Optimal Contracting with Subjective Evaluation

W. Bentley MacLeod*
Department of Economics and the Law School
The University of Southern California
Kaprilian Hall 300
3620 South Vermont Av.
Los Angeles, CA 90089-0253

August 2002

Abstract

This paper extends the standard principal-agent model to allow for subjective evaluation. The optimal contract results in more compressed pay relative to the case with verifiable performance measures. Moreover, discrimination against an individual implies lower pay and performance, suggesting that the extent of discrimination as measured after controlling for performance may under-estimate the level of true discrimination. Finally, the optimal contract entails the use of bonus pay rather than the threat of dismissal, hence neither "efficiency wages" nor the right to dismiss an employee are necessary ingredients for an optimal incentive contract. (JEL D800, J410, J700)

Like the parents in Garrison Keillor's Lake Woebegone, where all the children are above average, supervisors also have a tendency to judge their workers as above average, resulting in performance evaluations that are more compressed and less variable than actual performance. In this paper it is shown that such compression is a feature of the optimal contract between a risk neutral principal and a risk averse agent when rewards are based upon a subjective evaluation of performance. The extent to which the principal is able to reward the agent as a function of a subjective evaluation depends upon the degree to which the agent agrees with these evaluations. When the principal's and the agent's subjective evaluations concur, then one can implement the optimal contract, just as if evaluations were objective and verifiable. Conversely, when the principal's and agent's signals are uncorrelated, the optimal contract compresses evaluation into two levels - acceptable and unacceptable, with only the very worst performances receiving the unacceptable ranking. This latter result is consistent with

*I would like to thank Canice Prendergast for suggesting this problem and for helpful comments on a first draft. I also greatly appreciate the comments of Janet Currie, Jean-Jacques Laffont, Jonathan Levin, Joel Sobel, the two annonymous referees, and the audiences at numerous seminars. Bharat Bhole and Mehdi Farsi provided excellent research assistance, and I would like to thank the National Science Foundation for its support via grant SBE-9709333.

evidence documented in Canice Prendergast (1999), illustrating the reluctance of supervisors to distinguish between employees, particularly when it affects compensation.

The results also provide new insights into the impact of bias upon pay and performance. There is some direct evidence of bias in subjective evaluations of performance, yet, as Harry Holzer and David Neumark (2000) discuss, one finds that the addition of better controls for individual productivity tends to reduce, and, in some cases, eliminate discrimination as measured by the impact of race or gender upon wages. It is shown that these contrasting observations are consistent with this model, namely if the evaluations of the principal are biased against one of two equally skilled workers, then the discriminated against worker has both lower pay and performance.

Much of the previous literature on incentive contracts has focussed upon the problem of designing compensation schemes based upon verifiable measures of performance, where, as Bengt Holmström (1979) and Milton Harris and Arthur Raviv (1979) have shown, compensation should vary with any useful piece of information.[1] These results apply to jobs where objective measures of performance are available, yet as Prendergast (1999) observes, "most people don't work in jobs like these".[2] Rather, rewards such as bonuses and promotions often depend upon the subjective evaluations provided by an individual's employer. Moreover, even when verifiable measures of output are available, subjective evaluations of performance also affect compensation, particularly if one considers rewards such as promotions.[3]

The literature on subjective evaluation for the most part uses a repeated game analysis to provide conditions under which efficient contracts are possible when evaluations are subjective, and hence non-contractable. Clive Bull (1987) and W. Bentley MacLeod and James M. Malcomson (1989) show that when the principal and agent have the same beliefs regarding a subjective evaluation, then there exists, under the appropriate conditions, efficient self-enforcing contracts.[4] The case of a risk averse agent is considered by David G. Pearce and Ennio Stacchetti (1998), who extend the analysis of Baker et al. (1994), to show that the existence of contractible measures of performance can enhance the performance of implicit contracts. They consider the case in which the principal's and agent's subjective evaluations are perfectly correlated, and explore the implications for the inter-temporal structure of the equilibrium.

Jonathan Levin (March 2002) extends the MacLeod and Malcomson (1989) model to allow for imperfect subjective evaluation on the part of the principal, and demonstrates that the optimal contract takes a simple termination form, with punishment occurring when the principal's subjective evaluation of agent performance falls below a

[1] A major caveat to this result is the work of Bengt Holmstrom and Paul Milgrom (1991) who show that in some cases information may be ignored if this causes a decreases in performance along some other dimension of performance. For example, a speed of completion bonus may not be given in a construction contract if this has a large adverse impact on quality.

