MECS 570-1: Organizational Economics I

Michael Powell

Incentives in Organizations, Part IV: Relational Incentive Contracts

1 Relational Incentive Contracts (Updated: Jan 14 2017)

If an Agent's performance is commonly observed only by other members of his organization, or if the market is sure about his intrinsic productivity, then the career concerns motives we previously discussed cannot serve as motivation. However, individuals may form long-term attachments with an organization. In such long-term relationships, goodwill can arise as an equilibrium phenomenon, and fear of shattering this goodwill can motivate individuals to perform well and to reward good performance. This intuition is captured in models of relational contracts (informal contracts enforced by relationships). An entire section of the next course in this sequence will be devoted to studying many of the issues that arise in such models, but for now we will look at the workhorse model in the literature to get some of the more general insights.

The workhorse model is an infinitely repeated Principal-Agent game with publicly observed actions. We will characterize the "optimal relational contract" as the equilibrium of the repeated game that either maximizes the Principal's equilibrium payoffs or the Principal and Agent's joint equilibrium payoffs. A couple comments are in order at this point. First, these are applied models of repeated games and therefore tend to focus on situations where the discount factor is not close to 1, asking questions like "how much effort can be sustained in equilibrium?"

Second, such models often have many equilibria, and therefore we will be taking a stance on equilibrium selection in their analysis. The criticism that such models have no predictive power is, as Kandori puts it "... misplaced if we regard the theory of repeated games as a theory of informal contracts. Just as anything can be enforced when the party agrees to sign a binding contract, in repeated games [many outcomes can be] sustained if players agree on an equilibrium. Enforceability of a wide range of outcomes is the essential property of effective contracts, formal or informal." (Kandori, 2008, p. 7) Put slightly differently, focusing on optimal contracts when discussing formal contract design is analogous to focusing on optimal relational contracts when discussing repeated principal-agent models. Our objective, therefore, will be to derive properties of *optimal* relational contracts.

Description A risk-neutral Principal and risk-neutral Agent interact repeatedly in periods $t = 0, 1, 2, \ldots$. In period t, the Agent chooses an effort level $e_t \in E$ at cost $c(e_t) = \frac{c}{2}e_t^2$ that determines output $y_t = e_t \in Y$, which accrues to the Principal. The output can be sold on the product market for price p. At the beginning of date t, the Principal proposes a compensation package to the agent. This compensation consists of a fixed salary s_t and a contingent payment $b_t : E \to \mathbb{R}$ (with positive values denoting a transfer from the Principal to the Agent and negative values denoting a transfer from the Agent to the Principal), which can depend on the Agent's effort choice. The Agent can accept the proposal (which we denote by $d_t = 1$) or reject it (which we denote by $d_t = 0$) in favor of an outside option that yields per-period utility $\bar{u}$ for the Agent and $\bar{\pi}$ for the Principal. If the Agent accepts the proposal, the Principal is legally compelled to pay the transfer s_t, but she is not legally compelled to pay the contingent payment b_t.

Timing The stage game has the following five stages

1. P makes A a proposal (b_t, s_t).
2. A accepts or rejects in favor of outside opportunity yielding $\bar{u}$ to A and $\bar{\pi}$ to P.
3. P pays A an amount s_t.
4. A chooses effort $\hat{e}_t$ at cost $c(\hat{e}_t)$, which is commonly observed.
5. P pays A a transfer $\hat{b}_t$.

Equilibrium The Principal is not legally required to make the promised payment b_t, so in a one-shot game, she would always choose $\hat{b}_t = 0$ (or analogously, if $b_t < 0$, the Agent is not legally required to pay b_t, so he would choose $\hat{b}_t = 0$). However, since the players are engaged in a long-term relationship and can therefore condition future play on this transfer, nonzero transfers can potentially be sustained as part of an equilibrium.

