Influence Transfers, Performance, and Performance Ratings

Birger Wernerfelt *
Duncan Simester **
John R. Hauser *

February 1997 WP # 160-97

Sloan WP # 3949

* MIT Sloan School of Management
** Graduate School of Business, University of Chicago

© 1997 Massachusetts Institute of Technology

Sloan School of Management
Massachusetts Institute of Technology
38 Memorial Drive, E56-390
Cambridge, MA 02139-4307
Influence Transfers, Performance, and Performance Ratings†

by

Birger Wernerfelt*, Duncan Simester**, and John R. Hauser*

February, 1997

*MIT Sloan School, Cambridge, MA 02139
**Graduate School of Business, University of Chicago, Chicago, IL 60637
†We are indebted to George Baker and Jim Rebitzer for comments on an earlier version of this paper. The authors are in reverse alphabetical order.
Influence Transfers, Performance, and Performance Ratings

Abstract

Most large companies use formal systems in which supervisors rate the performance of employees. These systems are normally seen as reporting instruments which help to resolve moral hazard by revealing information about each employee's efforts. However, a number of studies have cast doubt on this interpretation by showing that actual ratings exhibit very little variance and thus do not distinguish between effort levels. We here identify a second function of rating systems, namely that they induce collusion which facilitates social transfers from ratees to raters. Not all of these transfers are inefficient "influence costs"; in some cases the value of the transfers received by the supervisor may greatly exceed the cost to the employee of their provision. We show how the firm can design reward systems such that only efficient transfers take place. We also demonstrate that, in the presence of such transfers, rating systems may indeed resolve moral hazard despite their lack of merit as reporting instruments.

To support our argument, we show that firms can design collusion-proof rating schemes. The fact that they do not use such schemes is consistent with our contention that collusion helps them take advantage of social transfers. Finally, we make an initial measurement of the cost and value of social transfers using a questionnaire study of one hundred and twenty-one managers. We find that the potential value created by transfers, defined as the perceived value to the supervisor less the perceived cost to the employee, can be quite substantial.
1. Introduction

Most large companies use subjective performance evaluation systems in which supervisors rate employees. Based on the belief that supervisors have better information about employees, these rating systems are seen as tools to control moral hazard. However, a number of studies (e.g. Medoff and Abraham, 1980) have pointed out that nominal ratings exhibit surprisingly little variance, thus casting doubt on the efficacy of the rating systems as reporting systems.

In the present paper we identify a second function of subjective rating systems; namely that they induce collusion between supervisors and employees which in turn facilitates social transfers from the employees to the supervisors. To the extent that these transfers are efficient they create value for the firm. Supervisors who enjoy their jobs, because they perceive that employees respect them and have a good understanding of their expectations, may require less monetary compensation from the firm. Although the firm may need to compensate employees for favoring their supervisors with additional respect, we offer evidence that the benefit of these services to the supervisor may exceed the cost to the employees.

Rating systems facilitate social transfers by allowing supervisors to compensate employees for their transfers with higher ratings and ultimately higher salaries from the firm. Endowing supervisors with the power to determine employee’s salaries is costless to the firm as long as it does not affect the actions that are implemented, the expected compensation of either agent, or the exposure of either agent to additional risk. We show that rating systems can achieve each of these goals and still allow firms to appropriate all the gains from the agents’ social trade. In doing so we also show that despite rating systems’ lack of efficacy as reporting systems, they may still provide effective incentives for resolving moral hazard. When the compensation received by an employee is
understood to include the cost of social transfers together with the monetary remuneration received from the firm, rating systems are capable of offering incentives to exert effort.

The intuition behind these results begins with the observation that the firm can appropriate the gains from the agents’ social trade by anticipating the social transfers and factoring them into the employment contracts. The firm reduces the supervisor’s salary by the value of the social transfers received while increasing the salary of the employee to compensate for the cost of the transfers provided. Second, the supervisor’s contract can include penalties for over-stating the employee’s effort, reducing the supervisor’s salary when observed output is low and the supervisor reported effort was high. Although the employee may offer social transfers to compensate the supervisor for these penalties, lower effort requires larger transfers. Hence, although ratings do not vary, employees still have an incentive to exert effort and the firm can control that incentive through the penalties it imposes on the supervisor. Third, making supervisors financially responsible for their ratings is costly if it simply transfers moral hazard from the employee to the supervisor. However, because the employee is paid directly and the slope of the supervisor’s contract is merely used to penalize over-reporting, we are able to separate the tension between allocating risk and inducing effort. In particular, when making a report on the equilibrium path the supervisor’s contract need not depend upon output, thus preventing risk from being transferred to the supervisor. It is sufficient that the supervisor prefers to make a different report unless the employee offers compensation.

Our main claims are that (1) subjective rating systems can implement efficient social transfers, and (2) they can do this while costlessly resolving moral hazard. We bolster these claims in two ways. First, we show that firms could also solve the moral hazard problems with collusion-proof contracts. The fact that such contracts are not used is consistent with our contention that some of the collusive activity benefits the firm.1

Secondly, we describe some empirical work on social transfers from ratees to raters. We

---

1 Although unobserved monetary transfers may also occur, Tirole notes that their use is more limited as they may yield verifiable evidence of dishonesty. This may be more true in some countries than in others.
identify four types of such transfers and perform initial measurements of their costs and values. Our study suggests that such transfers can create substantial value.

The paper contributes to three closely related streams of literature. Most importantly, it speaks to the literature on influence activities inside the firm. Tirole (1986, 1990) highlighted the pervasion of social transfers within firms. After citing widely from the sociology literature, he offers the example of a foreman who manipulates workers’ performance appraisals in return for the workers “refraining from activities such as unrest, going on strike, [and] leapfrogging for complaints.” Consistent with our arguments in the current paper, forbearing from each of these activities represents a cost to the worker and a benefit to the foreman. The literature on influence activities has been heavily influenced by Milgrom and Roberts (1990) who focus on the fact that these activities may be costly to the firm. Indeed the term “influence costs” hints at the inefficiency of intrafirm transfers. In the present paper we empirically document that there may be gains from trade between employees and we show that these can be realized at no direct cost to the firm.

The paper which is closest to ours in spirit is the recent piece by Prendergast and Topel (1996). They focus on the supervisor’s taste for power; the utility from being able to favor one employee over another. In their paper, the distortions caused by favoritism may be outweighed by the firm’s ability to pay a lower salary to the supervisor. Our results differ in two ways. First, we focus on transfers to the supervisor from the employees, rather than from the firm. Second, in our paper only the employees derive benefits from distorted performance ratings. In contrast, their results depend on supervisor’s deriving utility directly from the ratings given to employees.

We also contribute to the literature on the low variance in performance ratings. In our model all employees receive the same nominal rating and salary but those who (out-of-equilibrium) supply less effort will have to compensate with higher social transfers. In spite of the identical ratings, net compensation (social transfers and monetary remuneration)

---

2 We acknowledge that one could raise moral issues about these practices. We do not take a stand on those.
does vary with effort and the rating system can resolve moral hazard.\textsuperscript{3} This holds even if the social transfers are mere redistributions, as when employees pay money for higher ratings.