[2] Page 57.

[3] See Baker et al. (1994) on this point. In any given year about 20% - 30% of US workers receive some form of subjectively determined reward, while for most occupations only about 1% to 5% of workers report receiving performance pay in the form of commissions or piece rates. For more details see W. Bentley MacLeod and Daniel Parent (1999).

[4] Bull (1987) shows the existence of such equilibria assuming that firms have different unobservable characteristics, while MacLeod and Malcomson (1989) characterize the set of self-enforcing contracts in a model with symmetric information between the contracting parties.

threshold. In these models repeated interaction is needed to ensure that the principal and agent can credibly impose costs upon each other when either of them deviate from the implicit contract. These costs are generated when individuals either leave the relationship, or carry out inefficient actions for several periods that in effect punishes both parties. Given that these equilibrium actions lower the total value of the relationship they have the natural interpretation as the conflict that arises when there is cheating or perceived unfairness in a relationship.

In other words, during the contract formation stage the contracting parties *structure* the conflict so that it provides the appropriate incentives for performance, while at the same time keeping the expected costs of conflict as low as possible. Given that the potential for equilibrium punishment in repeated games is well appreciated, this paper takes this as given, and derives the optimal contract when there are no limits upon the costs that can be imposed upon either party *ex post.*[5] The work of Levin (March 2002) is extended to case of a risk averse agent, and it is shown that conflict in this model is part of optimal contract, with wage compression arising from the trade-off between providing incentives *ex ante,* and reducing the cost of conflict *ex post.*

The standard model of subjective evaluation in this literature supposes that the signal of performance is common knowledge to the contracting parties, but is not verifiable to a third party. This paper considers the general case in which the principal and agent have private (and hence subjective) measures of performance that are possibly correlated with each other. The optimal contract is structured to ensure that both parties have an incentive to reveal their private information, with the threat of conflict ensuring that the principal has an incentive to reveal favorable observations that result in higher compensation to the agent. The main features of the optimal contract can be summarized as follows:

1. The agent's compensation does not depend upon her own self-evaluation, but only upon the principal's evaluation of performance. This is consistent with the general recommendation in the management literature against the use of self-evaluations to set compensation.[6]

2. The agent's self evaluation does play an important role in lowering the costs of implementing a system of subjective evaluation. Under an optimal contract the agent imposes a cost upon the principal whenever she feels that her evaluation is unfair, that is whenever she is given a low ranking, when she believes her work is of high quality.

3. When the agent's beliefs regarding performance are only weakly correlated with the principal's, it is optimal for the principal to reward the agent for all performance evaluations above the lowest possible rating, and hence the agent imposes a cost upon the principal only when the worst evaluation is given.

[5] Though in section III.C the case with limits upon these costs is also considered.

[6] The management literature on compensation is voluminous, and cannot be adequately reviewed here. See George T. Milkovich and Jeffrey M. Neuman (1966) for a good review of standard compensation practices, particularly chapter 10 that addresses the problems with subjective evaluation. See also Milkovich and Alexandra K. Wigdor (1991), which provides a general review of the literature. Both sources (pages 69 and 374 respectively) emphasize the importance of self-evaluation for increasing the coherence between supervisor and employee expectations of performance.

In practice, supervisors are given multi-part scales, and asked to rank employees on a scale such as below average, average and above average. Result 3 would imply that the supervisor would in the extreme case use only two of the provided levels, and would pool the above average individuals with the average individuals. This would imply that more than half of the employees would receive the highest rating, a result that is consistent with some of the evidence on performance rankings.[7] In contrast, for organizations that have shared values regarding what constitutes good performance, one would expect to observe less pooling of evaluations, and more effective incentive pay. In the limit, when the beliefs of the principal and agent are perfectly correlated, there are no agency costs associated with the use of subjective evaluations.

Other work that is closely related to this paper is Prendergast (1993) and Canice Prendergast and Robert H. Topel (1996) who demonstrate the importance of subjective evaluation for understanding some features of organizations, such as the tendency of employees to agree with the views of their supervisors. In order to obtain closed form solutions these papers consider a restricted class of contracts, and leave open the question regarding the form of the optimal contract with subjective performance evaluation.[8] In particular, they suppose that compensation is a linear function of subjective evaluations, and hence as evaluations become more compressed, pay would not vary with subjective evaluations at all. In contrast, when compensation is allowed to depend upon subjective evaluations in a general manner, and there is little correlation between the agent's and the principal's evaluations, the optimal contract takes the form of a simple two part contract, as described above, with only individuals having the worst ranking not receiving a reward from the principal.