Whenever we consider repeated games, we will always try to spell out explicitly the variables that players can condition their behavior on. This exercise is tedious but important. Let $h_0^t = \left\{s_0, d_0, \hat{e}_0, \hat{b}_0, \ldots, s_{t-1}, d_{t-1}, \hat{e}_{t-1}, \hat{b}_{t-1}\right\}$ denote the history up to the beginning of date t. In this game, all variables are commonly observed, so the history up to date t is a public history. We will also adopt the notation $h_s^t = h^t \cup \{s_t\}$, $h_d^t = h_s^t \cup \{d_t\}$, and $h_e^t = h_d^t \cup \{\hat{e}_t\}$, so we can cleanly keep track of within-period histories. (If we analogously defined h_b^t, it would be the same as h_0^{t+1}, so we will refrain from doing so.) Finally, let $\mathcal{H}_0^t, \mathcal{H}_s^t, \mathcal{H}_d^t$, and $\mathcal{H}_e^t$ denote, respectively, the sets of such histories.

Following Levin (2003), we define a **relational contract** to be a complete plan for the relationship. It describes (1) the salary that the Principal should offer the Agent ($h_0^t \mapsto s_t$), (2) whether the Agent should accept the offer ($h_s^t \mapsto d_t$), (3) what effort level the Agent should choose ($h_d^t \mapsto \hat{e}_t$), and (4) what bonus payment the Principal should make $\left(h_e^t \mapsto \hat{b}_t\right)$. A relational contract is **self-enforcing** if it describes a subgame-perfect equilibrium of the repeated game. An **optimal relational contract** is a self-enforcing relational contract that yields higher equilibrium payoffs for the Principal than any other self-enforcing relational contract. It is important to note that a relational contract describes behavior on and off the equilibrium path.

Comment. *Early papers in the relational-contracting literature (Bull, 1987; MacLeod and Malcomson, 1989; Baker, Gibbons, and Murphy, 1994) referred to the equilibrium of the game instead as an implicit (as opposed to relational) contract. More recent papers eschew the term implicit, because the term "implicit contracts" has a connotation that seems to emphasize whether agreements are common knowledge, whereas the term "relational contracts" more* *clearly focuses on whether agreements are enforced formally or must be self-enforcing.*

The Program Though the stage game is relatively simple, and the game has a straightforward repeated structure, solving for the optimal relational contract should in principle seem like a daunting task. There are tons of things that the Principal and Agent can do in this game (the strategy space is quite rich), many of which are consistent with equilibrium play—there are lots of equilibria, some of which may have complicated dynamics. Our objective is to pick out, among all these equilibria, those that maximize the Principal's equilibrium payoffs.

Thankfully, there are several nice results (many of which are contained in Levin (2003) but have origins in the preceding literature) that make this task achievable. We will proceed in the following steps:

1. We will argue, along the lines of Abreu (1988), that the unique stage game SPNE is an optimal punishment.

2. We will show that optimal reward schedules are "forcing." That is, they pay the Agent a certain amount if he chooses a particular effort level, and they revert to punishment otherwise. An optimal relational contract will involve an optimal reward scheme.

3. We will then show that distribution and efficiency can be separated out in the stage game. Ex ante transfers have to satisfy participation constraints, but they otherwise do not affect incentives or whether continuation payoffs are self-enforcing.

4. We will show that an optimal relational contract is sequentially optimal on the equilibrium path. Increasing future surplus is good for ex ante surplus, which can be divided in any way, according to (3), and it improves the scope for incentives in the current period. Total future surplus is always maximized in an optimal relational contract, and since the game is a repeated game, this implies that total future surplus is therefore constant in an optimal relational contract.

5. We will then argue that we can restrict attention to stationary relational contracts. By (4), the total future surplus is constant in every period. Contemporaneous payments and the split of continuation payoffs are perfect substitutes for motivating effort provision and bonus payments and for participation. Therefore, we can restrict attention to agreements that "settle up" contemporaneously rather than reward and punish with continuation payoffs.

6. We will then solve for the set of stationary relational contracts, which is not so complicated. This set will contain an optimal relational contract.

