# MULTILATERAL CONTRACTING AND THE EMPLOYMENT RELATIONSHIP*

JONATHAN LEVIN

This paper studies the structure of the employment relationship in organizations. It investigates the trade-off firms face between making commitments to their workforce as a whole (multilateral relational contracts), and making more limited commitments to individuals or smaller groups of employees (bilateral relational contracts). Multilateral contracts bind the firm more strongly to implicit commitments, improving motivation, but are difficult to adjust in response to changes in the environment. Bilateral contracts make workforce changes easier to implement. The framework helps to explain the use of relative performance evaluation, why firms rely on temporary employees, and the adoption of two-tier workforces.

## I. INTRODUCTION

Why do firms hesitate to cut pay or lay off workers in economic downturns? Why do some firms promote the idea of lifelong employee commitment, while others explicitly back away from such policies or hire temporary workers with low tenure expectations? In this paper I present a simple model of employment as a long-term relationship between a firm and its workers and show that these phenomena, and others, can be understood as attempts by firms to optimally structure relational contracts with employees.[1] In particular, I consider the trade-off between emphasizing a broad commitment to all employees and making more targeted commitments to individual employees or groups of employees. The former, modeled as a multilateral relational contract, may increase motivation, but can limit flexibility and make changes more difficult than having separate bilateral contracts. I identify conditions that favor one arrangement over the other and discuss the implications of both for issues such as incentive provision and compensation.

The approach I take stresses the contractual relationship between firms and their workforce as a whole, rather than the

* This paper is a revised chapter of my Ph.D. thesis [Levin 1999]. I am indebted to Glenn Ellison, Jerry Hausman, and Bengt Holmström for their guidance and support. I have also benefited from the suggestions of Robert Gibbons, Richard Levin, Paul Oyer, Antonio Rangel, Korok Ray, Luis Rayo, and Steven Tadelis, and from the thoughtful comments of Edward Glaeser and a referee.

1. For papers that model various aspects of informal contracting between firms and workers, see, e.g., Shapiro and Stiglitz [1984], Bull [1987], MacLeod and Malcomson [1989, 1998], Kreps [1990], and Baker, Gibbons, and Murphy [2002].



*The Quarterly Journal of Economics,* August 2002

relationship between firms and individual employees. This is crucial for addressing many incentive issues. For instance, there is extensive evidence that managers believe selective pay cuts or layoffs will trigger widespread morale problems and lowered performance [Doeringer and Piore 1971; Bewley 1999]. Bewley argues that this is very difficult to reconcile with existing economic theories of employment. There are some striking examples of firms facing difficulties when they had to back away from what employees perceived as a blanket policy of employment security for those who performed well. Both Hewlett-Packard and IBM were cited for decades as having strong commitments to employment security; and both faced widespread employee discontent when market conditions forced them to lay off workers and change other aspects of their employment policies in the late 1980s and early 1990s [Mills and Friesen 1996; Rogers and Beer 1995]. Of course, they may be exceptional in having been forced to alter their policies so dramatically; there are many examples of firms that have upheld implicit promises of job security in the face of adverse circumstances by going to four-day weeks or flexible hours, or relocating or retraining workers.[2]

One way firms can avoid or mitigate the consequences of layoffs or other organizational changes is to set separate expectations for different employees or sets of employees. A stark example of this might be the use of temporary labor, as in the case of Microsoft. In the 1980s and 1990s Microsoft employed thousands of "permatemp" workers, many of whom enjoyed long tenures and worked on the same projects as regular workers. Yet permatemps were issued orange badges rather than the blue badges of regular workers, were not invited to the annual company picnic, and were prohibited from joining company social clubs or using company basketball courts. The president of a company that supplied some of Microsoft's permatemps explained to *The New York Times* that these practices served to align expectations: "when people know it's a temporary arrangement, someday when the assignment ends, there's not a sense of broken

2. A recent example, drawn more or less at random from the newspaper, is Charles Schwab's decision to ask half its employees to take a series of four-day weeks rather than contemplate layoffs ("Schwab Tells Some Workers to Stay Home," by Patrick McGeehan, *The New York Times,* January 31, 2001, Section C, Page 1).

trust."[3] Interestingly, following their respective bouts with layoffs in the late 1980s, both IBM and Hewlett-Packard also moved to increase their reliance on temporary workers.

In this paper I model employment as a repeated game between the firm and its employees, and view the process of setting worker expectations as the choice between different forms of relational contract. A relational contract describes prospective behavior in the repeated game and is self-enforcing if the behavior it describes constitutes an equilibrium. Under a multilateral contract, the firm maintains its promises because it fears that any deviation will lead to a breakdown in its relationship with all of its workers. In contrast, bilateral contracts have the feature that the firm may violate a promise to a small set of workers, but retain the goodwill of others.[4]

The benefit of multilateral contracting is familiar from the industrial organization literature on multimarket contact: it increases the sanctions following a deviation [Bernheim and Whinston 1990; Bendor and Mookherjee 1990]. In the employment context, this gives the firm greater credibility in promising to reward performance by its employees. Formally, rather than having a separate incentive constraint for each worker, a multilateral contract needs only to satisfy the sum of individual constraints. This allows a form of cross subsidization to support performance incentives that might not otherwise be self-enforcing. I show that this cross subsidization leads firms to balance incentives across the organization, and that it favors relative performance evaluation even if there is no correlation between the performance measures of different employees. I also show that firms can benefit most from it when they enjoy strong bargaining power vis-à-vis their employees and when employees' talents are substitutable rather than complementary in the production process.

The cost of being bound tightly to a set of implicit promises is that firms may find it difficult to extricate themselves from these commitments if the environment should change. In Section IV, I

3. See "Microsoft Leads the Way in Filling Jobs with 'Permatemps,'" by Stephen Greenhouse, *The New York Times,* March 30, 1998, Section D, Page 1.

4. The terminology follows Greif [1994], who contrasts the experience of medieval Genoese merchants, who relied on bilateral relationships to enforce contractual obligations, with that of the Maghribi traders, who relied on community-wide multilateral sanctions. Greif, Milgrom, and Weingast [1994] also consider forms of multilateral enforcement in their analysis of medieval trade. Bendor and Mookerjee [1990] study bilateral and multilateral enforcement in an $n$-player prisoner's dilemma.

extend the basic model to allow for changes in the productive environment and uncertainty regarding future changes. In response to shocks, the firm must adjust or rework its relational contracts, a process that requires communicating to all involved parties. I consider the implications if this process is imperfect, so there is always a danger of a misunderstanding that will lead to a breakdown of the relationship. This creates an inertia in contract structure: contracts will not adjust to small shocks. Moreover, multilateral contracts will tend to have higher costs of adjustment. A consequence is that if the environment is unstable, and the firm believes it will be difficult to widely and accurately communicate its motives to workers or suppliers, it can be optimal to explicitly separate workers or suppliers into tiers.[5] This seems to match informal explanations for a variety of observed contracting practices.

The next section presents the model and derives the incentive constraints that must be satisfied by bilateral and multilateral contracts. Section III considers stationary environments and outlines a few implications of optimal contracting. Section IV considers the trade-off between bilateral and multilateral contracts in a changing environment. The final section concludes. Proofs of the results are in the Appendix.