The agenda of the paper is as follows. The next section introduces the basic principal-agent model. Section 3 provides an analysis of the optimal contract with subjective evaluation. The implications of subjective evaluation for the theory of discrimination are discussed in section 4. The final section of the paper discusses the general implications of the model for the theory of organizations and employment practices such as efficiency wages.

1 I The Model

Consider a principal who offers a one period employment contract to an agent. If the agent accepts the contract, then she chooses effort $\lambda \in [0, 1]$, where λ is the probability that a benefit B is realized. The net benefit to the principal is:

$$\Pi = \lambda B - E\{W\},$$

where W are the dollar costs of employing the agent. The agent is risk averse, with preferences $U(c, \lambda)$ satisfying the following assumption:

Assumption A: The Bernoulli utility function of the agent satisfies $U(c, \lambda) = u(c) - V(\lambda)$, where $c > 0$, $\lambda \in [0, 1)$ and $u' > \varepsilon > 0$, $u'' < 0$, $\lim_{c \to 0} u(c) = -\infty$, $V' > 0$, $V'' > 0$ and $\lim_{\lambda \to 1} V(\lambda) = \infty$.

[7] See Milkovich and Newman (1996) page 364.

[8] See footnote 7 in Prendergast and Topel (1996).

It is assumed that the agent is risk averse, and that it is not possible for the agent to produce the benefit B with probability 1. If the agent rejects the contract offer, she is assumed to have a market opportunity yielding an expected utility of $\bar{u} > 0$. The condition $\lim_{c\to 0} u(c) = -\infty$, implies the Inada condition $\lim_{c\to 0} u'(c) = \infty$, and ensures that an optimal contract always entails positive consumption. When combined with the assumption that $u' > \varepsilon$, this ensures that it is always possible to implement any level of effort less than 1 and while simultaneously satisfying the individual rationality constraint.

It is also assumed that B is not directly observable, rather it corresponds to a complex good or service whose ultimate quality is difficult to determine. For example, B may be the return on a film that becomes a blockbuster; even though many individuals contribute to the success of the film (the secretaries, technicians etc.), except for some of the principals in the project, their compensation typically depends upon their perceived contribution to the project at the time their services are performed. The model can also be interpreted as a commercial contract between two firms, P and A, in which firm A supplies a service or good for firm P's production process, such as the design of some new manufacturing equipment. In this case the effort of firm A affects the probability that the new good produced by firm P is a success. In many, if not most cases, subcontractors, such as firm A, are not likely to be directly compensated as a function of the success of the final product.

Rather than observe whether or not B has been (or will be) realized, the principal observes a measure of performance $t \in T = \{1, ..., n\}$. The probability of t given the outcome is/will be a success is γ_t^H, while it is not/will not be a success with probability γ_t^L. Let $\gamma^k = \left[\gamma_1^k, .., \gamma_t^k, .., \gamma_n^k\right]^\top$ be the vector of probabilities when the performance is a success ($k = H$) or not ($k = L$).

Assumption B: (Generic Monotone Likelihood condition) $\gamma_{t+1}^H/\gamma_{t+1}^L > \gamma_t^H/\gamma_t^L > 0$ for $t = 1, .., n-1$, and for no t is it the case that $\gamma_t^H/\gamma_t^L = 1$.[9]

Given that the signals can always be relabeled so that the weak inequality holds, the only restrictive conditions are that these inequalities must be strict, and that all states occur with strictly positive probability. For the most part this is not a restrictive assumption, and allows for a somewhat more elegant derivation of the optimal contract. Given effort λ, the probability of observing signal t is given by $\gamma_t(\lambda) = \lambda\gamma_t^H + (1-\lambda)\gamma_t^L$, with $\gamma(\lambda) = \lambda\gamma^H + (1-\lambda)\gamma^L$ denoting the corresponding vector of probabilities. Notice that for $\lambda > 0$, all signals occur with positive probability, and hence the standard full support assumption is satisfied in this model.