In my view, while the restriction to stationary relational contracts is helpful for being able to tractably characterize optimal relational contracts, the important economic insights are actually that the relational contract is sequentially optimal and how this result depends on the separation of distribution and efficiency. The separation of distribution and efficiency in turn depends on several assumptions: risk-neutrality, unrestricted and costless transfers, and a simple information structure. In the next course in this sequence, we will return to these issues and think about settings where one or more of these assumptions is not satisfied.

Step 1 is straightforward. In the unique SPNE of the stage game, the Principal never pays a positive bonus, the Agent exerts zero effort, and he rejects any offer the Principal makes. The associated payoffs are $\bar{u}$ for the Agent and $\bar{\pi}$ for the Principal. It is also straightforward to show that these are also the Agent's and Principal's maxmin payoffs, and therefore they constitute an optimal penal code (Abreu, 1988). Define $\bar{s} = \bar{u} + \bar{\pi}$ to be the outside surplus.

Next, consider a relational contract that specifies, in the initial period, payments w and $b(\hat{e})$, an effort level e, and continuation payoffs $u(\hat{e})$ and $\pi(\hat{e})$. The equilibrium payoffs of this relational contract, if accepted are:

$$\begin{aligned} u &= (1-\delta)(w - c(e) + b(e)) + \delta u(e) \\ \pi &= (1-\delta)(p \cdot e - w - b(e)) + \delta \pi(e). \end{aligned}$$

Let $s = u + \pi$ be the equilibrium contract surplus. This relational contract is self-enforcing if the following four conditions are satisfied.

1. Participation:
$$u \geq \bar{u}, \pi \geq \bar{\pi}$$

2. Effort-IC:
$$e \in \underset{\hat{e}}{\operatorname{argmax}} \left\{ (1-\delta)\left(-c(\hat{e}) + b(\hat{e})\right) + \delta u(\hat{e}) \right\}$$

3. Payment:
$$\begin{aligned} (1-\delta)\left(-b(e)\right) + \delta \pi(e) &\geq \delta \bar{\pi} \\ (1-\delta)\, b(e) + \delta u(e) &\geq \delta \bar{u} \end{aligned}$$

4. Self-enforcing continuation contract: $u(e)$ and $\pi(e)$ correspond to a self-enforcing relational contract that will be initiated in the next period.

Step 2: Define the Agent's **reward schedule** under this relational contract by

$$R(\hat{e}) = b(\hat{e}) + \frac{\delta}{1-\delta} u(\hat{e}).$$

The Agent's no-reneging constraint implies that $R(\hat{e}) \geq \frac{\delta}{1-\delta}\bar{u}$ for all $\hat{e}$. Given a proposed effort level e, suppose there is some other effort level $\hat{e}$ such that $R(\hat{e}) > \frac{\delta}{1-\delta}\bar{u}$. Then we can define an alternative relational contract in which everything else is the same, but $\tilde{R}(\hat{e}) = R(\hat{e}) - \varepsilon$ for some $\varepsilon > 0$. The payment constraints remain satisfied, and the effort-IC constraint becomes easier to satisfy. Therefore, such a change makes it possible to weakly improve at least one player's equilibrium payoff. Therefore, it has to be that $R(\hat{e}) = \frac{\delta}{1-\delta}\bar{u}$ for all $\hat{e} \neq e$.

Step 3: Consider an alternative relational contract in which everything else is the same, but $\tilde{w} = w - \varepsilon$ for some $\varepsilon \neq 0$. This changes the equilibrium payoffs u, π to $\tilde{u}, \tilde{\pi}$ but not the joint surplus s. Further, it does not affect the effort-IC, the payment, or the self-enforcing continuation contract conditions. As long as $\tilde{u} \geq \bar{u}$ and $\tilde{\pi} \geq \bar{\pi}$, then the proposed relational contract is still self-enforcing.