The final contribution is to the general literature on collusion, pioneered by Tirole. This literature has previously recognized that side-contracts need not be costly and indeed may be beneficial if they enable privately informed employees to write risk-sharing contracts.\textsuperscript{4} While our results do not depend upon the supervisor or employee sharing risk, we make two contributions to this literature: demonstrating that there may be other gains from trade between agents, and identifying a collusion-proof contract that prevents collusion costlessly.

We go beyond a theoretical demonstration of the effects of social transfers and try to measure the importance of the phenomenon. To this end we administer a questionnaire to one hundred and twenty-one managers, asking them to trade off salary against the giving and taking of social transfers. To the extent that our data are reliable and valid, our findings suggest that these transfers may be creating substantial value, between $10,000 and $15,000 per manager in our sample.

The paper proceeds in Section II with a formal statement of our results. This is followed in Section III with a description of the design and findings in our empirical study, while Section IV contains discussion of our conclusions and contributions.

\textsuperscript{3} Interestingly, Prendergast and Topel's model also explains the low variance in performance ratings. It is an empirical question which is the more compelling explanation.

\textsuperscript{4} See for example Varian (1990) and Itoh (1993).
We consider a three-person game between a principal (P), a worker (A), and a supervisor (B). A exerts effort $a \in [0, \bar{a}]$, and the firm's expected output is $a$. P only observes $Q = a + e$, where $e$ has distribution $F$, with zero mean. A is paid $\alpha(r)$ based on the rating $r \in \mathbb{R}$ given by B, and B is paid $\beta(Q, r)$. A and B are risk-averse, and their utility functions are given by $U(x, a)$ for A and $V(y)$ for B, where $x$ and $y$ are monetary.

There is no adverse selection, and all functions are common knowledge. They are also assumed to be $C^3$, and we confine attention to situations where expected utilities are strictly concave functions of decision variables, $\frac{\partial V}{\partial y}$ is positive and bounded away from zero and $\beta$ is concave in $r$. So that a higher rating indicates higher effort, we also assume that $\frac{\partial^2 \beta}{\partial Q \partial r}$ is positive and bounded away from zero. These are neither strong nor unusual assumptions. We will finally normalize the participation constraints to zero expected utility and focus on contracts for which these constraints bind.

We define two games on this environment. In one, $(G^0)$, A and B cannot collude, and in the other $(G^1)$, they can. The sequence of events in $G^0$ is as follows:

1. $P$ announces $\alpha^0$ and $\beta^0$,
2. $A$ quits or selects $a$,
3. $B$ observes $a$,
4. $B$ quits or selects $r$,
5. $Q$ is realized and payments are made.

Let us pick a pair of contracts $(\alpha^0, \beta^0)$. By assumption we know that they implement participation and a unique $(\alpha^0, r^0)$ through first-order conditions. The content of this assumption is, among other things, that $\beta^0$ depends on $r$ in a nontrivial way. One could imagine considerations for fairness and believability imposing long-term costs of too high
or too low values of \( r \). We think of \( \beta^0 \) as including these. Given this, the first-order conditions in stage 4 are

\[
\frac{\partial}{\partial r} \int V(\beta^0[Q,r])dF = 0 ,
\]

This defines the truthful reporting function \( r^*(a) \), a \( C^2 \) function. We can find \( a \) from

\[
\frac{\partial}{\partial a} U(\alpha^0[r(a)],a) = 0.
\]

Define \( a^* \) as the profit-maximizing effort level that maximizes: \( a - x - y \); such that \( U(x^*,a^*) = 0 \) and \( V(y^*) = 0 \). Note that finding a pair of contracts \((\alpha,\beta)\) to implement \( a^*, x^*, \) and \( y^* \) is straight-forward under \( G^0 \) where the supervisor either reports truthfully or not at all. However, as long as \( \partial \alpha / \partial r \neq 0 \), \( A \) and \( B \) have incentives to collude. We investigate this possibility in game \( G^1 \).

We anticipate side contracts between \( A \) and \( B \) in which \( A \) agrees to compensate \( B \) for inflating \( B \)'s report of \( A \)'s effort. We will use \( s \) to denote \( B \)'s monetary equivalent valuation of the transfer, and \( t(s) \) to denote \( A \)'s monetary equivalent cost of supplying \( s \). It is assumed that \( t(s) \) is \( C^3 \), increasing and convex, and that the efficient transfer, \( s^* \), given by \( \partial t / \partial s = 1 \), is positive. The idea in this formulation is that \( A \) first will supply the most efficient transfers, and go to less and less efficient means as necessary. To model the contracting process, we assume that \( A \) makes a take-it-or-leave-it offer \((\hat{r},s)\) to \( B \) and that this sidecontract is enforceable. So the sequence of events is:
1. $P$ announces $\alpha^1$ and $\beta^1$.
2. $A$ quits or selects $a$,
3. $B$ observes $a$,
4. $A$ makes an offer $(\hat{r},s)$ to $B$
5. $B$ quits, takes the offer and selects $\hat{r}$, or refuses the offer and selects $r^-$,
6. $Q$ is realized and payments are made.

For given $\alpha^1, \beta^1, a^0$ we can find the smallest $s$ for which $B$ will accept the sidecontract $(\hat{r},s)$ from

$$\int V(\beta^1[a^0 + e, \hat{r}]+s) dF - \int V(\beta^1[a^0 + e, r^-]) dF = 0, \quad (3)$$

where $r^-$ is given by

$$\frac{\partial}{\partial r} \int V(\beta^1[a^0 + e, r]) dF = 0. \quad (4)$$

Note that Equation 3 captures the dependence between the costs incurred by the supervisor to inflate the employee's performance rating and the compensation paid to the supervisor by the employee. This induces $s(a^0, \hat{r})$ as a $C^2$ function, and we find $a^0, \hat{r}$ from

$$\frac{\partial}{\partial a} U(\alpha^1[\hat{r}] - t[s(a, \hat{r})], a) = 0, \quad (5)$$

$$\frac{\partial}{\partial r} U(\alpha^1[r] - t[s(a^0, r)], a^0) = 0. \quad (6)$$

Our main result is that the firm can implement $a^*, s^*$ in $G^1$ with a contract in which the employee receives the maximum performance rating for a wide range of effort levels.

---

5We assume that the second order conditions hold.
This implementation involves no transfer of rent or allocation of risk to the employee or supervisor. Because the employee's cost $t(s^*)$ is less than the supervisor's valuation $s^*$, the principal is better off in $G^1$ than in $G^0$.

**Proposition 1:** There exists $(\alpha, \beta)$ that implement $a^*, s^*$ in $G^1$ at lower expected cost than in $G^0$.

