observing the output, the principal decides whether to pay the bonus or not. In these contracts the principal need not report the realized output. Conditional on the principal's choice of wage payment the relationship then enters into either a cooperative or a conflict phase, which will be described in details below.

Let $(e^*, \varphi^*(y), M^*)$ solve Program (P). Recall that $s(\underline{e}) = E(y|\underline{e}) - g(\underline{e})$ is the surplus generated by the minimum effort $\underline{e}$. Note that $s_* \geqslant s_0 \geqslant s(\underline{e})$. Construction of the renegotiation-proof contract depends on the size of surplus loss when the agent chooses the minimum effort $\underline{e}$ for one period. There are two alternative cases.

Case A: $s_* - s(\underline{e}) \leqslant M^*$.

In this case, the surplus loss $s_* - s(\underline{e})$ is relatively small. The following strategies involve two phases and the contract starts with the cooperative phase at $t = 0$. Regardless of which phase the current play is in, the principal always proposes the same wage menu that contains just two wages $w_2 > w_1$, which are to be determined below, and the agent always accepts it. If one of them deviates, i.e. the principal offers the wrong menu or the agent rejects the correct menu, then the conflict phase will be played in the next period.

- *Cooperative Phase.* The agent chooses effort e^*; the principal pays the agent w_i, $i = 1, 2$, with probability $\beta_i(y)$ if realized output is y, where for all y, $\beta_1(y) + \beta_2(y) = 1$ and $\beta_1(y)$ is determined by

$$\beta_1(y) = \frac{\varphi^*(y)}{M^*}. \tag{7}$$

  If wage payment turns out to be w_2, continue with cooperative phase; if wage payment is anything other than w_2, switch to conflict phase as specified below.

- *Conflict Phase.* The agent chooses minimum effort $\underline{e}$ and the principal pays constant wage w_1 for one period. In the next period, with probability $1 - \alpha$ switch to cooperative phase; with probability α restart the conflict phase. (Probability α will be determined subsequently.)

Case B: $s_* - s(\underline{e}) > M^*$.

The loss of surplus by choosing minimum effort $\underline{e}$ is more severe than in Case A and such a punishment for one period is already too much. The proposed contract only carries out the punishment with probability less than one. The strategies again consist of two phases and the contract starts with the cooperative phase at $t = 0$. Regardless of which phase the current play is in, the principal always proposes the same wage menu that contains two wages $w_2 > w_1$ and the agent always accepts. If one of them deviates, i.e. the principal offers the wrong menu or the agent rejects the correct menu, then play the conflict phase in the next period.

- *Cooperative Phase.* The agent chooses effort e^*; the principal pays the agent according to schedule $(\beta_i(y))$, which is given by (7). If wage payment turns out to be w_2, continue with cooperation phase; otherwise, switch to conflict phase specified below.
- *Conflict Phase.* With probability α, the agent chooses effort $\underline{e}$ and the principal pays constant wage w_1 regardless of current performance, then switch to cooperative phase; with probability $1-\alpha$, start the cooperative phase.

In both cases, once the offer is accepted the contractual minimum wage w_1 must be paid at the end of the period; the principal only decides whether or not to pay the bonus $b = w_2 - w_1$. As discussed previously, w_1 can be chosen so as to satisfy participation constraints given players' default payoffs. Thus, to fully characterize the contracts one only needs to determine the probability of invoking the conflict phase, α, and the bonus b.

Remark 1. Here public randomization is used to simplify the construction of equilibrium strategies. It is possible to construct alternative strategies without using pubic randomization, as is done in a related paper (Zhao, 2009).

In what follows I offer a heuristic derivation of these parameters; details can be found in the appendix. Consider Case A (Case B can be analyzed similarly). The surplus in the conflict phase should equal the maximin surplus s_*, and the surplus in the cooperative phase, denoted s_2, should satisfy

$$\delta s_2 = \delta s_* + (1-\delta)M^*.$$

Moreover, by construction, s_*, s_2, and the transition probability α must satisfy the recursion

$$s_* = (1-\delta)s(\underline{e}) + \delta(\alpha s_* + (1-\alpha)s_2).$$

These conditions lead to the following value of probability α,

$$1-\alpha = \frac{s_* - s(\underline{e})}{M^*}.$$

To determine the bonus b, observe that the principal's payoff in a conflict phase, denoted by v_1, and her payoff in a cooperative phase, denoted by v_2, should satisfy two self-generating conditions as well as the incentive constraint as follows:

$$v_2 = E(y|e^*) - w_2,$$

$$v_1 = (1-\delta)(E(y|\underline{e}) - w_1) + \delta(\alpha v_1 + (1-\alpha)v_2),$$

$$\delta(v_2 - v_1) = (1-\delta)(w_2 - w_1).$$

These conditions pin down the bonus as

$$b = \frac{\delta(E(y|e^*) - E(y|\underline{e}))}{1+(1-\alpha)\delta}.$$

Finally, one can show that the probabilities $\beta_i(y)$, the transition probability α, and the bonus b defined in the above induce a payment schedule for the agent that exactly replicates the one that, together with e^*, solves Program (P). As such, the agent will have no problem choosing the proposed effort e^* in the cooperative phase.

Theorem 3. *The relational contract defined in the above, with the parameters $\beta_i(y)$ and $b = w_2 - w_1$ as determined, and an appropriately chosen $w_1 \in \Re$, is self-enforcing and renegotiation proof.*

Proof. See the appendix. □

Remark 2. A main purpose of this paper is to show how renegotiation-proof contracts can help explain some important characteristics of relational contracts such as the occurrence of "low morale," which happens when the agent chooses low effort in responding to low pay. Here low morale is "useful" because it imposes a cost on the principal and discourages her from falsely reporting the agent's output. This phenomenon by itself is not unfamiliar in repeated games with imperfect observations. What is more important is that the low-morale episodes are an integral part of renegotiation-proof contracts, even though the two parties always have the option to terminate the relationship. This is because the maximin surplus s_* is often larger than the default surplus from terminating the relationship, s_0, so termination simply is not renegotiation proof — this point will be explicitly demonstrated in the next section. Thus renegotiation-proof contracts offer an economic explanation for conflicts in organizations and in other types of long-term relationships where subjective evaluation and discretionary compensation play an important role in providing incentives.