## II. The Model

### II.A. Technology and Preferences

I consider a simple model of employment as an ongoing relationship. There is a single firm and $n$ workers, who can produce at dates $t = 1, 2 \ldots$. At each date, the firm chooses whether or not to employ each worker, and simultaneously workers choose whether or not they wish to be employed. Each employed worker $i$ then selects a performance level $e_t^i \in [0,\bar{e}]$, incurring a private cost $c^i(e_t^i)$. Costs are increasing and convex in performance, with $c^i(0) = 0$. Worker performance generates gross profits $y_t = y(\mathbf{e}_t)$ for the firm, where $\mathbf{e}_t \equiv (e_t^1, \ldots, e_t^n)$. Profits are increasing and concave in the performance of each

5. A somewhat related idea can be found in Fudenberg and Kreps [1987], who study a reputational model of entry, where a single large firm attempts to simultaneously deter entry in several markets. They show that the large firm may sometimes benefit if its competitors can only observe its actions in their particular market, rather than observing all its actions.

worker, $e^i$. In principle, worker inputs may be arbitrarily related, but much of the analysis will focus on the case where workers are *technologically independent,* so $y(\mathbf{e}) = y^1(e^1) + \ldots + y^n(e^n)$. Finally, I adopt the convention that if worker $i$ is not employed at time $t$, then $e_t^i \equiv \emptyset$, and let $y(\emptyset, e^{-i}) = y(0, e^{-i})$.[6]

The firm must resolve a basic incentive problem. How to induce its workers to perform, given that they must incur a private sacrifice? I take the view that ongoing interaction is crucial to resolving this problem. To this end, suppose that while the firm can monitor worker behavior, it cannot condition a formal incentive contract on performance. Rather, it may commit to pay a fixed wage $w_t^i \geq 0$ to worker $i$ if he is employed at $t$ and promise the worker a *discretionary* reward $b_t^i \geq 0$ conditional on performance. I treat both forms of compensation as cash payments, although they could naturally include nonpecuniary benefits. Then, worker $i$'s total payment at $t$, supposing that the firm delivers on the discretionary payment, is $W_t^i = w_t^i + b_t^i$.

All parties are risk-neutral. Given this, the firm's net profits at time $t$ are $y_t - \sum_i W_t^i$, while worker $i$'s payoff is $W_t^i - c^i(e_t^i)$ in the event he is employed and $\bar{u}^i \geq 0$ if he is not. All parties share a common discount factor $\delta$, meant to capture both pure time preference and any exogenous uncertainty about whether production opportunities will continue (in addition to the rate at which performance is assessed).[7] The joint surplus in a given period is (dropping the $t$ subscript)

$$s(\mathbf{e}) = y(\mathbf{e}) - \sum_{\{i:e^i \neq \emptyset\}} c^i(e^i) + \sum_{\{i:e^i = \emptyset\}} \bar{u}^i.$$

If production is separable, it is useful to isolate the contribution of each worker, writing $s(\mathbf{e}) = \sum_i s^i(e^i)$, where $s^i(e^i) = y^i(e^i) - c^i(e^i)$ if $i$ is employed and $s^i(\emptyset) = y^i(\emptyset) + \bar{u}^i$ if not.

In this environment the static equilibrium is straightforward. Because the firm cannot commit to reward performance, workers will do no more than the minimum, and the firm will do best not to produce at all. Ongoing interaction allows for a range of equilibria in which the firm honors its commitments and work-

6. This can be viewed as shorthand for the firm contracting out the work or hiring a replacement who performs poorly. The assumption may seem restrictive, but the analysis is virtually unchanged if the firm has an alternative that is better or worse than minimal performance; i.e., if $y(\emptyset, e^{-i}) \neq y(0, e^{-i})$.

7. The performance of all workers is assessed simultaneously. Asymmetries, such as if two workers performed in alternating periods, would create an additional motive for multilateral contracting.

ers are motivated. A *relational contract* describes behavior over time in the repeated game. Such a contract specifies performance $\mathbf{e} = (e^1, \ldots, e^n)$, wages $\mathbf{w} = (w^1, \ldots, w^n)$ and discretionary rewards $\mathbf{b} = (b^1, \ldots, b^n)$ to be used in each period, as well as behavior following a deviation.[8]

A relational contract is self-enforcing if it describes an equilibrium of the repeated game. There may be many possible self-enforcing contracts; I focus on the choice between two alternative regimes. Under a multilateral relational contract, any deviation by the firm leads all workers to revert to the static equilibrium. In essence, workers view any deviation by the firm, even one that does not harm them directly, as reason to withdraw their goodwill. I contrast this with the bilateral contracting regime, where workers view their relationship with the firm as distinct from the relationships of other workers. Under a set of bilateral relational contracts, workers revert to the static equilibrium only if the firm deviates from the initial agreement in a way that harms them directly.

### *II.B. Multilateral Relational Contracts*

Bewley [1999] and other management researchers argue that managers hesitate to cut pay or lay off employees because they fear that the result will be a widespread loss of morale among remaining employees. A natural explanation for this is that employees interpret such actions as evidence that a firm may not honor informal commitments in the future. Thus, even selective pay cuts or layoffs are perceived as violating a broader relational contract. This corresponds to an employment equilibrium in which any deviation from anticipated behavior by the firm triggers a general drop in performance—a multilateral relational contract.

Formally, a multilateral relational contract consists of a profile $(\mathbf{e},\mathbf{w},\mathbf{b})$ together with the expectation that if the firm deviates, all workers will cease to perform and the firm's (best-)response will be to stop employing them. Individual nonperformance leads the firm to withhold the shirking employee's bonus and terminate his employment.

8. I restrict attention to *stationary* contracts where behavior does not change on the equilibrium path of play. There is, however, no gain to nonstationary contracts [MacLeod and Malcomson 1989; Levin 2000].

PROPOSITION 1. There exists a self-enforced multilateral relational contract with performance $\mathbf{e}$ if and only if

$$\frac{\delta}{1-\delta}(s(\mathbf{e})-s(\varnothing)) \geq \sum_i c^i(e^i). \tag{1}$$

Self-enforcement requires that the future discounted surplus generated by production be greater than the present costs of all employees. Importantly, and this is the benefit of multilateral contracting, the condition is an aggregate one. There is no restriction on how the surplus generated by individual employees relates to their individual costs.

Implicit in such an arrangement is the idea that employees interpret any deviant behavior by the firm as threatening their own prospects, and also, of course, that they can observe these deviations (here, the assumption is that workers can observe *all* performance choices and payments). Both these points help shed light on human resource practices. For instance, encouraging informal communication or social interaction between employees, or emphasizing a broad commitment to lifetime employment (as firms such as IBM and Hewlett-Packard have historically done) is consistent with developing a multilateral relational contract.[9] Policies that discourage employee interaction, such as Microsoft's decision to prohibit temporary employees from recreational activities, are clearly antithetical to such a contract.

## *II.C. Bilateral Relational Contracts*

An alternative to multilateral contracting is to structure employment as a set of distinct commitments to individual workers or subsets of workers. For instance, law firms or consultancies may encourage junior associates to focus on their own careers and not worry too much about promises made to their peers. More generally, firms often have quite separate relationships with their blue- and white-collar workers, or with their full- and part-time employees. A characteristic of such situations is that if the firm behaves poorly toward one employee or set of employees, the goodwill of others need not be lost.

If relationships are technologically independent, this can be

9. Unions might also facilitate a coordinated response by workers if the firm behaves opportunistically. I will argue below, however, that employee bargaining power (which one might associate with unionization) can undermine multilateral contracting.

captured by treating the firm's relationship with each worker as a separate game, with a deviation by either side punished by termination of employment. An easy result is that performance $(e^1, \ldots, e^I)$ can be supported if and only if for all $i$

$$\frac{\delta}{1-\delta}\left(s^i(e^i) - s^i(\emptyset)\right) \geq c^i(e^i). \tag{2}$$

Production externalities add a few subtleties. First, with externalities, the value of a given worker depends on the performance of others. Second, if worker $i$ is terminated, the firm may want to alter its contracts with its remaining employees. And its ability to do so affects its initial incentives to deviate. With this in mind, I define bilateral contracts recursively. A profile $(\mathbf{e},\mathbf{w},\mathbf{b})$ is specified. Following a deviation involving a given set of employees, the affected employees are terminated. The firm then constructs a new (optimal) set of bilateral contracts with the remaining employed workers, possibly using lump-sum payments to redistribute surplus. Here, optimality of the contract means that it is surplus maximizing (within the set of self-enforced contracts). A set of bilateral contracts is self-enforcing if no employee can profitably deviate, and if the firm cannot profitably deviate against a set of employees and recontract optimally with the others.[10]

Define $s^I$ as the surplus from an optimal bilateral contract where the firm employs some set $I$ of workers.