**Proof:** We try $\alpha(r) = r + \theta_\alpha$, and $\beta(Q, r) = \phi Q (r - \tau) - \lambda r + \theta_\beta$, $r \in [\gamma, \tau]$, $(\lambda, \phi, \tau) \in \mathbb{R}_+^3$.

If $r = \tau$, $B$ bears no risk. So we want to implement $r^0 = \tau$, $a = a^*$, and $s = s^*$. We find $s$ from

$$\int V\left[\phi Q (r^0 - \tau) - \lambda r^0 + s + s\right] dF = 0$$  (3')

and note that $r^0 = \gamma$ if, when $r^0 = \gamma$

$$\int \frac{\partial V}{\partial y} (\phi Q - \lambda) dF < 0, \forall a.$$  (4')

Assuming for the moment that this holds, we use the implicit function theorem on (3') to get

$$\frac{\partial s}{\partial r} = \lambda - \phi \int \frac{\partial V}{\partial y} Q dF \left(\int \frac{\partial V}{\partial y} dF\right)^{-1}, \forall a.$$  (7)

From (6) we see that $r^0 = \tau$ if, when $r^0 = \tau$

These functions are clearly not unique. With more parameters one can also implement a desired rating.
\[
\frac{\partial U}{\partial x} \left( 1 - \frac{\partial t}{\partial s} \frac{\partial s}{\partial r} \right) > 0. 
\] (6')

This is true if, when \( r^0 = \tau \)

\[ 1+ \frac{\partial t}{\partial s}(\phi a - \lambda) > 0, \forall a. \] (6'')

Given the bounds on \( a \), there exists a maximum sidepayment \( \bar{s} \) which is payable when \( a = 0 \) and \( r^0 = \tau \). This implies an upper bound on \( \partial t / \partial s \), which we will denote \( \bar{t}_s \). We know that (6'') is satisfied if

\[ 1+ \bar{t}_s(\phi a - \lambda) > 0. \] (6''')

Comparing (4'') and (6''') we see that we can find pairs \((\phi, \lambda)\) such that both inequalities hold. Selecting a particular function \( \lambda(\phi) \) which does this, we can again use the implicit function theorem on (3') to get

\[ \frac{\partial s}{\partial a} = \phi(\gamma - \tau) \int \frac{\partial V}{\partial y} dF \left( \frac{\partial V}{\partial y} \right)^{-1}. \] (8)

Finally, we use \( \theta_a \) and \( \theta_\beta \) to keep participation constraints binding. Recall that we find \( a \) from (5):

\[ \frac{\partial U}{\partial x} \left( \frac{\partial t}{\partial s} \frac{\partial s}{\partial a} + \frac{\partial U}{\partial a} \right) = 0. \] (5')

Since \( \partial t / \partial s(s^*) = 1 \), (5') gives us \( a^* \) if, at \( a = a^* \), \( s = s^* \).
\[
\int \frac{\partial V}{\partial y} dF(\gamma - \tau) = \frac{\partial U}{\partial a} \left( \frac{\partial U}{\partial x} \right)^{-1} \frac{\partial V}{\partial y}, \quad (5'')
\]

while (3') gives us \( s^* \) if, at \( a = a^*, s = s^* \)

\[
V(-\bar{\lambda} \tau + s^* + \theta_\beta) = \int V(\phi Q[\gamma - \tau] - \bar{\lambda} \gamma + \theta_\beta) dF. \quad (3'')
\]

Note that the use of \( \theta_\alpha, \theta_\beta \) to keep the participation constraints binding imply that the LHS of (3'') and the RHS of (5'') are independent of \( (\lambda, \tau, \gamma, \phi) \). We fix \( \phi(\gamma, \tau) \) to satisfy (3'') and vary \( \gamma \) or \( \tau \) to satisfy (5'').

\[\text{---Q.E.D.}\]

The role of the social transfer in providing incentives for the employee to exert effort are highlighted in Equation (5'). When choosing whether to increase effort (increase \( a \)) the employee trades off increased disutility of effort with a reduction in the social transfer required to compensate the supervisor for setting \( r^0 = \tau \). The firm controls that incentive through the penalty it imposes in the supervisor’s contract for over-stating effort.

For a variety of reasons, this result should not be interpreted as a profitable way of resolving moral hazard. First, the supervisor is not free and the cost of supervision, less the value created by the influence transfer, is the net cost of addressing moral hazard. Second, in many cases the supervisor performs productive as well as monitoring functions. Our results do not help resolve the moral hazard problem associated with the supervisor’s productive actions.\(^7\) Third, our example suggests that the upper bound on the rating scale depends on the utility and production functions of the employee. If for reasons outside the

\[^7\text{Indeed our solution may not be implemented if the impact of the supervisor's productive actions on output cannot be separated from the impact of the employee's actions on output. We look at some aspects of such a model in Hauser, Simester, and Wernerfelt (1997).}\]
model the same rating scale is used for several employees with different utility or productive functions, \( a^* \) will not be implemented for all of them—although they may all receive the highest rating, depending on \((6'')\).

Note that the employee’s moral hazard problem becomes what appears to be an adverse selection problem for the supervisor. Adverse selection generally results in the principal transferring rent to the supervisor in high productivity states in order to ensure participation in low productivity states. However, this does not arise in the present case, because the productivity state is determined not by an act of nature, but rather by the level of effort exerted by the employee. By assuming in the proof that side-contracting on the employee’s effort level is not possible, our design of the employee’s wage schedule ensures that the productivity state follows the equilibrium path. We then get the result under collusion with Proposition 1. As a result, neither the principal nor the supervisor is concerned with the possibility of rent transfers to or from the supervisor at productivity levels off the equilibrium path.

The scheme used to prove Proposition 1 has many properties in common with schemes used in practice. In equilibrium the employee receives the highest rating. If the contract is offered to several employees with slightly different \( U \)'s, they will still all get a rating of \( \tau \) as long as \((6'')\) holds. Off-the-equilibrium path, if an employee chooses a lower than equilibrium effort level, he or she must compensate the supervisor with higher social transfers (Equation 8). In summary, the scheme implements efficient influence transfers and may do so in a situation where several employees are rated the same and get paid the same by the firm, but where nevertheless the more productive employees get higher net compensation.

It may seem far-fetched to some that the principal takes advantage of collusion between employee and supervisor. We will measure the possible gains from this in Section III, but first we push our point theoretically by showing that collusion is avoidable in the sense that the principal could design collusion-proof schemes. Since these schemes do not
look like anything observed in practice, this further supports our contention about the beneficial nature of the transfers. Specifically, we demonstrate the feasibility of designing a contract for the supervisor that ensures complete information revelation, without exposing the supervisor to risk, or transferring rent to the supervisor when the employee has performed favorably. We do so by discretizing the supervisor's contract and ensuring that any distortions from collusion are arbitrarily small.