When the signals are observable and contractible, this model is a standard principal-agent problem (Harris and Raviv (1979) and Holmström (1979)). The agent's contract is given by the vector of payments for each signal: $c = [c_1, ..., c_t, ..., c_n]^\top \in \Re_+^n$, where c_t is the agent's consumption in state t. The model satisfies the basic conditions for the principal-agent problem studied in Sanford J. Grossman and Oliver D. Hart (1983). They show that one can decompose the problem into two steps. First one

[9]The last assumption, $\gamma_t^H/\gamma_t^L \neq 1$ for any t, is a technical assumption that holds generically. It ensures that any signal t is unambiguously either a good or bad signal, and simplifies the characterization of the optimal contract.

determines the cost of eliciting a level of effort λ, denoted $C^*(\lambda)$, and then one can solve $\max_{\lambda\in[0,1]} \lambda B - C^*(\lambda)$ to determine the optimal effort level. The optimal contract that elicits effort λ at the lowest cost, $C^*(\lambda)$, solves:

$$C^*(\lambda) = \min_{c\in\Re^n_+} c^\top \gamma(\lambda), \tag{1}$$

$$u(c)^\top \gamma(\lambda) - V(\lambda) \geq \bar{u}, \tag{2}$$

$$\lambda \in \arg\max_{\tilde{\lambda}\in[0,1]} u(c)^\top \gamma\left(\tilde{\lambda}\right) - V\left(\tilde{\lambda}\right), \tag{3}$$

where $u(c) = [u(c_1), ..., u(c_n)]^\top$ is the vector of state contingent utilities, constraint 2 is the individual rationality constraint, while 3 is the incentive compatibility constraint ensuring that the agent has an incentive to select λ. As is standard in this literature, it is assumed that when several solutions to 3 exist, the agent selects the effort desired by the principal. As a matter of convention if there is no contract implementing λ, then $C^*(\lambda) = \infty$.

Grossman and Hart (1983) have shown that assumptions A and B ensure that a solution to this problem exists. In the absence of the incentive constraint 3, it is optimal to pay a risk averse agent a fixed wage, $c_t = w$ that satisfies $u(w) = \bar{u} + V(\lambda)$. However, if effort is not observable, and the agent is paid an income that is independent of performance, then she will set effort at the lowest possible level, $\lambda = 0$. To induce effort, the optimal contract entails compensation that is a function of the signal t. Holmström (1979) has shown that the extent to which consumption varies with t depends upon the quality of the signal, a result that is summarized in the following proposition.

Proposition 1 *Given assumptions A and B, then for* $\lambda \in [0,1)$, $\infty > C^*(\lambda) > 0$, *and the associated optimal contract,* c^*, *has the property that* $c^*_{t+1} > c^*_t$, *for* $t = 1, ..., n-1$, *whenever* $\lambda > 0$.

The proof of this proposition, as well as longer proofs for subsequent propositions are found the appendix.

An intuitive interpretation of this result is that observing a higher signal provides a stronger signal of good performances, and consequently the individual's pay should increase with t. Notice that risk aversion plays an important role in determining a unique optimal contract. In the case that the agent is risk neutral *any* contract satisfying the IC and IR constraints with equality is optimal. It follows from the arguments of Grossman and Hart (1983) that $C^*(\lambda)$ is lower semi-continuous and hence there is an optimal effort, λ^*, solving:

$$\max_{\lambda\in[0,1].} \lambda B - C^*(\lambda).$$

If C^* is twice differentiable at λ^* then one has $B = C^{*'}(\lambda^*)$, and the second order condition implies that the optimal effort level rises with B.

2 II The Optimal Contract with Subjective Evaluation

Most performance evaluations are at least partially subjective, whether it is in an employment context, or for a commercial contract that entails the receipt of goods or

services. Typically the principal has a right to observe performance and/or inspect goods and then decide if the quality is appropriate. However, while the principal may have a definite opinion regarding the quality of performance, it may be quite difficult to provide a corresponding objective measure that is sufficiently precise or observable to be enforceable by contract. For example, consider the quality of a research report, a work of art, or food prepared by a chef. In some cases, such as with food or a service, the quality of the good is not even stable over time, and hence even if the owner of a restaurant is certain that food quality is low, preserving the physical evidence for the purposes of contract enforcement may be impossible.