PROPOSITION 2. There exists a set of bilateral relational contracts that support performance $\mathbf{e}$ if and only if for any subset $I$ of employed workers

$$\frac{\delta}{1-\delta}\left(s(\mathbf{e}) - s^I\right) \geq \sum_{j \notin I} c^j(e^j). \tag{3}$$

The condition states that the *marginal* surplus generated by each subset of employees must be greater than their present

10. Self-enforcement requires subgame perfection (no one acting alone can profitably deviate) and a form of coalition-proofness (the firm cannot profitably deviate in concert with a subset of employees). Note that two issues are resolved in the definition. First, contractual revisions must themselves be self-enforcing. Second, the firm may only recontract with other employed workers. An alternative approach would be to allow it to recontract with unemployed workers as well. While this would not matter under technological independence, it would make enforcement more difficult if workers were substitutable.

costs. A detailed characterization requires a series of steps. Starting with the fact that $s^{\emptyset} = s(\emptyset)$, one finds the optimal contract with a single employee $i$, and hence $s^{\{i\}}$. In turn, the optimal contract with any two employees is derived—checking deviations against either one or both—and so on. Absent externalities, this sequential process is unnecessary. The set of self-enforcing bilateral contracts is described by (2).

## III. Optimal Relational Contracts

The benefit of establishing a multilateral contract is that the firm ensures that any deviant behavior it undertakes will have severe consequences. By committing itself more strongly, it can offer a broader range of incentives. This is seen formally by noting that (1) is just one of the (many) constraints that must hold under bilateral contracting. The next section shows that this commitment may come at a cost when the environment is unstable, if it makes necessary contractual adjustments costly. Before enriching the model in this direction, however, I first take up a few of the key implications of bilateral and multilateral contracting in the stationary environment developed above. I show that the two arrangements have divergent implications for the balance of incentives across the firm and for the structure of compensation. I also observe that two factors, the firm's bargaining power vis-à-vis its employees and the nature of production externalities, greatly influence the potential gains from multilateral contracting.

An important point is that these distinctions depend on there being a meaningful tension between present incentives not to perform and future gains from contracting. If the future is negligible ($\delta \approx 0$), no contract of any kind may be self-enforcing. On the other hand, when the future is infinitely important ($\delta \approx 1$), both bilateral and multilateral contracts (and many others) will allow first-best performance. I concentrate here on second-best situations—those where employment is possible, but incentive constraints matter.

### *III.A. Balancing Incentives*

As in the multimarket contact literature, multilateral contracting allows the firm to use some relationships to "cross-subsidize" others. In the employment context, optimal contracts involve a particularly transparent form of cross subsidization: the firm uses a multilateral contract to balance the shadow costs of incentive provision across the organization.

Assuming technological independence (a similar analysis applies with externalities), an optimal multilateral contract solves

$$\max_{\mathbf{e}} \sum_i s^i(e^i) \quad \text{subject to } \frac{\delta}{1-\delta} \sum_i (s^i(e) - s^i(\varnothing)) \geq \sum_i c^i(e^i).$$

The Kuhn-Tucker optimality conditions are[11]

$$\frac{dy^i}{de^i}(e^{i,M}) = \frac{1-\delta+\lambda}{1-\delta+\lambda\delta} \frac{dc^i}{de^i}(e^{i,M}),$$

where $e^{i,M}$ is the optimal performance for worker $i$ under the multilateral contract, and $\lambda \geq 0$ is the Lagrange multiplier on the incentive constraint.

Optimal bilateral contracts solve

$$\max_{\mathbf{e}} \sum_i s^i(e^i) \quad \text{subject to } \frac{\delta}{1-\delta} (s^i(e^i) - s^i(\varnothing)) \geq c^i(e^i) \quad \text{for all } i.$$

The corresponding Kuhn-Tucker conditions are

$$\frac{dy^i}{de^i}(e^{i,B}) = \frac{1-\delta+\lambda^i}{1-\delta+\lambda^i\delta} \frac{dc^i}{de^i}(e^{i,B}),$$

where $e^{i,B}$ is the optimal performance for $i$ under a bilateral contract, and $\lambda^i \geq 0$ is now a distinct multiplier that attaches to $i$'s constraint.

Applying the natural economic interpretation of the Lagrange multipliers, the optimal multilateral contract *equalizes* the shadow price of incentive provision across relationships. Since at both optima the shadow price on incentives for a given employee is proportional to the ratio of marginal benefits and costs of that employee's performance, a multilateral contract essentially has the effect of using relationships where this benefit-cost ratio is low to cross-subsidize those where it is high.[12]

11. I assume throughout this section that it is always possible and profitable to employ all workers. This is easily relaxed, but at some cost in terms of expositional clarity.

12. Bernheim and Whinston [1990] obtain an equalization result in their analysis of symmetric collusive equilibria with differentiated products (Section VII). In their case, optimal collusion equalizes the ratio of marginal profits from collusion to marginal profits from deviation across markets.

PROPOSITION 3. Multilateral contracting generates more surplus than bilateral contracting by equalizing the ratio of marginal benefits and marginal costs of performance across relationships.

The equalization of incentives suggests that multilateral contracts might give rise to a form of pay compression, such as is observed in many large organizations. However, while incentive equalization might translate into wage equalization under some conditions, it need not in general. Indeed, it need not even imply that worker's net marginal products ($dy^i/de^i - dc^i/de^i$) be equalized at the optimum.

### III.B. Structure of Compensation: Tournaments

A pervasive feature of many organizations is that compensation has elements of a tournament. Even apart from the explicit use of relative performance evaluation, promotions, the assignment of desirable tasks, and individual awards (e.g., "employee of the month") induce competition among employees. Incentive theory has traditionally argued that the role of relative performance evaluation is to provide insurance for workers whose performance measures are positively correlated.[13] Interestingly, relative performance evaluation turns out to be broadly optimal when employment is structured as a multilateral contract, if performance is measured imperfectly.[14] This result does not depend on risk aversion or on correlation in performance measurement. Rather, it results from the multilateral nature of the contract; with bilateral contracting, relative performance evaluation is not useful.

Suppose that each agent's performance $e^i$ is not observed directly, but generates an observable measure $x^i$ with continuous density $f^i(x^i|e^i)$. These stochastic measures are independent, and satisfy the Mirrlees-Rogerson conditions from incentive theory, so that each agent's performance choice can be characterized by a first-order condition. A relational contract specifies performance levels $e^i$, fixed payments $w^i$, and discretionary payments

13. Malcomson [1984] argues that rank-order tournaments are also appealing relative to piece rates because they do not require performance to be contracted on directly and because they give the employer no incentive to misrepresent outcomes.

14. An interesting contrast to this result is provided by Che and Yoo [2001]. They show that if the firm sets a static incentive scheme, and workers then attempt to cooperate over time to maximize their joint payoffs given this incentive scheme, then *joint* performance evaluation (positively correlating payments) can be optimal.

$b^i(x^1, \ldots, x^n)$ that may be contingent on all the performance measures.

PROPOSITION 4. Suppose that performance is measured imperfectly and there is technological independence. With bilateral contracts, each employee's compensation is independent and based on his own performance measure. With a multilateral contract, relative performance evaluation is optimal: specifically, the optimal incentive scheme is a modified tournament that awards a fixed prize to at most one employee.