5 Renegotiation and Incentive Contract

Renegotiation gears the objective of optimal contracting towards the maximin surplus. Does this make a difference? Is the maximin surplus s_* larger than the minimum continuation surplus of optimal contracts without renegotiation? Moreover, how does renegotiation change the incentive contract within a single period? I answer these questions in the present section.

Specifically, this section has two goals. The first is to characterize the wage schedule in renegotiation-proof contracts, which will provide insights into the effect of renegotiation on the incentive system; the second is to show by example that renegotiation does have a bite on optimal contract. For these purposes, I assume that$f_e(y)/f(y)$ is an increasing function of y for all e and that $F_{ee} > 0$. The former is the continuous version of the familiar monotone likelihood ratio condition; the latter is the so-called convexity of distribution function condition. These Mirrlees-Rogerson conditions ensure that the agent's incentive constraint (5) can be characterized by a first-order condition.

Recall that the principal's wage payment consists of a guaranteed base wage w_1 and a discretionary bonus b. The key component of the wage schedule is the probability $\beta_2(y)$ of paying bonus b, for each y. To characterize this probability, we revisit Program (P).

Let $(e^*, \varphi^*(\cdot), M^*)$ solve Program (P). Then given $e = e^*$ and assuming e^* is in the interior of $\mathbb{E}$, the pair $(\varphi^*(\cdot), M^*)$ solves the following[3]

Program (P')

$$\max_{\varphi(\cdot),M} \int_{\underline{y}}^{\overline{y}} (y - \varphi(y)) f(y|e) dy - g(e) - \frac{1-\delta}{\delta} M$$

s.t.

$$\int_{\underline{y}}^{\overline{y}} (-\varphi(y)) f_e(y|e) dy - g'(e) = 0,$$

$$M \geqslant \varphi(y) \geqslant 0, \ \forall y \in [\underline{y}, \overline{y}].$$

Note that the agent's incentive constraint now takes the form of a first-order condition. By previous construction, the probability $1 - \beta_2(y) = \varphi^*(y)/M^*$ is fully determined by the solution to this program. The detailed analysis of this problem, which requires some optimal control theory, is relegated to the appendix.

Theorem 4. *Suppose $(\varphi^*(\cdot), M^*)$ solves Program (P'). Then there exists $\hat{y}$ and $\lambda > 0$ such that*

$$\varphi^*(y) = \begin{cases} 0 & \text{if } y > \hat{y} \\ M^* & \text{if } y \leqslant \hat{y}, \end{cases} \tag{8}$$

[3] We drop the $*$ in e^* to lessen notational burden.

$$1+\lambda\frac{f_e(\hat{y})}{f(\hat{y})}=0, \tag{9}$$

$$\frac{1-\delta}{\delta}=-F(\hat{y})-\lambda F_e(\hat{y}), \tag{10}$$

and

$$-M^{*}F_e(\hat{y})=g'(e). \tag{11}$$

Moreover, as functions of δ, $M^{*}(\delta)$ increases and $\hat{y}(\delta)$ decreases as δ increases.

Proof. See the Appendix. □

Given effort e, the four conditions (8)-(11) pin down $\varphi(y), M^{*}, \hat{y}$, and λ. Combining Theorems 3 and 4, we see that the renegotiation-proof contract follows a simple cutoff rule: If performance is above the threshold $\hat{y}$, the agent is paid the base wage *and* the bonus and the relationship stays in cooperative phase. If performance falls below $\hat{y}$, the agent is only paid the base wage and the relationship enters into a conflict or "low morale" phase.

To demonstrate how renegotiation affects pay-performance relation and optimal contract in general, it is helpful to consider the contracting problem *without* renegotiation. In this case the maximum surplus that can be generated is found by solving the following program (also see Levin (2003)).

Program (P")

$$\max_{e,\varphi(\cdot)} s=\int_{\underline{y}}^{\overline{y}}(y-\varphi(y))f(y|e)dy-g(e)$$

s.t.

$$\int_{\underline{y}}^{\overline{y}}(-\varphi(y))f_e(y|e)dy-g'(e)=0$$

$$\delta(s-s_0)\geqslant(1-\delta)\varphi(y)\geqslant 0,\ \forall y.$$

The last is a self-enforcing constraint; it says the surplus that can be destroyed today, $(1-\delta)\varphi(y)$, cannot exceed the net surplus generated by the relationship from tomorrow onwards, $\delta(s-s_0)$—recall s_0 is the surplus from termination. A similar constraint is embedded in Program P'. Note that there is a floor on continuation surpluses with or without renegotiation. The difference is that without renegotiation the floor s_0 is exogenous; with renegotiation the floor s_* is endogenous.

Let $\check{s}$ be the lowest continuation surplus of optimal contracts without renegotiation. Clearly renegotiation makes an impact if the maximin surplus s_* is greater than $\check{s}$. We demonstrate this scenario in the following example. Assume output y is drawn from the interval $[0,1]$ according to the distribution function $F(y,e) = y^e$, where effort $e \in [0,\bar{e}]$. This function satisfies the Mirrlees-Rogerson conditions. Assume the constrained efficient effort is in the interior, so the incentive constraint is characterized by the first-order condition.

Proposition 1. *For this parameterized example, $s_* > \check{s}$.*

Proof Sketch. The following line of argument illustrates how renegotiation impacts the incentive contract.

Step 1. Without renegotiation the lowest continuation surplus $\check{s} = s_0$.