We consider two types of distortions: extortion and bribery. Extortion may occur if the supervisor can credibly threaten to under-report the performance of the employee. In return for forbearance from under-reporting, the supervisor may demand a portion of the additional salary that an honest report would pay the upstream employee. We overcome opportunities for extortion by ensuring that threats to under-report are imperfect. The employee can ignore the supervisor's demand, knowing that the supervisor acting in his or her own best interests will prefer not to under-report. Bribery may arise in the form of a side-contract under which the supervisor over-reports the employee's effort level, in return for a share of the additional compensation that results. In Proposition 1, the ratings inflation resulting from bribery is anticipated, and the participation and incentive compatibility constraints are adjusted to account for this distortion. In Proposition 2, bribery is averted by making selection of a misleading contract too costly to the supervisor. The maximum bribe that the employee would be willing to pay, in order to inflate reports of his effort level, is equal to the increase in compensation that the duplicity would yield. We show that it is possible to design contracts for which the potential gains from inflation are less than the cost to the supervisor of over-reporting. We state this result formally as Proposition 2, where \( r^* \) is a positive integer:

**Proposition 2:** If the supervisor's risk aversion does not change too rapidly as the supervisors' expected wealth changes, the principal can come arbitrarily close to implementing \( (a^0, r^*) \) in \( G^1 \) at the same expected cost as in \( G^0 \).
Proof: See Appendix.

In the proof, the supervisor's contract is discretized to comprise a finite menu of linear contracts. When the employee's effort level is higher, the supervisor prefers to choose a contract with a larger variable component and a smaller fixed component. The menu exploits this relationship to ensure that the loss from selecting an inappropriate contract exceeds any potential gains from bribery, while threats to under-report are imperfect. We then rotate the menu in order to make the slope of the equilibrium contract zero, so that in equilibrium the supervisor bears no risk. Indeed in equilibrium the supervisor does nothing. It is the anticipated reaction when the employee deviates from the equilibrium path and the subsequent punishment imposed on the employee that deter deviations.

Proposition 1 shows that influence transfers within the firm may be efficient, while Proposition 2 shows that firms could avoid these transfers. Since the rating schemes we see in practice do not have this collusion-proof property, one could take this as evidence that firms choose to allow collusion. We now proceed with an empirical study which suggests, at least in some cases, that efficient transfers may be possible.
3. Exploratory Empirical Study

To estimate the magnitude of the costs and values of influence transfers we undertook an initial empirical study using "conjoint analysis" (Luce and Tukey, 1964; Green and Rao, 1971; Green and Srinivasan, 1978). Conjoint analysis is a methodology for estimating utility functions based on hypothetical choices and ratings and is the quantitative technique most widely used by market research firms (Wittink and Cattin, 1989; Mahajan and Wind, 1992). The data come from subjects who are asked to rank or rate a number of alternatives, each of which is described by one of several levels on each of several attributes. For example, ["Car with 100 HP at $15,000, "Car with 150 HP at $20,000, "Car with 100 HP at $20,000", "Car with 150 HP at $15,000"]. After gathering this data, one uses a statistical technique such as regression analysis to estimate that multi-attribute utility function which best predicts the rankings or ratings. Naturally the use of the method depends on the scale properties of the data and the assumed error structures (Hauser and Shugan, 1980).

A critical first step is to identify the attributes, in this case the possible types of influence transfers. To do so we conducted one-and-a-half-hour interviews with five senior R&D executives. The interviews were taped, transcribed, and analyzed by the three authors. This process generated several (20-25) candidate attributes which were pilot tested in an MBA class. In the pilot test we tried to identify the most important attributes and experimented with alternative descriptions of each attribute. After making some difficult qualitative judgments, we settled on four types of social transfers which were valued by supervisors: having the respect of the employees, having employees forego interesting assignments to work for the supervisor, having the employees learn special

---

While they all worked in R&D, their backgrounds were quite different. One had twenty-five years of experience in aerospace. Another was a long-time employee of an independent research lab, dealing mostly with government contracts. A third worked primarily on the marketing side of R&D. A fourth had spent his career working for several very small computer companies in Silicon Valley and the last worked for a very large German company. All had very extensive experience with subjective performance ratings.
skills to work for the supervisor, and having employees spend time learning about the supervisor’s expectations and objectives. Contrary to our expectations, more overtly obsequious behaviour appeared to be valued less.

To measure the value and cost of these activities we used a salary attribute, and we also added a vacation time attribute. This serves both as a control measure and as a way to gauge “demand effects” (in which subjects value attributes highly because they feel that that is what is desired by the experimenter [Sawyer, 1975]).

We then designed eight hypothetical jobs for raters; for example, one such job was

A salary $3000 per year more than your current.

One week per year more vacation than your current.

You are not well-respected by the employees that work for you.

Employees will forego other interesting assignments just so they can work for you.

Employees will not spend a lot of time learning special skills just so they can work for you.

Employees have a good understanding of your expectations and objectives.

Each of the six attributes was either high or low in each job and the eight jobs were designed to make variation in the attributes orthogonal. A subject assigned to a rater role had to score all eight jobs on a 0-100 scale, with 0 for the worst and 100 for the best.

Subjects assigned to ratee roles were also given descriptions of eight jobs, but with the ratee side of the influence transfer attributes. For example

The same salary as your current.

The same vacation as your current.

You often have to spend time convincing managers that you respect them.

You seldom have to forego interesting assignments just to show managers that you want to work for them.
You often have to learn special skills just to persuade managers that you want to work for them.

You seldom have to spend time finding out about a manager's expectations and objectives.

The questionnaire was administered to 121 working executives enrolled in a weekend MBA program. The subjects’ average age was twenty-nine years, their average salary was $62,000 per year, and their average vacation time was three weeks per year. Consistent with our beliefs about the prevalence of performance rating systems, 93% of the subjects indicated that they received a rating on a twelve-, six-, or three-month basis. Thirty percent of the subjects were currently rating others.

For each group, say the sixty-one subjects who scored the rater jobs, we regressed the attribute levels (six dummies) on the scores. This allows us to compare the impact of $3,000 more in salary with low to high change in one of the other attributes. The results of this comparison are given in Table 1.

### Table 1

**Value of Influence Transfers**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Raters</th>
<th>Ratees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respect</td>
<td>$13,664</td>
<td>$2,859</td>
</tr>
<tr>
<td>Expectations</td>
<td>4,789</td>
<td>633</td>
</tr>
<tr>
<td>Special skills</td>
<td>2,190</td>
<td>1,295</td>
</tr>
<tr>
<td>Forego assignments</td>
<td>2,633</td>
<td>5,700</td>
</tr>
<tr>
<td>Sample size</td>
<td>488 (61 people)</td>
<td>480 (60 people)</td>
</tr>
</tbody>
</table>

We focus on two findings. First, the average value of influence transfers has a significant magnitude, summing up to almost $20,000 in our data. Second, different classes of influence transfers differ in their efficiency. Beyond these main conclusions, the results