The argument for the result proceeds as follows. Self-enforcement limits the level of discretionary pay the firm can credibly promise. With bilateral contracts, this places an independent restriction on each worker's compensation, but with a multilateral contract, self-enforcement limits only the *total* payment. So the firm essentially has a fixed pool of money (of endogenous size) to reward performance. A standard argument shows that promising an extra dollar to employee $i$ in the event of outcome $(x^1, \ldots, x^n)$ has incentive value $\mu^i(f_e^i/f^i)(x^i|e^i)$, where $\mu^i$ is the (positive) shadow price on $i$'s first-order condition for performance choice. Thus, to provide optimal incentives, the firm must compare the workers' weighted likelihood ratios, leading to a relative performance criterion. That the optimal contract has a tournament character follows from risk neutrality.[15] Under bilateral contracting, the firm still wants to reward each worker based on his or her likelihood ratio, but there is no motive for relative evaluation. The firm optimally gives each worker an independent "one-step" incentive scheme: paying the maximum reward following a good outcome ($x^i$ greater than some cutoff) and no reward following a bad outcome [Levin 2000].

## III.C. Employee Bargaining Power

One notable aspect of multilateral contracting is that its largest gains are realized only if the firm enjoys significant bargaining power vis-à-vis its employees. This point is of interest if one considers the role of unions or other employee organizations in supporting a relational contract. Because such institutions

15. With risk aversion, relative performance evaluation would still be beneficial, but it need not take such a stark form. One reason not to introduce risk aversion in this dynamic setting, however, is that in addition to creating a motive for smoothing payments, it also creates a motive to smooth intertemporally, leading to a much less tractable analysis.

facilitate coordinated action by workers, there is a plausible argument that they could play a positive role in committing firms to behave well toward employees. The model suggests that this effect can be counteracted if they simultaneously reduce the firm's bargaining power.

To incorporate bargaining power into the model, suppose that at the outset of each period, the firm and its employees bargain over the continuing gains to trade (the surplus over and above $s(\emptyset)$), setting compensation accordingly. Suppose that the firm is able to command a share $\beta$ of these gains, so its net profits in any period are $\beta(s(\mathbf{e}) - s(\emptyset))$, while employees capture the remainder. In particular, suppose that each worker $i$ captures a share $\beta^i$. If $\beta$ is higher, the firm will be able to promise larger end-of-period bonuses to reward performance, but on the other hand, workers will be less concerned with the threat of being fired.

PROPOSITION 5. Under technological independence, greater employee bargaining power (lower $\beta$) lowers the joint surplus under multilateral contracting but not under bilateral contracting.

To see the argument, note that the firm's claim on future surplus is what gives it an incentive to deliver on discretionary compensation (any $b^i > 0$). Similarly, any claim on future surplus gives an employee a reason to perform even if it is not in his short-term interest (even if $c^i(e^i) < b^i$). However, there is a key difference with multilateral contracting. If the firm claims the future surplus, it may make credible commitments to any employee. But if worker $i$ has a claim on some share of future surplus, it can be used only to enforce $i$'s present performance and not to motivate $j$. Therefore, cross subsidization relies on the firm being in a relatively strong bargaining position.

With bilateral contracts the inability to cross-subsidize is irrelevant. If the firm commands any share $\beta$ of the future gains, then by setting $b^i = \beta(s^i(e^i) - s^i(\emptyset))$ (and choosing $w^i$ to ensure the right division of surplus), it is always possible to find a contract that will enforce $e^i$ if and only if $e^i$ satisfies the general bilateral constraint (2). Thus, a change in bargaining power changes the structure of compensation—the firm relies more on discretionary bonuses and less on high up-front wages if it has greater bargaining power—but not optimal incentives.

### *III.D. Production Externalities*

One further issue concerns the types of workers that are best grouped under a multilateral contract. The model suggests that the benefits of multilateral contracting are likely to be greatest when employees are substitutes in the production process. The reason is that the availability of substitute labor tends to undermine the firm's ability to make bilateral commitments. This problem often does not arise when employee performance is complementary. In this case bilateral contracts may do just as well as multilateral. An implication is that the gains to grouping employees who perform similar tasks, or have similar talents, may be large relative to the gains from grouping disparate employees.

To develop this point, consider two workers with production given by $y(e^1,e^2)$. Workers are *complementary* (*substitutable*) if $y(e^i,e^j)$ has positive (negative) cross-partial derivatives. To state the next result, let $\hat{e}^i$ denote the performance that is induced in an optimal single-worker contract with $i$.

PROPOSITION 6. (i) If workers are substitutable, bilateral contracts are strictly worse than multilateral unless both achieve first-best.[16] (ii) If workers are complementary, bilateral contracts are as effective as multilateral whenever the optimal multilateral contract sets $e^i \geq \hat{e}^i$ for $i = 1,2$ and $\hat{e}^i$ is not first-best (i.e., does not maximize $s(e^i,\varnothing)$).

## IV. CONTRACTUAL ADJUSTMENT

The model developed above suggests that firms benefit from making broad multilateral commitments to their employees because such commitments are harder to back away from. A potential downside of being bound to a broad set of implicit promises is that organizational changes may become difficult or costly to implement. The idea that firms find it extremely difficult to adjust employment practices in response to macroeconomic or more specific shocks is a central theme of human resource studies [Doeringer and Piore 1971; Bewley 1999]. And there is anecdotal evidence that organizational changes can be particularly problematic at firms where employees have long-standing expecta-

16. Recently Spagnolo [1999] has shown a related result: if firms compete in two markets and have utility functions that are submodular in the profits from the two markets (e.g., have concave utility over the sum of profits), then multimarket contact will facilitate collusion.

tions of job security or particular benefits (Baron and Kreps [1999] provide some insightful examples). Rogers and Beer [1995], for instance, describe Hewlett-Packard's difficult move away from a long-time implicit promise of lifetime employment in response to increased demand volatility in the 1980s. They document the employee discontent and loss of morale at the time. Interestingly, one step taken by HP was to initiate a conservative hiring policy and rely more on temporary and "flex-force" workers [Rogers and Beer 1995]. As described in the Introduction, Microsoft more recently has adopted a strategy of hiring both temporary and permanent employees. Indeed, the popular press reports a growing tendency, particularly in high-tech, toward this sort of "two-tier" employment system. Abraham and Taylor [1996] provide some empirical evidence that one motivation for such contracting out can be to buffer the permanent workforce from shocks.[17]

In this section the model is extended to capture the idea that by targeting its promises to employees a firm can manage expectations about how it will respond to shocks. In contrast to the earlier setup, it is assumed that the environment is not stable but changes over time. Initially, the firm proposes a set of stationary contracts. It may change these contracts at the beginning of each period, but doing so involves communicating with all involved workers. I assume that this process is difficult in that it involves some probability of miscommunication. The assumption is meant to capture the idea that a successful relationship requires a high level of shared information, and that maintaining this level of mutual understanding when the environment changes is difficult. Indeed, Bewley [1999] cites managers as saying that their workers would not respond negatively to selective pay cuts or layoffs if these actions could be clearly explained and justified. Since Bewley's managers also report not making these changes, it seems likely that they believe such explanations are nontrivial. In the next section I simply assume that communication may fail, in which case workers who were supposed to be informed of the change are not. They then act contrary to the new agreement, or

17. There are other reasons why firms might want to contract out, some of which are discussed by Abraham and Taylor [1996]. Buffering the permanent workforce seems to have been one of Microsoft's motivations: "Microsoft was concerned about the volatile and unsettled nature of the industry, and with how to achieve the ultimate goal of avoiding the potential for layoffs or overstaffing," Microsoft lawyer, quoted in "Revenge of the Temps," by Kirstin Downey Grimsley, *Washington Post,* January 16, 2000, Section H, Page 1.

interpret the firm's actions as contrary to their (now outdated) expectations. In subsection IV.B I show how "miscommunication" can arise as an equilibrium phenomenon if there is asymmetric information about the necessity of layoffs.

Introducing a cost to contractual adjustment has two effects: first, even with a single agent, it introduces inertia into the way relational contracts adapt to shocks; second, it means that altering multilateral contracts may be more costly than altering bilateral contracts. I focus on the latter effect—while inertia is an important issue in structuring long-term relationships, a careful consideration would lead too far afield.