Even though performance may not be contractible, the subjective evaluation of performance is not arbitrary because good evaluations are likely to be highly correlated between individuals. This section explores the implications of such correlation for optimal contract design with subjective evaluation. Specifically, suppose that after the principal observes t, then the agent makes her personal evaluation of her own performance, denoted $s \in T$. As before, let γ_t^H and γ_t^L denote the probability that the principal observes signal t when performance is respectively H and L. .

If s provides additional information regarding performance λ, then the optimal contract should incorporate this information. To focus upon the role that s can play in enforcing an optimal contract, I follow Prendergast (1993) and assume that the likelihood of a particular s occurring is a function of t. Let the agent's probability of observing s conditional upon the principal observing t be $P_{ts} = \Pr\{s|t\}$. The implications of relaxing this assumption are considered in section III.

If evaluations are perfectly correlated then:

$$P_{ts} = I_{ts} \equiv \begin{cases} 1, & \text{if } t = s, \\ 0, & \text{if not.} \end{cases}, \tag{4}$$

while if the agent's signal has no information regarding the principal's evaluation then $P_{ts} = P_{\bar{t}s}$, for all $t, \bar{t}, s \in T$. The probability of the pair ts occurring in states H and L is $\Gamma_{ts}^H = P_{ts}\gamma_t^H$ and $\Gamma_{ts}^L = P_{ts}\gamma_t^L$, respectively. Given the effort λ by the agent, the *ex ante* unconditional probability of state ts is:

$$\Gamma_{ts}(\lambda) \equiv \lambda\Gamma_{ts}^H + (1-\lambda)\Gamma_{ts}^L. \tag{5}$$

The revelation principle is used characterize the optimal contract. It implies that for any contract $\Psi = \{c_{ts}, w_{ts}\}_{t,s\in T}$, where c_{ts} is the consumption of the agent in state ts, while w_{ts} is the wage paid by the principal in state ts, it must be the case that neither the principal, nor the agent have an incentive to misrepresent their information. The program for the optimal contract consists of adding the incentive constraints for the revelation of subjective information to the principal-agent problem with complete

contracts:

$$C^s(\lambda) \equiv \min_{\Psi \in \Re^{2n^2}} \sum_{t,s \in T} w_{ts} \Gamma_{ts}(\lambda), \tag{6}$$

$$\sum_{t,s \in T} u(c_{ts}) \Gamma_{ts}(\lambda) - V(\lambda) \geq \bar{u}, \tag{7}$$

$$\lambda \in \arg\max_{\tilde{\lambda} \in [0,1]} \sum_{t,s \in T} u(c_{ts}) \Gamma_{ts}\left(\tilde{\lambda}\right) - V\left(\tilde{\lambda}\right), \tag{8}$$

$$\sum_{s \in T} w_{ts} \Gamma_{ts}(\lambda) \leq \sum_{s \in T} w_{\bar{t}s} \Gamma_{ts}(\lambda), \ \forall t, \bar{t} \in T, \tag{9}$$

$$\sum_{t \in T} u(c_{ts}) \Gamma_{ts}(\lambda) \geq \sum_{t \in T} u(c_{t\bar{s}}) \Gamma_{ts}(\lambda), \ \forall s, \bar{s} \in T, \tag{10}$$

$$w_{ts} \geq c_{ts} \geq 0. \tag{11}$$

The new constraints for this program are 9 to 11. Constraint 9 requires that the principal's costs be lowest when he reports his true type. Expected cost conditional upon t is $\sum_{s \in T} w_{ts} \Gamma_{ts}(\lambda) / \sum_{s \in T} \Gamma_{ts}(\lambda)$. If $\sum_{s \in T} \Gamma_{ts}(\lambda) \neq 0$, then this cancels on both sides, while if $\sum_{s \in T} \Gamma_{ts}(\lambda) = 0$, then $\Gamma_{ts}(\lambda) = 0$ for $s \in T$, and the inequality is automatically satisfied.