---

9For specific managers in our subject pool, the value may be much higher.
should be interpreted carefully. Let us start with the bad news. First, the data are noisy. This is best seen in the value-of-vacation, which should be the same for both groups, but is $2437 per week for those given the hypothetical role of raters and $1214 per week for those given the role of ratees. The result is partially driven by a small group of extreme outliers, but there is a broad tendency for the subjects scoring the ratee cards to value vacation less. Second, in most companies a single rater will rate several ratees. With one ratee, three of the four attributes can serve as bases for efficient transfers, but with five ratees incurring the cost of the transfer and only one supervisor enjoying the benefits, only one attribute can be the basis for efficient transfers. This leads us to the third point which begins the good news: it should be remembered that the model is linear and that the line is estimated between a pair of low/high points. Suppose we make the reasonable assumption that the values are concave functions, and costs are convex functions on a per attribute basis. In that case the efficient transfers may not be “high” on each attribute, but some intermediate level. So by focusing on the high points, our study understates the true gains from social transfers. Fourth, for ease of measurement, we concentrated on a subset of the attributes and omitted several important attributes. This causes us to underestimate the amount of efficient influence transfer. Based on qualitative post test input from subjects, we also feel that our forego attribute is mis-specified and most likely picks up a combination of other attributes. The attribute we sought was a description by the raters that they valued getting high quality employees to work on their projects. If they do not “take care of their people” with high ratings, they found that employees were “busy” when it came time to staff one of their projects. This is very salient for the R&D managers from whom we got the attributes. However, our phrasing did not capture the construct well.
4. Discussion

We have argued that performance rating systems may help firms implement desirable social transfers between employees and supervisors, and that this does not interfere with the control of moral hazard. The argument has three components. First, we show that rating schemes with oft-observed properties can implement efficient social transfers while still controlling moral hazard. Secondly, we show that the schemes could have been constructed to avoid these transfers. Thirdly, we present empirical evidence to the effect that the value of these transfers is significant.

In the process of making our argument, we also contribute to an explanation of the low variance in nominal performance ratings. Following our model, employees who perform poorly are penalized by having to make larger social transfers, rather than by lower nominal salaries.

Our theoretical results depend upon two critical assumptions. The first assumption is that the rating system is used solely to control moral hazard and not adverse selection. In both of our results, control of the equilibrium path is important to avoid transferring risk or rent to the supervisor. If employees vary in their productivity or disutility of effort, the contracts we have proposed will need to be tailored to each employee. These contracts could be designed to yield different effort levels and/or ratings for each employee. If the firm is not fully informed about employee heterogeneity, then adverse selection problems arise not just in tailoring the incentive contracts but also in making promotion and dismissal decisions. Resolving these adverse selection problems suggests recourse to a screening mechanism. Indeed, subjective performance ratings may have a separate role in promotion and dismissal decisions by providing information to resolve adverse selection. Simplicity would argue for designing a screening mechanism that depended upon a second set of ratings, however, it is possible to imagine screening mechanisms that rely upon the same ratings. Hence, while adverse selection may hinder the design of incentive contracts (to
resolve moral hazard) the incentive contracts themselves need not hinder resolution of adverse selection.

The second critical assumption is that the employee and supervisor cannot side-contract upon the employee's effort level. Control of the equilibrium path (through the efforts of the employee) is necessary to prevent exposing the supervisor to risk or enabling the supervisor to capture rents. When the supervisor and employee are able to contract on the employee's effort, control of the equilibrium path may be lost. Although restricting side-contracting on the employee's effort level is common in the literature, some readers may find this troublesome because we anticipate side-contracting to inflate supervisors' reports. Enforceability of side-contracts generally pre-supposes repeated interaction, the indivisibility of which argues for either all side contracts being enforceable or none being so. In response, we make two observations. First, it is generally much harder to specify effort over a period than to specify a single rating. Second, enforceability depends upon more than just repeated interaction. For example, it is possible that the supervisor's report is observed throughout the firm, but the employee's choice of effort is observed by just the employee and the supervisor. Hence, while the threat of censure by colleagues may be sufficient to enforce an agreement on the supervisor's report it may not be available to enforce a contract on the employee's level of effort.

In terms of future theoretical research it would be interesting to look at the role of influence transfers in situations where economic choices implicitly serve as ratings. For example, when a downstream profit-center manager buys intermediate goods from an upstream profit-center manager, the amount purchased implicitly rates the performance of the latter. One could imagine that such situations also lead to social transfers. In the long run it is possible that this could add to our understanding of aspects of transactions between firms.

In terms of future empirical research, it should be clear that our study has limitations and could be refined substantially, both in terms of the measures and the
sample. More generally, however, one should be able to identify both efficient and inefficient transfers and test whether the former are more likely to take place.
Appendix

We prove our second result that, under weak conditions, the principal can prevent collusion and, with the supervisor truthfully revealing the employee's effort \((r = r^*)\), come arbitrarily close to implementing any \(a^0\). The approximation involves discretizing the set of possible effort levels. We label the discrete actions in increasing order by \(\{a_1, a_2, \ldots, a_n\} \subset \mathbb{R}\), and with a slight abuse of notation we will use \(a_i\) to denote both the point \(a_i\) and the interval \([a_i, a_{i+1})\).\(^{10}\) Moreover, let \(a_j\) define the interval that includes \(a^0\). It is useful to first prove the following Lemma.

**Lemma:** If \(B\) has constant absolute risk aversion and \(F\) is normal, then the principal can dominant-strategy-implement \((a^0, r^*)\) in \(G^i\) at the same expected cost as in \(G^0\).

**Proof:** Consider a linear reward scheme \(L_j(\epsilon, \alpha, \beta)\) in which \(B\) rates \(A\) by picking from a family of functions \(\beta^i(Q) = \eta^i Q + \omega^i, \; \eta^i \in \mathbb{R}, \; \omega^i \in \mathbb{R}, \; i = 1, 2, \ldots, n\), and \(A\) is paid \(\alpha^i\) if \(B\) picks \(\beta^i\), where

\[
\begin{align*}
U(\alpha^i + \epsilon, a_i) &= 0, \; i \neq j, \\
U(\alpha^i, a_j) &= 0.
\end{align*}
\]  

(12)

The largest sidepayment \(s_i^+\) which \(A\) can offer \(B\) to report implicitly \(a_k\) when \(a_i\) was chosen, is given by

\[
t(s_i^+) = \alpha^+ - \alpha^i,
\]

(13)

and the certainty equivalent of \(B\)'s utility in this case is

\(^{10}\)It will become obvious that \(A\) never will select \(a \in (a_i, a_{i+1})\) in equilibrium, so a set of \(n\) ratings can fully reveal efforts. In practice, most firms use such discrete rating scales (e.g. Collier, 1977).
\[ \eta^i a_i + \omega^i + s_i^i - \frac{1}{2} \rho \eta^i \sigma^2, \quad (14) \]

where \( \rho \) is the coefficient of absolute risk aversion and \( \sigma^2 \) is the variance of \( F \). In order for "truth-telling" to be a dominant strategy, \( L_j \) must necessarily satisfy the following conditions.