### *IV.A. Adjustment Costs and Contracting*

Suppose that there are two workers who initially, if employed, generate gross profits $y^i(e^i)$. With Poisson probability $1 - \gamma$, a one-time shock arrives. Following the shock, the firm requires only the first worker—the second has negative value. At this point, the firm will let the second worker go and potentially adjust the contract of the first. I assume that a contract can be adjusted successfully with probability $1 - \eta$.

The firm can avoid contractual adjustment by using bilateral contracts. The optimal contract with the first worker solves

$$\max_e s^1(e) \quad \text{subject to} \ \frac{\delta}{1-\delta}\,(s^1(e) \ - \ s^1(\varnothing)) \ \geq \ c^1(e),$$

while the contract with the second worker solves

$$\max_e s^2(e) \quad \text{subject to} \ \frac{\gamma\delta}{1-\gamma\delta}\,(s^2(e) - s^2(\varnothing)) \geq c^2(e).$$

Suppose that these contracts result in performance $\mathbf{e}^B = (e^{1,B}, e^{2,B})$. There is no need to explain changes following the shock: the first worker's contract remains the same; while the second is let go.

Now suppose that the initial contract is multilateral. Following the shock, the firm must lay off the second worker, and can successfully explain the layoff with probability $1 - \eta$. If successful, future performance will be $e^{1,B}$. If communication is not successful, the firing of the second worker will be interpreted as a breach of the initial contract in which case the first will sanction the firm by quitting. Prior to the shock, the effort profile of the optimal multilateral contract is the solution to

$$\max_{e^1, e^2} s^1(e^1) + s^2(e^2)$$

$$\text{subject to } \frac{\delta}{1-\delta}\left(\xi[s^1(e^1) + s^2(e^2) - s^1(\varnothing) - s^2(\varnothing)] + (1-\xi)(1-\eta)[s^1(e^{1,B}) - s^1(\varnothing)]\right) \geq c^1(e^1) + c^2(e^2),$$

where $\xi = \gamma(1 - \delta)/(1 - \gamma\delta)$ is the present value share of future periods before the shock. Denote the performance level that solves this program as $\mathbf{e}^M = (e^{1,M}, e^{2,M})$.

PROPOSITION 7. A multilateral contract is optimal if and only if the gains from cross subsidization are greater than the cost of contractual adjustment; i.e., if

$$s(\mathbf{e}^M) - s(\mathbf{e}^B) \geq \frac{\delta(1-\gamma)}{1-\delta}\eta(s^1(e^{1,B}) - s^1(\varnothing)).$$

The gains from multilateral enforcement depend on the value of balancing initial incentives as well as the expected waiting time until the environment changes. The more stable the environment, the greater the gain from multilateral enforcement. On the other hand, the expected adjustment costs are increasing in the probability of miscommunication $\eta$ and in the value of the first relationship, but decreasing in the expected waiting time until the shock arrives. If the initial environment is stable and there exist gains to equalizing incentives, a multilateral contract will be optimal. If the environment is unstable, and especially if the first worker is valuable, the firm prefers a bilateral contracting regime.

*IV.B. Asymmetric Information and Adjustment Costs*

The model above incorporates adjustment costs by simply assuming that communication is imperfect, so the firm cannot costlessly restructure its relational contract with the worker who remains after the layoff. Similar dynamics arise with no direct assumption about imperfect communication, but rather as an equilibrium phenomenon, if there is asymmetric information about the necessity of layoffs, so the firm cannot credibly communicate its intentions. The underlying logic is similar to Green and Porter [1984], where asymmetric information necessitates equilibrium price wars among colluding firms. The novel implication is that these equilibrium sanctions can be so costly in terms of forgone surplus from the relationship with "core" workers that the gains to providing incentives for "peripheral" workers are swamped.

Consider the same situation, where the firm initially can employ two workers, but with Poisson probability $1 - \gamma$ must lay off the second worker. The firm learns the necessity of this layoff immediately prior to paying the previous period's bonus. That is, if the second worker must be laid off in period $t$, the firm learns of this prior to paying $b_{t-1}$. The first worker does not observe whether conditions necessitate a layoff, but does observe the firm's behavior toward the second worker.

For clarity, assume that performance is a binary choice, so that $e^i \in \{0,\bar{e}\}$. With bilateral contracts, the performance of the first worker can be ensured if and only if

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

while the performance of the second worker can be ensured if and only if

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

There are several cases depending on whether these constraints are satisfied, but the interesting situation is where only the performance of the first worker can be ensured under bilateral contracts.[18] Assume that this is the case, the per-period surplus under bilateral contracting is

$$s^B = s^1(\bar{e}) + s^2(\varnothing).$$

A multilateral contract works by specifying compensation $(\mathbf{w},\mathbf{b})$ and enforcing performance as follows. If either worker fails to perform, the firm withholds discretionary compensation and fires the worker. If the firm deviates against the first worker, both workers quit. If the firm fails to pay or fires the second worker, the second worker departs, and the first departs with probability $\alpha$. With prob-

18. If neither constraint is satisfied, then neither bilateral or multilateral contracting can induce any performance, while if both are satisfied, then bilateral contracts already achieve the first-best and there is no benefit to multilateral contracting. If the second constraint is satisfied, but the first fails, the first worker will necessarily depart when the second worker does. In this case, multilateral contracting may be beneficial if it can induce the first worker to perform prior to the layoff.

ability $1 - \alpha$, the first remains as an employee at the same compensation, and that employment relationship continues.[19]

Under a multilateral contract, the second worker's performance is supported by surplus generated by the second worker, and by the first worker's threat to quit if there is mistreatment. This sort of cross subsidization is feasible whenever

$$(s^1(\bar{e}) - s^1(\varnothing)) + (s^2(\bar{e}) - s^2(\varnothing)) \geq \frac{1-\delta}{\delta}\, c^1(\bar{e}) + \frac{1-\gamma\delta}{\gamma\delta}\, c^2(\bar{e}).$$

However, cross subsidization now comes at a cost. In equilibrium the firm will find it necessary to terminate the second worker, and the first will quit inefficiently with probability $\alpha$. This cost of adjustment must be set large enough so that the firm does not have an incentive to terminate the second worker unnecessarily and claim necessity. Formally, $\alpha$ must be chosen so that

$$\alpha(s^1(\bar{e}) - s^1(\varnothing)) + (s^2(\bar{e}) - s^2(\varnothing)) \geq \frac{1-\gamma\delta}{\gamma\delta}\, c^2(\bar{e}).$$

The optimal contract selects the smallest $\alpha$ that satisfies the constraint.

The surplus from a multilateral contract that induces performance by both workers (normalized by $(1 - \delta)$) will be

$$s^M = \frac{1-\delta}{1-\delta\gamma}\,(s^1(\bar{e}) + s^2(\bar{e})) + \frac{\delta(1-\gamma)}{1-\delta\gamma}\,((1-\alpha)s^1(\bar{e}) + \alpha s^1(\varnothing) + s^2(\varnothing)).$$

Substituting in the optimal $\alpha$ gives

$$s^M = s^1(\bar{e}) + s^2(\bar{e}) - \frac{1-\gamma}{\gamma}\, c^2(\bar{e}).$$

PROPOSITION 8. Multilateral contracting improves on bilateral contracting if and only if

$$s^2(\bar{e}) - s^2(\varnothing) \geq \frac{1-\gamma}{\gamma}\, c^2(\bar{e}).$$

19. A clever alternative that this rules out is for the firm to promise a transfer (or pay raise) to the first worker in the event that it fires the second. Such an arrangement can solve the problem of the firm wanting to deviate by firing just the second worker. However, it succeeds only by giving the first worker a corresponding incentive to get rid of the second worker. So if the first worker can induce the second to quit, the cost of cross subsidization immediately returns, and the analysis is identical to what is done here.

As in the case of exogenous communication problems, a multilateral contract is preferred only if the environment is sufficiently stable ($\gamma$ is large), or if the gains from cross subsidization ($s^2(\bar{e}) - s^2(\varnothing)$) are large.