Define the value s^* to be the maximum total surplus generated by any self-enforcing relational contract. The set of possible payoffs under a self-enforcing relational contract is then $\{(u, \pi) : u \geq \bar{u}, \pi \geq \bar{\pi}, u + \pi \leq s^*\}$. For a given relational contract to satisfy the self-enforcing continuation contract condition, it then has to be the case that for any equilibrium effort e,

$$(u(e), \pi(e)) \in \{(u, \pi) : u \geq \bar{u}, \pi \geq \bar{\pi}, u + \pi \leq s^*\}.$$

Step 4: Suppose the continuation relational contract satisfies $u(e) + \pi(e) < s^*$. Then $\pi(e)$ can be increased in a self-enforcing relational contract, holding everything else the same. Increasing $\pi(e)$ does not affect the effort-IC constraint, it relaxes both the Principal's participation and payment constraints, and it increases equilibrium surplus. The original relational contract is then not optimal. Therefore, any optimal relational contract has to satisfy $s(e) = u(e) + \pi(e) = s^*$.

Step 5: Suppose the proposed relational contract is optimal and generates surplus $s(e)$. By the previous step, it has to be the case that $s(e) = e - c(e) = s^*$. This in turn implies that optimal relational contracts involve the same effort choice, e^*, in each period. Now we want to construct an optimal relational contract that provides the same incentives for the agent to exert effort, for both players to pay promised bonus payments, and also yields continuation payoffs that are equal to equilibrium payoffs (i.e., not only is the action that is chosen the same in each period, but so are equilibrium payoffs). To do so, suppose an optimal relational contract involves reward scheme $R(\hat{e}) = \frac{\delta}{1-\delta}\bar{u}$ for $\hat{e} \neq e^*$ and

$$R(e^*) = b(e^*) + \frac{\delta}{1-\delta} u(e^*).$$

Now, consider an alternative reward scheme $\tilde{R}(e^*)$ that provides the same incentives to the agent but leaves him with a continuation payoff of u^*:

$$\tilde{R}(e^*) = \tilde{b}(e^*) + \frac{\delta}{1-\delta} u^* = R(e^*).$$

This reward scheme also leaves him with an equilibrium utility of u^*

$$\begin{aligned} u^* &= (1-\delta)(w - c(e^*) + b(e^*)) + \delta u(e^*) = (1-\delta)(w - c(e^*) + R(e^*)) \\ &= (1-\delta)\left(w - c(e^*) + \tilde{R}(e^*)\right) = (1-\delta)\left(w - c(e^*) + \tilde{b}(e^*)\right) + \delta u^*. \end{aligned}$$

Since $\bar{u} \leq u^* \leq s^* - \bar{\pi}$, this alternative relational contract also satisfies the participation constraints.

Further, this alternative relational contract also satisfies all payment constraints, since by construction,

$$\tilde{b}(e^*) + \frac{\delta}{1-\delta} u^* = b(e^*) + \frac{\delta}{1-\delta} u(e^*),$$

and this equality also implies the analogous equality for the Principal (since $s^* = u^* + \pi^*$ and $s^* = u(e^*) + \pi(e^*)$):

$$-\tilde{b}(e^*) + \frac{\delta}{1-\delta} \pi^* = -b(e^*) + \frac{\delta}{1-\delta} (\pi(e^*)).$$

Finally, the continuation payoffs are (u^*, π^*), which can themselves be part of this exact same self-enforcing relational contract initiated the following period.

Step 6: The last step allows us to set up a program that we can solve to find an optimal relational contract. A stationary effort level e generates total surplus $s = e - c(e)$. The Agent is willing to choose effort level e if he expects to be paid a bonus b satisfying

$$b + \frac{\delta}{1-\delta} (u - \bar{u}) \geq c(e).$$

That is, he will choose e as long as his effort costs are less than the bonus b and the change in his continuation payoff that he would experience if he did not choose effort level e. Similarly, the Principal is willing to pay a bonus b if

$$\frac{\delta}{1-\delta}\left(\pi-\bar{\pi}\right)\geq b.$$

A necessary condition for both of these inequalities to be satisfied is that

$$\frac{\delta}{1-\delta}\left(s-\bar{s}\right)\geq c\left(e\right).$$

This condition is also sufficient for an effort level e to be sustainable in a stationary relational contract, since if it is satisfied, there is a b such that the preceding two inequalities are satisfied. This pooled inequality is referred to as the **dynamic-enforcement constraint.**