\[ \eta^i a_i + \omega^i - \frac{1}{2} \rho \eta^{(i^+)} \sigma^2 > \eta^{(i^+)} a_i + \omega^{(i^+)} + s^{(i^+)} - \frac{1}{2} \rho \eta^{(i^+)} \sigma^2 \quad (15) \]

\( (A \) cannot pay \( B \) to report \( a_{i+m} \) when \( a_i \) was chosen), and

\[ \eta^{i^+} a_{i+1} + \omega^{i^+} - \frac{1}{2} \rho \eta^{(i^+)} \sigma^2 > \eta^{(i^+)} a_{i+1} + \omega^{(i^+)} - \frac{1}{2} \rho \eta^{(i^+)} \sigma^2 \quad (16) \]

\( (B \) prefers \( \beta^{i+m} \) over \( \beta^i \) when \( a_{i+m} \) is chosen). We look first at \( m = 1 \) and turn (15), (16) into equalities by adding positive constants on the right sides. For simplicity, we add the same \( c > 0 \) to both. We can then solve for the recursive relationships

\[ \eta^{i+1} = \eta^i + \frac{2c + s^{i+1}}{a_{i+1} - a_i}, \quad (17) \]

\[ \omega^{i+1} = \omega^i - \frac{2c + s^{i+1}}{a_{i+1} - a_i} + \frac{1}{2} \rho \left( \eta^{(i^+)} - \eta^i \right) \sigma^2 + c. \quad (18) \]

To ensure that \( B \) carries no risk, we set \( \eta^i = 0 \). From this starting point we can solve (17), (18) for all \( i \).
We now need to show that the series thus constructed will satisfy (15) for \( m > 1 \).

To this end, we use (17), (18) to get

\[
\omega^{i+m} = \omega^i - \sum_{t=i}^{i+m-1} a_t \frac{2c + s^{i+1}_t}{a_{t+1} - a_t} - mc - s^{i+m} + \frac{1}{2} \rho \sigma^2 \left( \eta^{(i+m)^2} - \eta^{i^2} \right),
\]

(19)

\[
\eta^{i+m} = \eta^i + \sum_{t=i}^{i+m-1} \frac{2c + s^{i+1}_t}{a_{t+1} - a_t}.
\]

Substituting into (15) gives

\[
mc + \sum_{t=i}^{i+m-1} \frac{2c + s^{i+1}_t}{a_{t+1} - a_t} (a_t - a_i) > 0.
\]

(21)

which is true since \( a \) is increasing.

A similar exercise allows us to write (16) for \( m > 1 \) as

\[
(m+1)c + \sum_{t=i-m}^{i-1} \frac{2c + s^{i+1}_t}{a_{t+1} - a_t} (a_{t+1} - a_{t+1}) \geq 0.
\]

(22)

So \( L_j(e, \alpha, \beta) \) with \( \alpha, \beta \) given by (17), (18) has truth telling as a dominant strategy for \( B \).

Given truth telling, \( \alpha \) is constructed to make \( a^0 \) a dominant strategy choice for \( A \). Since

\( \eta^i = 0 \), \( B \) bears no risk in equilibrium and we have implemented \( a_j \) without risk costs.

—Q.E.D.

This Lemma depends on \( B \) having constant absolute risk aversion and \( F \) being normal. However, under much weaker assumptions, the result holds approximately.

Suppose that the third derivative of \( B \)'s utility function \( V \) is continuous and bounded and that the third absolute central moment of \( F \) is of smaller order than \( F \)'s variance \( \sigma^2 \). In this
case we can use a Taylor series expansion (Pratt, 1964) to approximate the analog of (14) for small values of \( \sigma^2 \):

\[
\eta^k a_i + \omega^k + s^k - \frac{1}{2} \eta^k \sigma^2 \frac{V_{11}}{V_1},
\]

where \((V_1, V_{11})\) are the first and second derivatives of \(V\). For notational convenience we define \( \rho_i^k \) as the negative of absolute risk aversion

\[
\rho_i^k \equiv \frac{V_{11}}{V_1} \eta^k a_i + \omega^k,
\]

where \((V_1, V_{11})\) are the first and second derivatives of \(V\). For notational convenience we define \( \rho_i^k \) as the negative of absolute risk aversion

\[
\rho_i^k \equiv \frac{V_{11}}{V_1} \eta^k a_i + \omega^k,
\]

\[
\rho_{i,s}^k \equiv \frac{V_{11}}{V_1} \eta^k a_i + \omega^k.
\]

Substitution yields the following sufficient conditions for (21) and (22).

\[
\sum_{i,m+1}^i \left[ \left( \eta^{i+1} - \eta^i \right)(a_i - a_i) + \frac{1}{2} \eta^{i+1} \sigma^2 (\rho_i^i - \rho_{i-1}^i) \right] + (m+1)c + \frac{1}{2} \eta^{i+1} \sigma^2 (\rho_i^{i+1} - \rho_{i+1}^{i+1}) \geq 0,
\]

\[
\sum_{i,m+1}^i \left[ \left( \eta^{i+1} - \eta^i \right)(a_i - a_i) + \frac{1}{2} \eta^{i+1} \sigma^2 (\rho_i^i - \rho_{i+1}^i) \right] + (m+1)c + \frac{1}{2} \eta^{i+1} \sigma^2 (\rho_i^{i+1} - \rho_{i+1}^{i+1}) \geq 0.
\]

Everything else equal, (21')-(22') will hold if \( \frac{V_{11}}{V_1} \) does not change too rapidly. Because we can discretize \(a\) arbitrarily finely, we have shown that if \( \frac{V_{11}V_1 - V_{11}^2}{V_1^2} \) are small in the sense defined by (21') and (22'), then the principal can prevent collusion and come arbitrarily close to implementing \(a^0\).

**Proposition 2:** If the supervisor’s risk aversion does not change too rapidly as the supervisors’ expected wealth changes, the principal can come arbitrarily close to implementing \((a^0, r^*)\) in \(G^1\) at the same expected cost as in \(G^0\).
References


Appendix: Proofs

Following the text we use a hat (^) to indicate the rating ($\hat{r}$), D's actions ($\hat{d}$), and the incremental profits ($\hat{\pi}$) that result when D optimizes its expected utility in the absence of a side payment (second expression in Equation 2). This means that the downstream agent's reaction to $u$ implies three continuously differentiable functions, $\hat{r}(u)$, $d(u)$, and $\hat{\pi}(u)$, which tell us how D would react to U's choice of actions and how expected profit would be affected if there were no side payment. We use a tilde (\~) to indicate the rating ($\tilde{r}$), D's actions ($\tilde{d}$), and incremental profits ($\tilde{\pi}$) that result when side payments are allowed under a take-it-or-leave-it offer from U to D on a side-payment-for-rating contract, $(s, \tilde{r})$.