### *IV.C. Two-Tier Contract Structures*

The difficulty of contractual adjustment can be used to explain the use of "two-tier" (or multitier) contract structures. Suppose that the firm initially employs $n$ workers, each of whom produces $s(e) + z^i$. With Poisson probability $1 - \gamma$, the environment changes, and the firm requires only $m < n$ workers. Returning to the case of exogenous adjustment costs (a similar result can be given with endogenous adjustment costs), suppose that the firm faces a probability $\eta$ of miscommunicating with any agent who is retained, and whose contract the firm wishes to adjust (these probabilities can be independent or correlated).

The following proposition shows that of the many possible initial contract structures, only a relatively few could possibly be optimal.

PROPOSITION 9. There is always an optimal contract structure involving either a single multilateral contract with all employees, or a pair of contracts with two distinct groups of workers. With two groups, one is stable and composed only of workers to be retained, while the other includes all agents to be released and possibly workers to be retained as well.

The argument is a simple one. If the firm contracts separately with two groups of workers, all of whom are to be retained, it can do at least as well by reaching a unified agreement with both groups. On the other hand, if the firm contracts separately with two groups of workers, both of which contain employees who are to be released, it will have to recontract with any worker in either group who is to be retained. Thus, joining the groups initially in a single contract involves no additional cost and potential multilateral enforcement gains. It follows that there will be an optimal contract with at most one group that includes agents to be released and (potentially) another group that contains only agents to be retained.

## V. Conclusion

Many aspects of employment can be usefully thought of as elements of a long-term relational contract between a firm and its workers. Such contracts capture the idea that workers are motivated by the prospect of continuing employment or by rewards that are never formally guaranteed in advance, while firms are discouraged from short-term expropriation by the prospect of alienating employees. In such a setting, firms face important choices in how to structure workers' expectations. By encouraging workers to view its behavior broadly, a firm can effectively commit itself to better treatment of employees. Such a strategy, however, risks widespread problems if the firm wants to make changes at a later date. If the environment is volatile in ways that cannot be perfectly planned for in advance, it may be better to make more targeted commitments and create separate relationships with different employees or sets of employees.

The dynamics described by the model apply best to firms where workers have relatively high expectations about the firm's behavior (i.e., where a relational contract is a good description of the employment relationship). Consider, for instance, the case of Lincoln Electric Company, a manufacturer of arc welding equipment that is widely noted for its use of piece-rate incentives and high productivity. Lincoln's recruitment materials emphasize that the company has not laid off workers in over 50 years and has paid an annual year-end bonus to each worker for over 60 years. The cost of backing away from these policies, even by making targeted layoffs or withholding the bonus of some deserving workers, would presumably be quite large. Indeed, in an article recounting Lincoln's experience in 1992 when it suffered large losses from poorly planned overseas expansion, former CEO Donald Hastings stressed that management's crucial concern was finding sufficient resources to pay the annual bonus for U. S. workers. As he explained it, "if we didn't pay the bonus, the whole company might unravel" [Hastings 1999].

The difficulties of backing away from a perceived broad commitment to employees are further illustrated by the many case studies of firms that have found themselves forced to make layoffs or other changes when their market position changed. Mills and Freisen [1996] document IBM's experience in the late 1980s and early 1990s, when it instituted widespread layoffs for the first time in response to operating losses. Although IBM offered no formal guarantee of employment security, Mills and Freisen argue that the layoffs were widely perceived as violating an implicit promise to the workforce and that this led to significant discontent and loss of morale.

Interestingly, like Hewlett-Packard, one of IBM's responses following layoffs was to increase reliance on flex-force and part-time employees. At both companies, there is reason to suspect that one motivation for doing so was to allow full-time workers a measure of job security as in the above model. I have argued above that Microsoft could have had similar motives as well. Given the striking rise in temporary and contingent employment over the last decade, an intriguing empirical question is how important this buffering explanation is relative to other motives for using contingent labor.[20]

Moving beyond employment, the general trade-off considered in this paper is relevant for other organizational decisions as well. For instance, many researchers have stressed that long-term relationships are a central feature of the supply contracting that takes place in Japanese manufacturing [Dore 1983; Williamson 1985; Taylor and Wiggins 1997]. A number of aspects of these relationships are usefully thought of as multilateral. As an example, Williamson [p. 121] describes Toyota's practice of organizing suppliers into loose-knit associations that serve as a forum to discuss reputationally relevant information. Dore reports similar arrangements in other Japanese enterprise groups. Part of the function of these supplier groups seems to be to help promote coordinated sanctions in the event of questionable behavior by a manufacturer.

An in-depth study of the Japanese automotive and electronic sectors by Asanuma [1989] suggests a further subtlety. Asanuma observes the practices cited by Williamson [1985] and Dore [1983], but notes that manufacturers tend to class their suppliers into different tiers. Relationships with first-tier suppliers are generally secure, but there may be significant turnover among lower-tier suppliers. Asanuma hypothesizes that one value of

20. Among the other explanations for the increase in contingent labor is that courts have increasingly limited the ability of firms to terminate "full-time" employees—in essence turning what was previously a relational contract into a formal legal commitment. Another is that having two modes of employment allows firms to screen different types of workers. The popular press emphasizes contingent labor as a cost-saving device, claiming that firms prefer hiring contingent workers to whom they can pay low wages and benefits. It is not obvious how the last story fits with standard economic theory.

having marginal suppliers is that they act as a buffer that helps to protect relations with upper-tier suppliers, something that is particularly important in periods of uncertain or intermittent demand. Thus, there seem to be both bilateral and multilateral elements at work.

A further related problem concerns customer relationships. Since Klein and Leffler [1981], it has been common to view a firm's reputation for quality as an implicit contract with consumers. To capture this, the model must be turned around, so that the "workers" pay the "firm" to perform but the choice between bilateral and multilateral contracts is still relevant. It can be interpreted, for instance, as the choice between maintaining a set of distinct and separate brands and building a more general reputation for quality. An in-depth analysis of this problem would be interesting to pursue.

## Appendix

*Proof of Proposition 1.* The argument follows standard repeated game logic. Worker $i$ will be willing to perform in each period if and only if

$$\frac{\delta}{1-\delta}(W^i - c^i(e^i) - \bar{u}^i) \geq \max\{0, c^i(e^i) - b^i\},$$

while the firm is willing to make bonus payments in each period if any only if

$$\frac{\delta}{1-\delta}\left(y(\mathbf{e}) - \sum_i W^i - y(\varnothing)\right) \geq \sum_i b^i.$$

Summing the worker's incentive constraints for $i = 1, \ldots, n$, and the firm's constraint gives the stated condition as a necessary condition for $(\mathbf{e},\mathbf{w},\mathbf{b})$ to be supported. To see that it is sufficient, let $w^i = \bar{u}^i + k^i$ and $b^i = c^i(e^i) - \delta k^i$ for any $0 \leq k^i \leq c^i(e^i)/\delta$. QED

*Proof of Proposition 2.* Consider a contract given by $(\mathbf{e},\mathbf{w},\mathbf{b})$. No worker individually prefers to choose zero performance and quit if and only if for all $i$,

$$\frac{\delta}{1-\delta}(W^i - c^i(e^i) - \bar{u}^i) \geq \max\{0, c^i(e^i) - b^i\}. \tag{4}$$

In addition, the firm can always identify a subset $I$ of employees, withhold the discretionary payments to other employees, and then recontract with those in $I$, terminating the others. For this to be unprofitable, it must be that for each subset $I$ of employees,

$$(5)\qquad \frac{\delta}{1-\delta}\left(\left(s(\mathbf{e}) - \sum_{j\notin I}(W^i - c^i(e^i))\right) - \left(s^I - \sum_{j\notin I}\bar{u}^j\right)\right) \geq \sum_{j\notin I} b^j.$$

Summing (4) and (5) implies the stated inequality, making it a necessary condition for self-enforcement. On the other hand, if the stated inequality holds, payments $(\mathbf{w},\mathbf{b})$ are easily found to satisfy the individual worker constraints, and each of the firm constraints (take $w^i = \bar{u}^i$ and $b^i = c^i(e^i)$ for each employed worker. QED

*Proof of Proposition 4.* I assume that the Mirrlees-Rogerson conditions hold: for all $i$, $f^i(x|e^i)$ has the monotone likelihood ratio property ($f^i_e/f^i$ is increasing in $x$), and $F^i(x|c^{-1}(q))$ is convex in $q$.