First, the solution $(e, \varphi(y))$ to Program P'' also has a cutoff form as in Program P':

$$\varphi(y) = \begin{cases} 0 & \text{if } y > \tilde{y} \\ m & \text{if } y \leqslant \tilde{y}. \end{cases}$$

for some $\tilde{y}$; Then by the incentive constraint,

$$m = -\frac{g'(e)}{F_e(\tilde{y},e)} = -\frac{g'(e)}{\tilde{y}^e \log(\tilde{y})}.$$

Moreover, the self-enforcing constraint $\delta(s - s_0) \geqslant (1-\delta)\varphi(y)$ *must* bind for $0 \leq y \leq \tilde{y}$. To see this, notice that the cutoff-form solution entails a surplus loss equal to

$$L(\tilde{y}) \equiv \int_{\underline{y}}^{\tilde{y}} \varphi(y) f(y|e) dy = mF(\tilde{y},e) = -\frac{g'(e)F(\tilde{y},e)}{F_e(\tilde{y},e)} = -\frac{g'(e)}{\log(\tilde{y})}.$$

Therefore, by choosing $\tilde{y}$ arbitrarily small and hence m arbitrarily large, the loss can be made arbitrarily small. Of course m cannot be arbitrarily large due to the self-enforcing constraint; instead the constraint must bind at the solution so $m = \delta(s - s_0)/(1-\delta)$. Thus $\check{s} = s_0$, as we can construct a contract achieving surplus s, with minimum continuation surplus being $s - \frac{1-\delta}{\delta} m = s_0$.

In the remaining steps, we show that the maximin surplus $s_* > s_0$. Intuitively, we can increase the cutoff $\tilde{y}$ and decrease the maximum surplus destruction m, which has the potential to raise the minimum continuation surplus $s - \frac{1-\delta}{\delta} m$. The problem is that such a plan also reduces the surplus s, leaving the net outcome uncertain. The remaining steps will clear up this uncertainty and verify that the plan indeed works.

First, given an arbitrary cutoff point y, define the following schedule

$$\varphi(y) = \begin{cases} 0 & \text{if } y' > y \\ m = -\frac{g'(e)}{F_e(y,e)} = -\frac{g'(e)}{y^e \log(y)} & \text{if } y' \leqslant y. \end{cases}$$

Given this schedule, the agent's incentive constraint is satisfied for the same effort. Define

$$\begin{aligned}\Delta(y) &= \delta(s-s_0)-(1-\delta)m \\ &= \delta\left[\int_0^1 zf(z|e)dz - g(e) - L(y) - s_0\right] + (1-\delta)\frac{g'(e)}{y^e \log(y)} \\ &= \delta\left[\int_0^1 zf(z|e)dz - g(e) - s_0\right] + \delta\frac{g'(e)}{\log(y)} + (1-\delta)\frac{g'(e)}{y^e \log(y)}.\end{aligned}$$

Then the self-enforcing constraint is equivalent to $\Delta(y) \geq 0$. Recall that at the solution to Program P", the cutoff $\tilde{y}$ satisfies $\Delta(\tilde{y}) = 0$.

Step 2. The derivative $\Delta'(y)$ satisfies strict single-crossing: as the cutoff y increases, $\Delta'(y)$ changes sign at most once at a single point from positive to negative, i.e. $\Delta'(y_1) \leq 0 \Longrightarrow \Delta'(y_2) < 0, \forall y_2 > y_1$ and $\Delta'(y_2) \geq 0 \Longrightarrow \Delta'(y_1) > 0, \forall y_1 < y_2$.

This can be seen from

$$\begin{aligned}\Delta'(y) &= -\delta\frac{g'(e)}{y(\log(y))^2} - (1-\delta)(1+e\log(y))\frac{g'(e)}{y^{e+1}(\log(y))^2} \\ &= \frac{g'(e)}{y(\log(y))^2}\left(-\delta-(1-\delta)\frac{1+\log(y^e)}{y^e}\right).\end{aligned}$$

The result follows since $\frac{1+\log(y^e)}{y^e}$ strictly increases in y.

Step 3. $\Delta'(\tilde{y}) > 0$.

Suppose $\Delta'(\tilde{y}) < 0$. Choose a cutoff y' slightly below $\tilde{y}$. Since $\Delta(\tilde{y}) = 0$, we have $\Delta(y') > 0$. In other words, the lower cutoff still satisfies the self-enforcing constraint. But it also lowers the loss and hence raises the net surplus! This contradicts the fact that $\tilde{y}$ is the optimal cutoff. The case $\Delta'(\tilde{y}) = 0$ is nongeric, as the two variables $e, \tilde{y}$ would need to satisfy three equations (the other two are the incentive and self-enforcing constraints). If this case does arise, then there is a nearby pair $(\hat{e}, \hat{y})$ that satisfies the incentive constraint and satisfies $\Delta(\hat{y}) > 0$; the next step then applies to $\hat{e}$ and $\hat{y}$ (in place of y' below).

Step 4. $s_* > s_0$.

Now raise the cutoff from $\tilde{y}$ to some $y' = \tilde{y} + \varepsilon$ and let $m' = \frac{g'(e)}{F_e(y', e)} < m$. Then, since $\Delta'(\tilde{y}) > 0$ and $\Delta(\tilde{y}) = 0$, we have $\Delta(\tilde{y} + \varepsilon) > 0$. Thus this new schedule satisfies the agent's incentive constraint for the same effort and it relaxes the self-enforcing constraint

$$\delta(s-s_0) \geqslant (1-\delta)\varphi(y).$$

We can then construct a low-morale contract as in the previous section, with the lowest continuation surplus equal to $s - \frac{1-\delta}{\delta}m' > s_0$. Then of course the maximin $s_* > s_0$.

Taking Steps 1 and 4 together, we have $s_* > \check{s}$. □

The above example and the argument for Proposition 1 demonstrate how renegotiation affects optimal contract. The need to maximize the minimum continuation surplus generally compresses variation in continuation surpluses. This further requires a lower-powered incentive contract *in every period*. Thus, compared with optimal contract without renegotiation, renegotiation-proof contract punishes the agent less severely but for a larger range of low performance levels, or equivalently, rewards the agent less generously for a small set of high performance levels. (In the example, renegotiation rasies the cutoff $\tilde{y}$ and extends the punishment region $[0,\tilde{y}]$.) In this sense, the incentive contract becomes *low powered* because of renegotiation-proofness. Thus one could say that the desire to reduce conflict leads to the adoption of low-powered incentive systems. This point, often informally invoked in the discussion of performance pay, finds some theoretical underpinnings in the present model.

Finally, a technical point is that as the principal and agent become more patient, the cost of renegotiation decreases. In particular, conflict becomes less often (threshold $\hat{y}$ decreases) but the cost of a conflict (the size of M^*) becomes larger. The contract thus gets closer to the first-best contract.