If D can do at least as well with the contract as without, then this implies Constraint (C1)

\[ EU_d[w(\tilde{r}, \tilde{\pi}) - c_d(\tilde{d}) + s(u, \tilde{r})] \geq EU_d[w(\hat{r}, \hat{\pi}) - c_d(\hat{d})] . \]

For a given $v$ and $w$, the upstream agent will select $u$ and $\tilde{r}$ to maximize its expected utility according to Constraint (C2).

\[ EU_u[\tilde{r}, u] = EU_u[v(\tilde{r}) - c_u(u) - s(u, \tilde{r})] . \]

To earn non-zero profits, the firm must retain its employees. Let $\tilde{U}_u$ and $\tilde{U}_d$ be the minimum expected utilities that U and D require to participate. Thus, to induce U and D to participate, the firm must set $v$ and $w$ such that:

\[ EU_u[v(\tilde{r}) - c_u(u) - s(u, \tilde{r})] \geq \tilde{U}_u , \]
\[ EU_d[w(\tilde{r}, \tilde{\pi}) - c_d(\tilde{d}) + s(u, \tilde{r})] \geq \tilde{U}_d . \]

Because the firm is maximizing expected profits it will try to keep wages as low as is feasible. Formally this means that the firm will attempt to select $v$ and $w$ such that the constraints are binding.

RESULT 1. For incentive systems in which the rewards to the upstream agent are increasing in the downstream agent's rating, the upstream agent will provide a positive side payment to the downstream agent unless the firm sets a binding upper bound on the rating (or otherwise precludes side payments). Even with a binding upper bound, there may be side payments.

Proof. Consider a given $u$. Let $F(e)$ be the distribution function for $e$. Consider first the case where there is no upper bound on $r$ or it is not binding. U will seek to maximize the expression in (C2) subject to the conditions imposed by D's maximization problems in (C1). Let $\Gamma_u = v - c_u - s$. Differentiating (C2) we obtain:

\[ \int \frac{\partial U_u}{\partial \Gamma_u} \frac{\partial [v(\tilde{r}) - c_u(u) - s(u, \tilde{r})]}{\partial \tilde{r}} dF = 0 . \]

By assumption, $\partial U_u / \partial \Gamma_u > 0$. The error, $e$, appears in $w$, but it does not appear in $v$, $c_u$, or $s$. Thus, this integral can be zero if and only if $\partial U_u[v - c_u - s] / \partial \tilde{r} = 0$. Thus, this first order condition holds if and only if:

\[ \frac{\partial v(\tilde{r})}{\partial \tilde{r}} - \frac{\partial s(u, \tilde{r})}{\partial \tilde{r}} = 0 . \]
Let $\Gamma = w - c + s$. We now use implicit differentiation on (C1) recognizing that the right-hand side (RHS) does not depend on $\bar{r}$.

\[ \frac{\partial U_d}{\partial \Gamma_d} \left[ \frac{\partial \Gamma_d}{\partial w} \frac{\partial w(\bar{r}, \pi)}{\partial \bar{r}} + \frac{\partial \Gamma_d}{\partial s} \frac{\partial s(u, \bar{r})}{\partial \bar{r}} \right] dF = 0. \]

By assumption $\frac{\partial U_d}{\partial \Gamma_d}$, which depends on $e$, is positive. Furthermore, $s(u, \bar{r})$ does not depend upon $e$, and $\frac{\partial \Gamma_d}{\partial w} = \frac{\partial \Gamma_d}{\partial s} = 1$. Hence, Equation (A2) becomes:

\[ \int \frac{\partial U_d}{\partial \Gamma_d} dF = -\frac{\partial s(u, \bar{r})}{\partial \bar{r}} \int \frac{\partial U_d}{\partial \Gamma_d} dF. \]

From (A1) we have $\frac{\partial s(r, u)}{\partial r} = \frac{\partial v(\bar{r})}{\partial \bar{r}}$. But $\frac{\partial v(\bar{r})}{\partial \bar{r}} > 0$ by assumption. Thus, the RHS of (A3) is negative:

\[ \int \frac{\partial U_d}{\partial \Gamma_d} dF < 0. \]

By similar arguments we use implicit differentiation on (C1) to obtain

\[ \int \frac{\partial U_d}{\partial \Gamma_d} dF = 0. \]

Finally, we differentiate the left-hand side (LHS) of (A4) or (A5) to demonstrate that the second derivative with respect to $r$ is negative (concave) because $\frac{\partial U_d}{\partial \Gamma_d} > 0$ and $\frac{\partial^2 w}{\partial \bar{r}^2} < 0$. Thus, we have shown that the first derivative of a concave function is negative at $\bar{r}$ and zero at $\hat{r}$. Hence, $\bar{r} > \hat{r}$.

The side payment is positive by (C1). In the case where $r$ is constrained, we add a Lagrange multiplier, $-\lambda (r - \hat{r})$, to $U$'s optimization problem. This might allow a solution of the form $\frac{\partial v}{\partial \bar{r}} = \lambda$ and $\frac{\partial s}{\partial \bar{r}} = 0$. If $u$ were limited to a finite set, we obtain a result similar to that for Lagrange multipliers by using a series of piecewise functions for $w$ such that each piece corresponds to a different action by $U$.

Finally, the reader can verify side payments for $v = v_0 + v_1 r$ and $w = w_0 + w_1 (1 - r) + w_3 r \pi$ for $s \in [0, 1]$ and sufficiently large $v_1$ and $w_1$. ■

**RESULT 2.** If the firm can preclude side payments and choose a reward system, $v$ and $w$, such that the upstream and downstream agents, acting in their own best interests, choose actions $u^0$ and $d^0$, then there exists a reward system such that (1) there is no loss of profits to the firm and (2) the upstream and downstream agents, still acting in their own best interests, choose $u^0$ and $d^0$ even though they are free to make side payments in return for higher ratings.

**Proof.** Let us first denote the game in which the firm precludes side payments by $G^0$ and the game in which side payments are allowed as $G^1$. Formally, we show that if $(v^0, w^0)$ implement $(u^0, d^0, r^0)$ in $G^0$, and the participation constraints bind, then there exist real numbers $\gamma, \Theta_u, \Theta_d$ such that $(v^1, w^1) = (0.5 \gamma^2 + \gamma + \Theta_u, w^0 - 0.5 \gamma^2 - \gamma + \Theta_d)$ leads to $U$ and $D$ choosing $(u^0, d^0, r^0)$ in $G^1$ at the same expected cost to the firm. We will assume that $w^0$ satisfies the same assumptions as $w$ and that utility is strictly increasing in monetary payments (with the slope bounded away from zero). Finally we also require that $\forall d, \pi$ is increasing in $u$ when $u = u^0$, and that the integrals over $f(e)$ are defined correctly. These are all rather weak (and natural) assumptions.

We first generalize the reward and cost functions for the upstream and downstream agents as:

$U_u = U_u(x, u)$ and $U_d = U_d(y, d)$ where the first argument in each function represents the money
transfers resulting from the side payments and wage payments and the second argument reflects the
disutility of effort. That is: \( x = v - s \) and \( y = w + s \).