Suppose that contracting is multilateral. I claim that a necessary and sufficient condition for a performance profile $\mathbf{e}$ to be enforced is that there are payments $(\mathbf{w},\mathbf{b})$, with $b^i \geq 0$ a function of $(x^1, \ldots, x^n)$, such that for all $i$ with $e^i \neq \varnothing$:

$$(6)\qquad e^i \in \arg\max_{\tilde{e}^i} \mathbb{E}[b^i(\mathbf{x})|(\tilde{e}^i,e^{-i})] - c(\tilde{e}^i),$$

and for all $\mathbf{x}$,

$$(7)\qquad \frac{\delta}{1-\delta}(s(\mathbf{e}) - s(\varnothing)) \geq \sum_i b^i(\mathbf{x}).$$

The first condition is an incentive compatibility condition for each agent's effort choice. The second is a self-enforcement constraint. To see that these conditions are sufficient, set $w^i = \bar{u}^i$. It is easily checked that agents are willing to choose $e^i$ and that the firm will deliver payments rather than terminate all relationships. As for necessity, clearly (6) must hold. And if (7) fails, the firm will renege on payments for some $\mathbf{x}$, or not all workers will agree to be employed.

The optimal contract solves

$$\max_{\mathbf{e},\mathbf{b}(\mathbf{x})} s(\mathbf{e}) \quad \text{subject to (6),(7).}$$

Under the Mirrlees-Rogerson conditions, the incentive compatibility conditions can be replaced by a first-order condition for each agent's performance choice:

$$\frac{d}{de^i}\{\mathbb{E}[b^i(\mathbf{x})|(e^i,e^{-i})] - c^i(e^i)\} = 0.$$

Let $\lambda(\mathbf{x}) \geq 0$ denote the Lagrange multiplier on the enforcement constraint, (7), and $\mu^i > 0$ denote the multiplier on each first-order condition. Then if the optimal contract sets $b^i(\mathbf{x}) > 0$ *and* $b^j(\mathbf{x}) > 0$ for some $\mathbf{x}$,

$$\mu^i \frac{f_e^i}{f^i}(x^i|e^i) = \mu^j \frac{f_e^j}{f^j}(x^j|e^j) = \frac{\lambda(\mathbf{x})}{\Pi_i f^i(x^i|e^i)}.$$

Given that $f^i$ has the MLRP, the first term is increasing in $x^i$. So fixing $x^{-i}$, there can be at most $n - 1$ values of $x^i$ such that $b^i(\mathbf{x}) > 0$ and $b^j(\mathbf{x}) > 0$ for some $j$. Thus, except on a set of (lebesque) measure zero, at most one agent gets a positive bonus. The optimal contract sets

$$b^i(\mathbf{x}) = \begin{cases} \dfrac{\delta}{1-\delta}(s(\mathbf{e}) - s(\varnothing)) & \text{if } \mu^i\left(\dfrac{f_e^i}{f^i}\right) > 0 \text{ and } \mu^i\left(\dfrac{f_e^i}{f^i}\right) > \mu^j\left(\dfrac{f_e^j}{f^j}\right) \\ & \qquad \text{for all } j \neq i, \\ 0 & \text{otherwise.} \end{cases}$$

That is, if the outcome $\mathbf{x}$ is such that all agents have a negative likelihood ratio, all get zero discretionary payment. If at least one agent has a positive likelihood ratio, the maximum payment is given to the agent with the highest weighted likelihood ratio: $\mu^i(f_e^i/f^i)(x^i|e^i)$.

For the bilateral case, Levin [2000] shows that agent $i$ optimally receives the maximum payment $\delta/(1-\delta)(s^i(e^i) - s^i(\varnothing))$, if and only if his own likelihood ratio, $f_e^i/f^i$, is positive. QED

*Proof of Proposition 5.* Consider the multilateral case, where the firm captures a share $\beta$, and each worker $i$ a share $\beta^i$ of the per-period surplus $s(\mathbf{e}) - s(\varnothing)$. For worker $i$ to cooperate with a contract that specifies performance $\mathbf{e}$ and bonuses $\mathbf{b} \geq 0$ requires

$$\frac{\delta}{1-\delta}\beta^i(s(\mathbf{e}) - s(\varnothing)) \geq \max\{0, c^i(e^i) - b^i\}.$$

For the firm to make rewards requires that

$$\frac{\delta}{1-\delta}\,\beta(s(\mathbf{e})-s(\varnothing)) \geq \max\left\{0,\ \sum_i b^i\right\}.$$

Recall that **e** can be sustained for some configuration of bargaining power if and only if the sum of these constraints holds (Proposition 1). With fixed bargaining weights, however, for many performance profiles **e**, worker *i*'s constraint will necessarily be slack if $\beta^i$ is sufficiently large. When this is the case, even if **e** satisfies the general multilateral constraint from Proposition 1, it may be impossible to sustain with bargaining as the structure of payments may not permit slack to be shifted between the different constraints. This problem never arises if the firm has sufficient bargaining power. If $\beta = 1$, then choosing $b^i(e^i) = c^i(e^i)$ for all $i$ ensures that all constraints will hold if their sum does. This issue also does not arise with bilateral contracts, where we model the firm as bargaining for a share $\beta$ of each individual surplus $s^i(e^i) - s^i(\varnothing)$, as the payment scheme described in the text succeeds in balancing slack between the constraints. QED

*Proof of Proposition 6.* An optimal bilateral contract solves

$$\max_{\mathbf{e}} s(e^1,e^2) \tag{8}$$

$$\text{subject to } \frac{\delta}{1-\delta}\,(s(e^1,e^2) - s(\varnothing,\varnothing)) \geq c^1(e^1) + c^2(e^2),$$

$$\text{and } \frac{\delta}{1-\delta}\,(s(e^1,e^2) - s(\hat{e}^i,\varnothing)) \geq c^j(e^j) \quad \text{for } i = 1,2, \tag{9}$$

where $s(\hat{e}^i,\varnothing)$ is the joint surplus and $\hat{e}^i$ the optimal performance if the firm employs only worker $i$. Only the first constraint need hold under a multilateral contract, so this arrangement is superior exactly when the possibility of deviating against one worker and recontracting with the other constrains incentives.

As $\hat{e}^i$ is optimal if the firm contracts with only $i$,

$$\frac{\delta}{1-\delta}\,(s(\hat{e}^i,\varnothing) - s(\varnothing,\varnothing)) \geq c^i(\hat{e}^i). \tag{10}$$

This holds with equality so long as $\hat{e}^i$ is not first-best (i.e., does not maximize $s(e^i,\varnothing)$). If $\hat{e}^i$ is not first-best, then it is strictly less than first-best, and (10) is strictly satisfied for any $e^i < \hat{e}^i$.