6 Further Discussion on Renegotiation-Proofness

Definition 1 lets the agents consciously ponder the consequence of renegotiating an ongoing contract and in the process reach a decision not to renegotiate it. In fact, Definition 1 can be justified by equilibrium play with the use of the following axiom.

> *Broken-Pact Axiom.* If in period t the act of renegotiation by the current incarnation of the agents imply that payoffs in some set B are unacceptably low but every cooperative agreement beyond static Nash must prescribe some continuation payoff in B, then "the pact is broken": the parties believe that cooperation is no longer possible and they will simply play the best static Nash forever (in the current setting this means terminating the relationship).

Now if σ satisfies Definition 1, then $\inf \mathscr{S}(\sigma) \geqslant \inf \mathscr{S}(\sigma')$, for every self-enforcing σ' (Lemma 1). At each h^t, an attempt to renegotiate $\sigma|h^t$ will make any surplus less than or equal to $s(\sigma|h^t)$ unacceptable. But since $s(\sigma|h^t) \geqslant \inf \mathscr{S}(\sigma) \geqslant \inf \mathscr{S}(\sigma')$, $\forall \sigma'$, every self-enforcing contract, including repetition of static Nash, has some continuation surplus that is no more than $s(\sigma|h^t)$; thus renegotiating $\sigma|h^t$ would invoke the broken-pact axiom and cause the relationship to terminate, resulting in the surplus s_0, which is no more than $\inf \mathscr{S}(\sigma) \leqslant s(\sigma|h^t)$ in the first place. Given such a prospect, the two parties would choose not to renegotiate any $\sigma|h^t$; hence σ is renegotiation proof! In conclusion, this approach imposes an explicit constraint on players' desire to renegotiate an ongoing agreement, which is the fear that future generations will do the same, to the harm of the players' current welfare.

Relation to the Literature

There are a number of renegotiation-proof concepts proposed in the literature (see Bergin and MacLeod (1993) for a survey). I have already discussed one family of concepts, including the popular strong and weak renegotiation-proofness concepts proposed by Farrell and Maskin (1989) in Section 3. These concepts share a common characteristic: the set of renegotiation-proof payoffs is a subset of its own Pareto frontier. As explained in Section 3, in the present model this means continuation surpluses must be constant across all histories, which leaves no room for incentivizing the agent.

The Pearce renegotiation-proof concept in Definition 1 embodies a certain kind of history dependence. Bergin and MacLeod (1993) introduce a related concept, *recursive efficiency*, which also exhibits history dependence. Briefly speaking, an agreement φ specifies a set of permissible payoffs $\varphi(h)$ at every history h. Let $\mathscr{B}(\varphi,h)$ be the set of payoffs that can be generated at history h by continuation payoffs specified by the agreement. Then the agreement is recursively efficient if for all h the specified payoffs $\varphi(h)$ belong to the efficient frontier of $\mathscr{B}(\varphi,h)$. Recursive efficiency depends on the agreement φ that is already in place. For instance, in the present model repetition of static Nash is recursively efficient: At every history, the agreement specifies the same pair of default payoffs (u_0,v_0). The only payoff vector that can be generated by such continuation payoffs is (u_0,v_0).

Bergin and MacLeod (1993) also introduce another concept, which is closer to the concept used here. A self-enforcing contract is *undominated* if there does not exist another self-enforcing contract that delivers higher surplus at some history and delivers at least the same amount of surplus at all subsequent histories. This concept treats the contract in place as a social norm; the challenging contract must perform at least as well at all histories. In comparison, Definition 1 does not place that much emphasis on the contract in place but instead requires the challenging contract itself to be logical consistent, i.e. if the parties propose to move to a new contract σ' then all continuation surplus of the new contract should be at least as high as $s(\sigma')$ itself; otherwise, some continuation contract of σ' suffers the same challenge from σ' itself. With infinite horizon, contracts satisfying Definition 1 are easier to characterize (cf. Lemma 1).

Finally, most of the studies in the renegotiation-proof literature, including the present paper, try to identify a set of payoffs that are immune from further renegotiation, but leave open the question as to which particular payoff point (and associated equilibrium) should be chosen. It is possible to add an initial stage during which the agents would bargain over which point should be chosen. The outcome will then depend on the agents' relative bargaining power. MacLeod and Malcomson (1989) characterize self-enforcing contracts in a market context where one party has all the bargaining power due to unbalanced demand and supply. In general, fully incorporating bargaining into the renegotiation process is a challenging exercise. Abreu et al. (1993) combine renegotiation with equal bargaining power in symmetric games and show that this combination implies strong symmetry of the renegotiation-proof equilibria. More ambitiously, bargaining and (re)negotiation may be regarded as the general means for selecting the ultimate equilibrium path in repeated games. See Abreu and Pearce (2007) for advances on this front.

7 Conclusion

This paper studies a simple repeated principal-agent model with private evaluation. To provide incentive for the agent to exert effort, relational contracts necessarily destroy surpluses in the form of inefficiently low effort following low pay. When the two parties cannot bind themselves to such ex post inefficient agreements, they try to negotiate away from them. The effective outcome that can survive such renegotiations is called renegotiation proof. Renegotiation-proofness imposes an endogenous floor on the surplus from any continuation contract. This floor generally is above the default level of surplus from terminating the relationship. As such, recurrent low morale becomes an indispensable part of renegotiation-proof contracts. These results shed light on periodic conflicts in organizations and in business relationships where subjective and private evaluations are important and commitment to harsh punishments is difficult to sustain.

My results also show that in such environments the incentive contracts may appear to be low powered compared with the situation when evaluation is subjective and renegotiation is absent. Since low-powered incentive contracts and subjective evaluation are commonly observed at the same place, an explanation for their co-existence is important. As pointed out in the introduction, subjective evaluation does not automatically render incentive contracts useless, because there are ways to circumvent the issue of verifiability and enforceability and to provide essentially the same kind of incentives for the agent as the case when evaluation is objective. By taking into account of the possibility of renegotiation, the result in this paper is able to offer an explanation for the co-existence of low-powered incentive contracts and subjective evaluation.

An important extension of the present model is to allow risk-averse agents. The contract then has to take into consideration intertemporal risk sharing.