Note that under the new reward system only D incurs risk and this risk is due to \( w^0 \) (as under the old reward system). Hence, the risk premium paid by the firm will be unchanged if we implement \((u^0, d^0, r^0)\) in \( G^1 \). Moreover, if we can implement \((u^0, d^0, r^0)\) in \( G^1 \), we can adjust \( \theta_u, \theta_d \) until both participation constraints bind. Therefore, we need only show that we can choose to implement \((u^0, d^0, r^0)\) in \( G^1 \) with \((0.5y^2 + \gamma + \theta_u, w^0 - 0.5y^2 - \gamma + \theta_d)\). To do so we first show that if we can implement \( u^0 \) and \( d^0 \) with \((0.5y^2 + \gamma + \theta_u, w^0 - 0.5y^2 - \gamma + \theta_d)\) then, when the participation constraints are binding, we will also implement \( r^0 \). Next we show that if we can implement \( u^0 \) and \( r^0 \) with \((0.5y^2 + \gamma + \theta_u, w^0 - 0.5y^2 - \gamma + \theta_d)\) then, when the participation constraints are binding, we will also implement \( d^0 \). Finally we show that we can choose a \( \gamma \) so that \((0.5y^2 + \gamma + \theta_u, w^0 - 0.5y^2 - \gamma + \theta_d)\) implements \( u^0 \) given \( d^0 \) and \( r^0 \).

Note that \( r^0 \) is determined by D in \( G^0 \) from the following first order condition:

\[
(A6) \quad \int \frac{\partial U_u}{\partial y} \left[ \frac{\partial w^0}{\partial \pi} \frac{\partial \pi}{\partial r} + \frac{\partial w^0}{\partial r} \right] dF = 0.
\]

Under \( G^1 \), \( r \) is determined by \( U \) maximizing \( \int U_u(v^1 - s, u) dF \), which yields the following first order condition:

\[
(A7) \quad \int \frac{\partial U_u}{\partial x} \left( \gamma - \frac{\partial s}{\partial r} \right) dF = 0.
\]

Using the implicit function theorem on (C1) we get:

\[
(A8) \quad \frac{\partial s}{\partial r} = \left( \int \frac{\partial U_u}{\partial y} \gamma dF - \int \frac{\partial U_u}{\partial y} \left[ \frac{\partial w^0}{\partial \pi} \frac{\partial \pi}{\partial r} + \frac{\partial w^0}{\partial r} \right] dF \right) \left( \int \frac{\partial U_u}{\partial y} dF \right)^{-1} = 0.
\]

Inserting this into (A7) yields:

\[
(A9) \quad \frac{\partial U_u}{\partial x} \left( \gamma - \gamma + \int \frac{\partial U_u}{\partial y} \left[ \frac{\partial w^0}{\partial \pi} \frac{\partial \pi}{\partial r} + \frac{\partial w^0}{\partial r} \right] dF \left[ \int \frac{\partial U_u}{\partial y} dF \right]^{-1} \right) = 0.
\]

This is zero if (A6) holds, since all derivatives are evaluated at their participation constraints. So if we can use \((0.5y^2 + \gamma + \theta_u, w^0 - 0.5y^2 - \gamma + \theta_d)\) to implement \( u^0 \) and \( d^0 \) in \( G^1 \), we will also implement \( r^0 \). Note that this result holds even if there is an upper bound on \( r \) that constrains \( r^0 \) in \( G^1 \). Although (A6) will not hold in those circumstances, because the LHS of (A9) equals the LHS of (A6), we know that the constraint will also be binding in \( G^1 \).

We turn now to implementation of \( d^0 \), and make the analogous argument that if \((0.5y^2 + \gamma + \theta_u, w^0 - 0.5y^2 - \gamma + \theta_d)\) implements \( u^0 \) and \( r^0 \) in \( G^1 \), we will also implement \( d^0 \). Note that \( d^0 \) is determined by D in \( G^0 \) from the following first order condition:

\[
(A10) \quad \int \frac{\partial U_d}{\partial y} \left( \frac{\partial w^0}{\partial \pi} \frac{\partial \pi}{\partial d} \right) dF + \int \frac{\partial U_d}{\partial d} dF = 0.
\]

Under \( G^1 \), \( d \) is determined by D maximizing \( \int U_d(w^1 + s, d) dF \), which yields the following first order condition:
This holds if (A10) holds since all derivatives are again evaluated at their participation constraints. So we now need only show that we can find a $\gamma$ that implements $u^0$ using 

$$\int \frac{\partial U_u}{\partial y} \left( \frac{\partial w^0}{\partial \pi} \frac{\partial \pi}{\partial d} \right) dF + \int \frac{\partial U_u}{\partial d} dF = 0.$$ 

Under these constraints, only the second \( \left( \frac{\partial w^0}{\partial \pi} \right) \left( \frac{\partial \pi}{\partial u} \right) \) term changes as we change $\gamma$. On the other hand (C3) tells us that the change is continuous. Suppose first that $\gamma = 0$. In this case, $s = 0$, $\tilde{r} = r^0$, and $d = d^0$, so the LHS of (A12) reduces to $\frac{\partial U_u}{\partial d} < 0$. Consider next what happens as $\gamma \to \infty$. Using the implicit function theorem on (C3), we see that $\frac{d\tilde{r}}{dy}$ is negative iff:

$$\int \frac{\partial^2 U_u}{\partial y^2} \left[ \frac{\partial w^0}{\partial \pi} \frac{\partial \pi}{\partial r} + \frac{\partial w^0}{\partial r} - \gamma \right] (\gamma + \tilde{r}) dF + \int \frac{\partial U_u}{\partial y} dF > 0. \tag{A13}$$

We would like to show that $\tilde{r} \to -\infty$ as $\gamma \to \infty$ (we assume that D’s threat point $\tilde{r}$ has no lower bound). Let us assume, to derive a contradiction, that this is not the case. This implies that (A13) is not satisfied, $(\gamma + \tilde{r}) > 0$ and $\frac{\partial w^0}{\partial \pi}$ is increasing as $\gamma \to \infty$. Note that because $w^0$ and $\pi$ are concave in $r$, $\frac{\partial w^0}{\partial r}$ and $\frac{\partial \pi}{\partial r}$ are non-increasing as $\gamma$ increases. Hence for sufficiently large $\gamma$ we know that $(\gamma + \tilde{r}) > 0$ and $\frac{\partial w^0}{\partial \pi} > 0$. Thus we have derived a contradiction because if 

$$\left( \frac{\partial w^0}{\partial \pi} \right) \left( \frac{\partial \pi}{\partial r} \right) + \frac{\partial w^0}{\partial r} - \gamma < 0.$$ 

Hence we know that $\tilde{r} \to -\infty$ as $\gamma \to \infty$. Since by assumption $\frac{\partial w^0}{\partial \pi} \to -\infty$ as $\tilde{r} \to -\infty$, this will eventually make the LHS of (A12) positive. So by continuity there exists a $\gamma$ such that (A12) holds and this $\gamma$ implements $u^0$.\hfill\(\blacksquare\)