($i$) Suppose that efforts are substitutes. I claim that the optimal bilateral contract sets $e^1 \leq \hat{e}^1$ and $e^2 \leq \hat{e}^2$. Let $e^{i*}(e^j)$ denote the level of $e^i$ that maximizes $s(e^i,e^j)$ for any given $e^j$. By Topkis' Theorem, $e^{i*}(e^j)$ is decreasing in $e^j$. It follows that if $\hat{e}^i$ is first-best (i.e., if $\hat{e}^i = e^{i*}(\varnothing)$), the optimal bilateral contract cannot set $e^i > \hat{e}^i$, since in this case, lowering $e^i$ to $e^{i*}(e^j)$ (for whatever $e^j$ was specified) would increase surplus and relax all constraints. This proves the claim if $\hat{e}^1,\hat{e}^2$ are both first-best. Suppose that $\hat{e}^1$ is not first-best, but $\hat{e}^2$ is. From above, restrict attention to $e^2 \leq \hat{e}^2$. Then if $e^1 > \hat{e}^1$, (9) will fail:

$$\frac{\delta}{1-\delta}(s(e^1,e^2) - s(\varnothing,\hat{e}^2)) \leq \frac{\delta}{1-\delta}(s(e^1,e^2) - s(\varnothing,e^2))$$

$$\leq \frac{\delta}{1-\delta}(s(e^1,\varnothing) - s(\varnothing,\varnothing)) < c^1(e^1).$$

Last, suppose that neither $\hat{e}^1,\hat{e}^2$ are first-best. Then if $e^2 \leq \hat{e}^2$ and $e^1 > \hat{e}^1$, the same argument shows again that (9) must fail. And if $e^1 > \hat{e}^1$ and $e^2 > \hat{e}^2$, it is easy to verify that (8) must fail. Thus, regardless of whether $\hat{e}^1,\hat{e}^2$ are first-best, it must be that $e^1 \leq \hat{e}^1$ and $e^2 \leq \hat{e}^2$ at the bilateral optimum. Furthermore, one of these inequalities must be strict since (8) fails with $e^1 = \hat{e}^1$ and $e^2 = \hat{e}^2$ when (10) binds, and if $\hat{e}^i$ is first-best, it is optimal to lower $e^i$ below $\hat{e}^i$.

To complete the proof, observe that (8) will be slack at the bilateral optimum. Indeed, whenever (9) holds for $e^1 < \hat{e}^1$ and $e^2 \leq \hat{e}^2$,

$$\frac{\delta}{1-\delta}(s(e^1,e^2) - s(\varnothing,\varnothing)) = \frac{\delta}{1-\delta}(s(e^1,e^2) - s(\hat{e}^1,\varnothing))$$

$$+ \frac{\delta}{1-\delta}(s(\hat{e}^1,\varnothing) - s(\varnothing,\varnothing)) \geq c^2(e^2) + c^1(\hat{e}^1),$$

so (8) must be slack. Thus, a multilateral contract strictly improves on bilateral contracts unless both achieve first-best.

($ii$) Suppose that efforts are complements, and suppose that the optimal multilateral contract sets $e^1 \geq \hat{e}^1$ and $e^2 \geq \hat{e}^2$. Because $e^1,e^2$ satisfy (8), they must also satisfy (9), since

$$\frac{\delta}{1-\delta}\,(s(e^1,e^2) - s(\hat{e}^i,\varnothing)) = \frac{\delta}{1-\delta}\,(s(e^1,e^2) - s(\varnothing,\varnothing))$$

$$-\frac{\delta}{1-\delta}\,(s(\hat{e}^i,\varnothing) - s(\varnothing,\varnothing)) \geq c^1(e^1) + c^2(e^2) - c^i(\hat{e}^i) \geq c^j(e^j).$$

So the multilateral contract optimum also satisfies the bilateral constraints.

More generally, with complementary performance, the multilateral optimum may have $e^i > \hat{e}^i$ and $e^j < \hat{e}^j$ (it will never have $e^1 \leq \hat{e}^1$ and $e^2 \leq \hat{e}^2$), while the bilateral optimum always has $e^i \geq \hat{e}^i$ and $e^j \geq \hat{e}^j$. In this event, which requires a significant asymmetry between employees, the multilateral contract does better. QED

STANFORD UNIVERSITY

## REFERENCES


Abraham, Katharine and Susan Taylor, "Firms' Use of Outside Contractors: Theory and Evidence," *Journal of Labor Economics,* XIV (1996), 394–424.

Asanuma, Banri, "Manufacturer-Supplier Relationships in Japan and the Concept of Relation-Specific Skill," *Journal of the Japanese and International Economies,* III (1989), 1–30.

Baker, George, Robert Gibbons, and Kevin J. Murphy, "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics,* CXVII (2002), 39–84.

Baron, James, and David Kreps, *Strategic Human Resources: Frameworks for General Managers* (New York, NY: John Wiley & Sons, 1999).

Bendor, Jonathan, and Dilip Mookherjee, "Norms, Third-Party Sanctions, and Cooperation," *Journal of Law, Economics, and Organization,* VI (1990) 33–63.

Bernheim, B. Douglas, and Michael Whinston, "Multimarket Contact and Collusive Behavior," *Rand Journal of Economics,* II (1990), 1–26.

Bewley, Truman, *Why Wages Don't Fall During a Recession* (Cambridge, MA: Harvard University Press, 1999).

Bull, Clive, "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics,* CII (1987), 147–159.

Che, Yeon-Koo, and Seung-Weon Yoo, "Optimal Incentives for Teams," *American Economic Review,* XCI (2001), 525–541.

Doeringer, Peter, and Michael Piore, *Internal Labor Markets and Manpower Analysis* (Lexington, MA: Heath, 1971).

Dore, Ronald, "Goodwill and the Spirit of Market Capitalism," *British Journal of Sociology,* XXXIV (1983), 459–482.

Fudenberg, Drew, and David Kreps, "Reputation in the Simultaneous Play of Multiple Opponents," *Review of Economic Studies,* LIV (1987), 541–568.

Green, Edward, and Robert Porter, "Noncooperative Collusion under Imperfect Price Information," *Econometrica,* LII (1984), 87–100.

Grief, Avner, "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies," *Journal of Political Economy,* CII (1994), 912–950.

Grief, Avner, Paul Milgrom, and Barry Weingast, "Coordination, Commitment and Enforcement: The Case of the Merchant Guild," *Journal of Political Economy,* CII (1994), 745–776.

Hastings, Donald, “Lincoln Electric’s Harsh Lessons from International Expansion,” *Harvard Business Review* (May–June 1999), 3–11.
Klein, Benjamin, and Keith Leffler, “The Role of Market Forces in Assuring Contractual Performance,” *Journal of Political Ecomony,* LXXXIX (1981), 615–641.
Kreps, David, “Corporate Culture and Economic Theory,” *Perspectives on Positive Political Economy,* James Alt and Kenneth Shepsle, eds. (Cambridge: Cambridge University Press, 1990).
Levin, Jonathan, “Relational Contracts, Incentives and Information,” Ph.D. Thesis, Massachusetts Institute of Technology, 1999.
——, “Relational Incentive Contracts,” Stanford University, 2000.
MacLeod, Bentley, and James Malcomson, “Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment,” *Econometrica,* LVII (1989), 447–480.
MacLeod, Bentley, and James Malcomson, “Motivation and Markets,” *American Economic Review,* LXXXVIII (1998), 388–411.
Malcomson, James, “Work Incentives, Hierarchy, and Internal Labor Markets,” *Journal of Political Economy,* XCII (1984), 486–507.
Mills, D. Quinn, and G. Bruce Friesen, *Broken Promises* (Boston, MA: Harvard Business School Press, 1996).
Rogers, Gregory, and Michael Beer, “Human Resources at Hewlett-Packard,” Harvard Business School Cases 9-495-051 and 9-495-052, 1995.
Shapiro, Carl, and Joseph Stiglitz, “Equilibrium Unemployment as a Worker Discipline Device,” *American Economic Review,* LXXIV (1984), 433–444.
Spagnolo, Giancarlo, “On Interdependent Supergames: Multimarket Contact, Concavity, and Collusion,” *Journal of Economic Theory,* LXXXIX (1999), 127–139.
Taylor, Curtis, and Steven Wiggins, “Competition or Compensation: Supplier Incentives under the American and Japanese Subcontracting Systems,” *American Economic Review,* LXXXVII (1997), 598–618.
Williamson, Oliver, *The Economic Institutions of Capitalism* (New York, NY: Free Press, 1985).