Also, in the present paper my analysis focuses on public equilibrium and renegotiation is allowed in every period. It would be interesting to consider the case when the principal could delay her report on the agent's performance and accordingly renegotiation is permitted only infrequently. From the work of Abreu et al. (1991), Fuchs (2007), and others we know that infrequent information release will enhance efficiency in the current model. On the other hand, infrequent information release also relies on harsher punishments to enforce cooperation. Then renegotiation could play a bigger role as it limits the severity of punishment and thereby restrict the delay of information release. This is a potentially interesting avenue for future research.

Appendix: Proofs Omitted from the Main Text

Proof of Theorem 1. Let $\bar{s} = \sup \mathscr{S}(\sigma)$ be the least upper bound of continuation surpluses of σ. Fix an $\varepsilon > 0$. Let $\sigma|h^t$ be a continuation contract that delivers a pair of payoffs (u, v) such that the total surplus $s = u + v \in [\bar{s} - \varepsilon, \bar{s}]$.

Since the continuation contracts $\sigma|(h^t, w(y), \hat{y}(y))$ are self-enforcing, the payoffs (u, v) delivered by $\sigma|h^t$ can be decomposed into the current payoffs plus the continuation payoffs $u(w(y), \hat{y}(y)), v(w(y), \hat{y}(y))$ (to lessen notational burden, I suppress dependence on h^t whenever possible) as follows

$$u = \int_{\underline{y}}^{\bar{y}} \left[(w(h^t, y) - g(e(h^t)))(1-\delta) + \delta u(w(y), \hat{y}(y))\right] f(y|e(h^t))dy$$

$$v = \int_{\underline{y}}^{\bar{y}} \left[(y - w(h^t, y))(1-\delta) + \delta v(w(y), \hat{y}(y))\right] f(y|e(h^t))dy.$$

Moreover, the continuation payoffs and current strategies $e(h^t), w(h^t, y), \hat{y}(h^t, y)$ should satisfy the principal's incentive constraint (PIC) and the agent's incentive constraint (AIC).

Recall that the continuation surpluses are given by $s(w(y), \hat{y}(y)) = u(w(y), \hat{y}(y)) + v(w(y), \hat{y}(y))$. The current surplus then is given by

$$s = \int_{\underline{y}}^{\bar{y}} \left[(1-\delta)(y - g(e(h^t)) + \delta s(h^t, w(y), \hat{y}(y))\right] f(y|e(h^t))dy.$$

Let $\underline{s} = \inf_y s(h^t, w(y), \hat{y}(y))$ and let $m = \bar{s} - \underline{s}$. Since $\underline{s} = \bar{s} - m \le s + \varepsilon - m$, we have

$$\underline{s} \leqslant \int_{\underline{y}}^{\bar{y}} \left[(1-\delta)(y - g(e(h^t))) - \delta(\bar{s} - s(h^t, w(y), \hat{y}(y)))\right] f(y|e)dy + \delta \underline{s} - (1-\delta)m + \varepsilon$$

Regrouping terms and dividing both sides by $(1-\delta)$, we have

$$\underline{s} \leqslant \int_{\underline{y}}^{\overline{y}} \left[y - \frac{\delta}{1-\delta}(\bar{s} - s(h^t, w(y), \hat{y}(y))) \right] f(y|e(h^t))dy - g(e(h^t)) - m + \frac{\varepsilon}{1-\delta}$$

Now let

$$\varphi(y) = \frac{\delta}{1-\delta}(\bar{s} - s(h^t, w(y), \hat{y}(y)))$$

$$M = \frac{\delta}{1-\delta}m.$$

We then observe that $e(h^t), \varphi(y)$ and M satisfy

$$\underline{s} \leqslant \int_{\underline{y}}^{\overline{y}} (y - \varphi(y)) f(y|e(h^t))dy - g(e(h^t)) - \frac{1-\delta}{\delta}M + \frac{\varepsilon}{1-\delta}$$

$$0 \leqslant \varphi(y) \leqslant M, \forall y.$$

Moreover, subtracting $\bar{s}$ from the agent's objective function in (4), we can rewrite the agent's incentive constraint (4) as

$$e(h^t) \in \arg\max_{e} \int_{\underline{y}}^{\overline{y}} (-\varphi(y)) f(y|e)dy - g(e).$$

Comparing the three preceding conditions with Program (P), we conclude that $\underline{s} - \frac{\varepsilon}{1-\delta} \leqslant s_*$ for all $\varepsilon > 0$. Therefore,

$$\inf \mathscr{S}(\sigma) \leqslant \underline{s} \leqslant s_*.$$

□

Proof of Theorem 3. Specifically, consider two alternative cases.

Case A: $s_* - s(\underline{e}) \leqslant M^*$.

The parameters $\alpha, b = w_2 - w_1$ will be determined below.

In this self-enforcing contract, let (u_2, v_2), (u_1, v_1) be the payoffs of the agent and the principal for the cooperative and conflict phases respectively. For this contract to be renegotiation proof, we need the continuation surpluses for the two phases to satisfy the following conditions:

$$s_1 \equiv u_1 + v_1 = s_* \tag{12}$$

$$\delta s_2 \equiv \delta(u_2 + v_2) = \delta s_* + (1-\delta)M^*. \tag{13}$$

Provided that the continuation payoffs are no less than the default payoffs, the proposed strategies need to satisfy the following necessary and sufficient conditions:

$$u_2 = \int \left\{ \sum_{i=1}^{2} \beta_i(y)[(1-\delta)w_i + \delta u_i] \right\} f(y|e^*)dy - (1-\delta)g(e^*) \tag{14a}$$

$$u_1 = (1-\delta)(w_1 - g(\underline{e})) + \delta(\alpha u_1 + (1-\alpha)u_2) \tag{14b}$$

$$v_2 = \int \left\{ \sum_{i=1}^{2} \beta_i(y)[(1-\delta)(y-w_i) + \delta v_i] \right\} f(y|e^*)dy \tag{14c}$$

$$v_1 = (1-\delta)(E(y|\underline{e}) - w_1) + \delta(\alpha v_1 + (1-\alpha)v_2) \tag{14d}$$

$$u_2 \geqslant \int \left\{ \sum_{i=1}^{2} \beta_i(y)[(1-\delta)w_i + \delta u_i] \right\} f(y|e)dy - (1-\delta)g(e), \quad \forall e \in \mathbb{E} \tag{14e}$$

$$u_1 \geqslant (1-\delta)(w_1 - g(e)) + \delta(\alpha u_1 + (1-\alpha)u_2), \quad \forall e \in \mathbb{E} \tag{14f}$$

$$-(1-\delta)w_2 + \delta v_2 = -(1-\delta)w_1 + \delta v_1 \geq -(1-\delta)w + \delta v_1, \forall w \geq w_1. \tag{14g}$$

Equations (14a) – (14d) ensure that the payoffs (u_2, v_2), (u_1, v_1) can be generated by the proposed strategies; (14e), (14f), (14g) are the incentive constraints for the agent and the principal.