The Program: Putting all this together, then, an optimal relational contract will involve an effort level that solves

$$\max_{e} pe-\frac{c}{2}e^{2}$$

subject to the dynamic-enforcement constraint:

$$\frac{\delta}{1-\delta}\left(pe-\frac{c}{2}e^{2}-\bar{s}\right)\geq\frac{c}{2}e^{2}.$$

The first-best effort level $e^{FB}=\frac{p}{c}$ solves this problem as long as

$$\frac{\delta}{1-\delta}\left(pe^{FB}-\frac{c}{2}\left(e^{FB}\right)^{2}-\bar{s}\right)\geq\frac{c}{2}\left(e^{FB}\right)^{2},$$

or

$$\delta\geq\frac{p^{2}}{2p^{2}-2c\bar{s}}.$$

Otherwise, the optimal effort level e^{*} is the larger solution to the dynamic-enforcement constraint, when it holds with equality:

$$e^* = \frac{p}{c}\left(\delta + \sqrt{\frac{p^2\delta^2 - 2\delta\bar{s}c}{p^2}}\right).$$

For all $\delta < \frac{p^2}{2p^2 - 2c\bar{s}}$, $\delta + \sqrt{\frac{p^2\delta^2 - 2\delta\bar{s}c}{p^2}} < 1$, so $e^* < e^{FB}$.

Comment. *People not familiar or comfortable with these models often try to come up with ways to artificially generate commitment. For example, they might propose something along the lines of, "If the problem is that the Principal doesn't have the incentives to pay a large bonus when required to, why doesn't the Principal leave a pot of money with a third-party enforcer that she will lose if she doesn't pay the bonus?" This proposal seems somewhat compelling, except for the fact that it would only solve the problem if the third-party enforcer could withhold that pot of money from the Principal if and only if the Principal breaks her promise to the Agent. Of course, this would require that the third-party enforcer condition its behavior on whether the Principal and the Agent cooperate. If the third-party enforcer could do this, then the third-party enforcer could presumably also enforce a contract that conditions on these events as well, which would imply that cooperation is contractible. On the other hand, if the third-party enforcer cannot conditionally withhold the money from the Principal, then the Principal's reneging temptation will consist of the joint temptation to* (a) *not pay the bonus she promised the agent and* (b) *recover the pot of money from the third-party enforcer.*

Further Reading The analysis in this section specializes Levin's (2003) analysis to a setting of perfect public monitoring and no private information about the marginal returns to effort. Levin (2003) shows that in a fairly general class of repeated environments with imperfect public monitoring, if an optimal relational contract exists, there is a stationary relational contract that is optimal. Further, the players' inability to commit to payments enters the program only through a dynamic enforcement constraint. Using these results, he is able to show how players' inability to commit to payments shapes optimal incentive contracts in moral-hazard settings and settings in which the agent has private information about his marginal returns to effort.

MacLeod and Malcomson (1998) show that the structure of payments in an optimal relational contract can take the form of contingent bonuses or efficiency wages. Baker, Gibbons, and Murphy (1994) show that formal contracts can complement relational contracts, but they can also crowd out relational contracts. We will explore a number of further issues related to relational-incentive contracts later in the course.

The motivation I gave above begins with the premise that formal contracts are simply not enforceable and asks what *equilibrium* arrangement is best for the parties involved. Another strand of the relational-contracting literature begins with the less-stark premise that formal contracts are costly (but not infinitely so) to write, and informal agreements are less costly (but again, are limited because they must be self-enforcing). Under this view, relational contracts are valuable, because they give parties the ability to adapt to changing circumstances without having to specify in advance just how they will adapt (Macaulay, 1963). Baker, Gibbons, and Murphy (2011) and Barron, Gibbons, Gil, and Murphy (2015) explore implications of relational *adaptation*, and the former paper also considers the question of when adaptation should be governed by a formal contract and when it should be governed through informal agreements.