Determining α. Adding (14b) to (14d) yields the recursion for surpluses

$$s_1 = (1-\delta)s(\underline{e}) + \delta(\alpha s_1 + (1-\alpha)s_2). \tag{15}$$

By design, $s_1 = s_*$ and $\delta s_2 = \delta s_1 + (1-\delta)M^*$. Replacing s_2 with this expression in (15), we can solve for α as follows,

$$\alpha = 1 - \frac{s_* - s(\underline{e})}{M^*}. \tag{16}$$

Determining $b = w_2 - w_1$. The base wage w_1 can be adjusted without affecting equilibrium strategies as long as continuation payoffs are no less than the default payoffs.

By the principal's incentive constraint (14g) and equation (13), it follows

$$(1-\delta)w_2 + \delta u_2 - (1-\delta)w_1 - \delta u_1 = \delta s_2 - \delta s_1 = (1-\delta)M^*. \tag{17}$$

Eq. (17), together with the agent's payoff recursions (14a), (14b), will pin down the bonus $w_2 - w_1$. Specifically, plug equation (17) into (14a) to eliminate u_1 and we have

$$\begin{aligned} u_2 &= \int \{(1-\delta)w_2 + \delta u_2 - \beta_1(y)[(1-\delta)w_2 + \delta u_2 - (1-\delta)w_1 - \delta u_1]\} f(y|e^*)dy \\ &\quad - (1-\delta)g(e^*) \\ &= \int \{-\beta_1(y)(1-\delta)M^*\} f(y|e^*)dy + (1-\delta)(w_2 - g(e^*)) + \delta u_2 \end{aligned}$$

By the definition of $\beta_1(y)$ in (7), this implies

$$u_2 = w_2 - E(\varphi^*|e^*) - g(e^*). \tag{18}$$

Substituting this back into equation (17) to solve for u_1 as follows:

$$u_1 = w_2 - E(\varphi^*|e^*) - g(e^*) - \frac{(1-\delta)}{\delta}(M^* - (w_2 - w_1)). \tag{19}$$

To find the desired bonus $w_2 - w_1$, substitute for u_1, u_2 using (18) and (19) in (14b) to get the following,

$$w_2 - w_1 - [E(\varphi^*|e^*) + g(e^*) - g(\underline{e})] = \frac{1-\alpha\delta}{\delta}[M^* - (w_2 - w_1)]. \tag{20}$$

Since by assumption the agent weakly prefers e^* to $\underline{e}$ given schedule $-\varphi(\cdot)$, we have

$$-E(\varphi^*|e^*) - g(e^*) \geqslant -E(\varphi^*|\underline{e}\,) - g(\underline{e}) > -M^* - g(\underline{e}) \tag{21}$$

then $M^* > E(\varphi^*|e^*) + g(e^*) - g(\underline{e})$. It follows that left-hand side of (20) is greater than the right-hand side if $w_2 - w_1 = M^*$ and is less than the right-hand side if $w_2 - w_1 = E(\varphi^*|e^*) + g(e^*) - g(\underline{e})$; moreover the left-hand (right-hand) side is increasing (decreasing) and continuous in $w_2 - w_1$. Therefore the unique bonus solving (20) must lie in the interval $(E(\varphi^*|e^*) + g(e^*) - g(\underline{e}), M^*)$. Specifically, by using equation (16) to substitute for α in equation (20), the bonus can be found as follows:

$$b = \frac{M^*[E(y|e^*) - E(y|\underline{e})]}{M^* + s_2 - s(\underline{e})}, \tag{22}$$

which, by using the value of α in (16), is equal to

$$b = \frac{\delta(E(y|e^*) - E(y|\underline{e}))}{1 + (1-\alpha)\delta},$$

as given in the text. Note that the bonus is a fraction of the extra output created relative to the minimum level.

The agent's payoffs then can be determined (up to the base wage w_1) by substituting the bonus $w_2 - w_1$ into equations (18) and (19). By construction, these payoffs satisfy the agent's payoff recursions (14a), (14b). Incentive constraint (14f) obviously follows from (14b). Incentive constraint (14e) also holds because by construction the agent's payoff when choosing an effort e is a linear transformation of the agent's objective function in Program P and e^* is the agent's best choice there.

To complete the proof, it only remains to find the principal's payoffs v_2, v_1 that satisfy payoff recursions (14c) and (14d) and incentive constraint (14g).

By (14c) and (14g), one can solve for the principal's payoffs as follows:

$$v_2 = E(y|e^*) - b - w_1, \tag{23}$$

$$v_1 = E(y|e^*) - \frac{b}{\delta} - w_1. \tag{24}$$

Now using definition of α and solution of $b = w_2 - w_1$, it is straightforward although a bit tedious to verify that these payoffs satisfy equation (14d). Moreover, base wage w_1 can be chosen so that $u_1 \geqslant u_0$ and $v_1 \geqslant v_0$. It follows that the contract is self-enforcing on the equilibrium path, i.e. as long as the correct wage menu is offered.

To show that the principal has no incentive to offer any other wage menu and the agent has no incentive to reject the correct menu, we note that if either of them indeed deviates in such a way then the conflict phase starts for sure in the next period, generating continuation surplus s_*. On the other hand, if they do not deviate then there is always a chance that the cooperative phase will start in the next period so the expected continuation surplus is strictly higher than s_*. Thus, for δ close enough to one, any short-term gain from deviation is overwhelmed by potential loss from switching to the conflict phase.

In summary, the proposed contract σ is indeed self enforcing. By construction, the minimum surplus equals the maximin value: $\inf \mathscr{S}(\sigma) = s_*$. Thus the contract is also renegotiation proof.

Case B: $s_* - s(\underline{e}) > M^*$ or equivalently $\delta(\bar{s} - s(\underline{e})) > M^*$.

The idea again is to generate surplus $s_1 = s_*$ in the conflict phase and $s_2 = s_* + \frac{1-\delta}{\delta} M^*$ in the cooperative phase. These surpluses should satisfy the following recursion:

$$s_1 = \alpha [(1-\delta)(E(y|\underline{e}) - g(\underline{e})) + \delta s_2] + (1-\alpha) s_2. \tag{25}$$

These conditions pin down α as

$$\alpha = \frac{M^*}{\delta(s_* - s(\underline{e})) + (1-\delta)M^*}.$$

Then following similar arguments as in Case A, which is omitted, one can find the unique bonus $b = w_2 - w_1$. □

Proof of Theorem 4. To formulate the problem as an optimal control problem, first define an auxiliary function $k\colon [\underline{y},\, \bar{y}] \to \Re$ by

$$k(y) = \int_{\underline{y}}^{y} -\varphi(y) f_e(y|e) dy.$$

Let φ be the sole control variable and $k(y)$, $M(y)$ be the two state variables. The constraints of the optimal control problem are given by the two laws of motion,

$$\dot{k} = -\varphi(y) f_e(y|e), \tag{26a}$$

$$\dot{M} = 0, \tag{26b}$$

a pair of constraints on surplus destruction $\varphi(y)$,

$$f(y)\varphi(y) \geq 0, \tag{26c}$$

$$f(y)(M(y) - \varphi(y)) \geq 0, \tag{26d}$$

and the boundary conditions,

$$k(\underline{y}) = 0 \;\text{ and }\; k(\bar{y}) = g'(e), \tag{26e}$$

$$M(\bar{y}) > 0. \tag{26f}$$

Note that the boundary condition $k(\bar{y}) = g'(e)$ is the agent's incentive constraint.

Let λ, μ be the co-state variables assigned to k and M respectively; let $p(y)$, $q(y)$ be the multipliers for constraints $f(y)\varphi(y) \geq 0$ and $f(y)M(y) \geq f(y)\varphi(y)$ respectively.

Then the Hamiltonian of the control problem is given by (the constant term $g(e)$ is omitted)

$$H = (y - \varphi(y))f(y) - \lambda(y)\varphi(y)f_e(y) - \frac{1-\delta}{\delta}M(y)f(y),$$

and the Lagrangian is given by

$$\begin{aligned} L &= H + pf(y)\varphi(y) + qf(y)(M - \varphi(y)) \\ &= (y - \varphi(y))f(y) - \lambda\varphi(y)f_e(y) - \frac{1-\delta}{\delta}Mf(y) + pf(y)\varphi(y) + qf(y)(M - \varphi(y)). \end{aligned}$$

By the Pontryagin Maximum Principle, the necessary and sufficient conditions for a piecewise continuous function φ and piecewise differentiable functions k and M to solve the problem include the original constraints of the control problem, (26a) – (26f), and the following additional conditions (for example, see Leonard and Long, 1992):
The first-order condition on the control variable φ,

$$\frac{\partial L}{\partial \varphi} = f(y)\left(-1-\lambda\frac{f_e(y)}{f(y)}+p(y)-q(y)\right)=0, \tag{27a}$$

the complementary slackness conditions,

$$p(y)\geq 0,\ p(y)f(y)\varphi(y)=0, \tag{27b}$$

$$q(y)\geq 0,\ q(y)f(y)(M(y)-\varphi(y))=0, \tag{27c}$$

the laws of motion of the two co-state variables λ and μ,

$$\dot{\lambda}=-\frac{\partial L}{\partial k}=0, \tag{27d}$$

$$\dot{\mu}=-\frac{\partial L}{\partial M}=f(y)\left(\frac{1-\delta}{\delta}-q(y)\right), \tag{27e}$$

and the boundary conditions,

$$\mu(\underline{y})=\mu(\bar{y})=0. \tag{27f}$$

To derive the results in the theorem, first we observe that λ is constant and satisfies $\lambda > 0$. To see why, suppose $\lambda \leq 0$. Then for any y with $\varphi(y)>0$, we have $p(y)=0$ by the complementary slackness condition (27b), and then must have $f_e(y)>0$ by (27a). But then

$$k(\bar{y})=\int_{\underline{y}}^{\bar{y}} -\varphi(y)f_e(y|e)dy<0,$$

so the agent's incentive constraint $k(\bar{y})=g'(e)$ cannot be satisfied.

Rewrite (27a) as

$$1+\lambda\frac{f_e(y)}{f(y)}=p(y)-q(y).$$

It follows that

$$1+\lambda\frac{f_e(y)}{f(y)}>0 \implies p(y)>0 \implies \varphi(y)=0$$

$$1+\lambda\frac{f_e(y)}{f(y)}<0 \implies q(y)>0 \implies \varphi(y)=M^*.$$

Given $\lambda>0$ and the fact that the likelihood ratio $\frac{f_e(y)}{f(y)}$ is increasing in y, there exists $\hat{y}$ with $f_e(\hat{y})<0$ such that

$$1+\lambda\frac{f_e(\hat{y})}{f(\hat{y})}=0 \tag{28}$$

and

$$\varphi(y)=\begin{cases} 0 & \text{if } y>\hat{y} \\ M^* & \text{if } y\leqslant\hat{y}. \end{cases}$$

It follows that the agent's incentive constraint $k(\bar{y})=g'(e)$ becomes

$$\int_{\underline{y}}^{\bar{y}}-\varphi(y)f_e(y|e)dy=-M^*F_e(\hat{y})=g'(e). \tag{29}$$

Moreover, since $q(y)=0$ for $y>\hat{y}$ and $p(y)=0$ for $y\leq\hat{y}$, by the first-order condition (27a) we have

$$-f(y)q(y)=\begin{cases} 0 & \text{if } y>\hat{y} \\ f(y)+\lambda f_e(y) & \text{if } y\leqslant\hat{y}. \end{cases} \tag{30}$$

Integrate $\dot{\mu}$ using (27e), use the boundary conditions (27f), and then apply the above condition (30) to get

$$0=\int_{\underline{y}}^{\bar{y}}\left(\frac{1-\delta}{\delta}f(y)-f(y)q(y)\right)dy=\frac{1-\delta}{\delta}+\int_{\underline{y}}^{\hat{y}}(f(y)+\lambda f_e(y))dy, \tag{31}$$

which reduces to

$$\frac{1-\delta}{\delta}=-F(\hat{y})-\lambda F_e(\hat{y}). \tag{32}$$

Using the cutoff condition (28) to substitute for λ in (32), we obtain the following condition that determines the cutoff point $\hat{y}$ implicitly as a function of δ:

$$\frac{1-\delta}{\delta}=-F(\hat{y})+\frac{f(\hat{y})}{f_e(\hat{y})}F_e(\hat{y}). \tag{33}$$

Differentiating (33) w.r.t. δ yields

$$\frac{d}{d\delta}\left(\frac{1-\delta}{\delta}\right)=\left\{-f(\hat{y})+\frac{d}{d\hat{y}}\left(\frac{f}{f_e}\right)\cdot F_e(\hat{y})+\frac{f(\hat{y})}{f_e(\hat{y})}\cdot f_e(\hat{y})\right\}\frac{d\hat{y}}{d\delta}$$
$$-\frac{1}{\delta^2}=\frac{d}{d\hat{y}}\left(\frac{f}{f_e}\right)\cdot F_e(\hat{y})\cdot\frac{d\hat{y}}{d\delta}$$

Since $\frac{d}{d\hat{y}}\left(\frac{f}{f_e}\right) < 0$ and $F_e < 0$ at $\hat{y}$, we have

$$\frac{d\hat{y}}{d\delta} < 0.$$

It then follows that

$$\frac{d\lambda}{d\delta} = -\frac{d}{d\hat{y}}\left(\frac{f}{f_e}\right) \cdot \frac{d\hat{y}}{d\delta} < 0$$

and from $M^* = -g'(e)/F_e(\hat{y})$ and $f_e(\hat{y}) < 0$ that

$$\frac{dM^*}{d\delta} = \frac{g'(e) \cdot f_e(\hat{y})}{(F_e(\hat{y}))^2} \cdot \frac{d\hat{y}}{d\delta} > 0.$$

□

References

Abreu, Dilip, "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, November 1988, *56(6)*, pp. 383–96.

Abreu, Dilip, Paul Milgrom, and David G. Pearce, "Information and Timing in Repeated Partnerships," *Econometrica*, 1991, *59(6)*, pp. 1713–33.

Abreu, Dilip and David G. Pearce, "A Perspective on Renegotiation in Repeated Games," in R. Selten, ed., *Game Equilibrium Models*, Berlin: Springer Verlag, 1991.

Abreu, Dilip and David G. Pearce, "Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts," *Econometrica*, 2007, *75(3)*, pp. 653–710.

Abreu, Dilip, David G. Pearce, and Ennio Stacchetti, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 1990, *58(5)*, pp. 1041–64.

Abreu, Dilip, David G. Pearce, and Ennio Stacchetti, "Renegotiation and Symmetry in Repeated Games," *Journal of Economic Theory*, 1993, *60(2)*, pp. 217–240.

Baker, George, Robert Gibbons, and Kevin J. Murphy, "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics*, 1994, *109*, pp. 1125–56.

Bergin, James and W. Bentley MacLeod, "Efficiency and Renegotiation in Repeated Games," *Journal of Economic Theory*, 1993, *v61, n1*, 42–73.

Bernheim, B. Douglas and Debraj Ray, "Collective Dynamic Consistency in Repeated Games," *Games and Economic Behavior*, 1989, *v1, n4*, 295–326.

Bull, Clive, "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 1987, *102(1)*, pp. 147–59.

Farrell, Joseph and Eric Maskin, "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1989, *v1, n4*, 327–60.

Fuchs, William, "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 2007, *97(4)*, pp. 1432–48.

Fudenberg, Drew, David K. Levine, and Eric S. Maskin, "The Folk Theorem with Imperfect Public Information," *Econometrica*, 1994, *62(5)*, pp. 997–1039.

Green, Edward J. and Robert H. Porter, "Non-Cooperative Collusion Under Imperfect Price Information," *Econometrica*, January 1984, *52*, pp. 87–100.

Kandori, Michihiro, "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 2002, *102(1)*, pp. 1–15.

Leonard, D. and N. V. Long, *Optimal Control Theory and Static Optimization in Economics*. Cambridge University Press, 1992.

Levin, Jonathan, "Relational Incentive Contracts," *American Economic Review*, 2003, *93(3)*, pp. 835–57.

MacLeod, W. Bentley, "Optimal Contracting with Subjective Evaluation," *American Economic Review*, 2003, *93*, pp. 216–40.

MacLeod, W. Bentley and James M. Malcomson, "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 1989, *57(2)*, pp. 447–80.

Milkovich, George T. and Jeffrey M. Newman, *Compensation*, Chicago: Irwin, 1996.

Mirrlees, James A., "Notes on Welfare Economics, Information and Uncertainty," in M. Balch, D. McFadden, and S. Wu, eds., *Essays in Equilibrium Behavior under Uncertainty*, North-Holland, 1974.

Pearce, David G., "Renegotiation-Proof Equilibria: Collective Rationality and Intertemporal Cooperation," Yale University, Cowles Foundation Discussion Papers #855, 1987.

Pearce, David G. and Ennio Stacchetti, "The Interaction of Implicit and Explicit Contracts in Repeated Agency," *Games and Economic Behavior*, 1998, *23(1)*, pp. 75–96.

Prendergast, Canice, "The Provision of Incentives in Firms," *Journal of Economic Literature*, Mar 1999, *36(1)*, pp.7–63.

Radner, Roy, "Repeated Principal-Agent Games with Discounting," *Econometrica*, 1985, *53(5)*, pp. 1173–98.

Ray, Debraj, "Internally Renegotiation-proof Equilibrium Sets: Limit Behavior with Low Discounting," *Games and Economic Behavior*, 1994, *6*, pp. 162–177.

Zhao, Rui, "Productive Low Morale," *Economics Letters*, 2009, *103(1)*.