Similarly constraint 10 requires the agent to weakly prefer to report s truthfully. Finally, constraint 11 requires that the consumption of the agent is less than or equal to the payment of the principal. One would normally suppose that the wage payment is equal to consumption, $w_{ts} = c_{ts}$, however in that case it would be impossible to elicit any effort:

Proposition 2 *Suppose that $c_{ts} = w_{ts}$ for all $t, s \in T$, then for $\lambda > 0$,the cost of effort is undefined, and hence the only possible solution entails no effort ($\lambda = 0$), and a fixed wage contract $\bar{w} = u^{-1}(\bar{u})$.*

Proof. If $c_{ts} = w_{ts}$ then after making their subjective evaluations, the principal and agent play a constant-sum game when making their reports. From the min-max theorem such a game has a unique value and hence the agent's compensation cannot depend upon t. ■

Therefore in this principal agent model it is impossible to elicit subjective information under the hypothesis that the contract is budget balancing.[10] When the budget balancing constraint is relaxed, it is straightforward to establish the existence of a contract implementing effort λ. For example, if $c^* = \left[c_1^*, ., c_t^*, .., c_n^*\right]$ is a consumption contract implementing effort level λ, then the contract $c_{ts} = c_t^*$ and $w_{ts} = \max_{t \in T} c_t^* = w$, for all $t, s \in T$, satisfies the constraints 7 to 11. Under this contract, the report of the principal does not affect her wage payment, w, while the agent's information is ignored.

Formally, this can be achieved by having the principal pay the difference $w - c_t^*$ to a third party, however this is not the only mechanism available. As discussed above,

[10]The constraints that a balanced budget place upon the set of implementable contracts is an important theme in the theory of incentives. See Jerry R. Green and Jean-Jacques Laffont (1979) for a review of the early literature, and John Moore (1992) for the more recent literature.

the difference $w - c_t^*$ also has the interpretation as equilibrium conflict in repeated relationship, in which the principal or agent engages in costly and mutually unproductive behavior. This might be work to rule behavior by unionized employees, the threat to leave a job, sabotage at the firm, etc. It is well know that such conflict is a ubiquitous feature of any organization yet, as Milkovich and Newman (1996) discuss, such conflict is typically viewed as being counter productive.

In contrast, this proposition demonstrates that in the absence of some method of generating costs in a relationship it is impossible to implement performance incentives based upon subjective evaluation alone. Rather, the analysis suggests that the role of a good subjective evaluation system is not the elimination of socially wasteful conflict, but rather to find an optimal trade-off between the imposition of costs *ex post* on the relationship and the provision of performance incentives. The next proposition establishes the existence of such contracts, and a basic characterization of their structure.

Proposition 3 *Given Assumptions A and B, then for all* $\lambda \in [0, 1)$, *there is a cost minimizing contract,* $\Psi^S = \left\{c_{ts}^S, w_{ts}^S\right\}_{t,s\in T}$, *implementing effort* λ, *with the property that* $c_{ts}^S = c_{t\hat{s}}^S$ *for all* $t, s, \hat{s} \in T$.

This proposition establishes the existence of an optimal contract implementing effort λ based upon the fact that the constraint set is closed and the set of potential candidate solutions can be bounded. Secondly the consumption that the agent receives depends only upon the evaluation of the principal, a result that follows from the fact that the agent's signal is not more informative than the principal's, combined with the linearity of the incentive constraints. From the arguments of Grossman and Hart (1983) it also follows that $C^S(\lambda)$ is lower semi-continuous and hence an optimal effort level exists and is the solution to:

$$\max_{\lambda\in[0,1]} \lambda B - C^S(\lambda).$$

The expected deadweight loss from using subjective evaluation is $\sum_{t,s\in T}\left(w_{ts}^S - c_t^S\right)\Gamma_{ts}(\lambda)$, which is strictly positive if and only if $C^S(\lambda) > C^*(\lambda)$. The next three subsections explore the implications of placing additional structure upon the set of beliefs.

2.1 II.A Perfect Correlation

Consider first the case of perfect correlation between the principal's and agent's beliefs regarding their assessments, with $P_{ts} = I_{ts}$ as defined by equation 4. In this case it is straightforward to show that one can implement the optimal complete contract.

Proposition 4 *If the principal's and agent's signals are perfectly correlated, then the incentive constraints, 9 and 10 are not binding, hence the optimal contract with subjective evaluation is the same as the optimal principal-agent contract with verifiable information.*

Proof. Let $\left\{c_t^*\right\}$ be the optimal complete contract, and set $w_{tt}^* = c_t^*$, and for $t \neq s$ set $w_{ts}^* = \max_{t\in T} c_t^* + k$, where $k > 0$. From proposition 1 we know that the agent

automatically satisfies her incentive constraints. The cost to the principal who reports t', given that he has observed t is:

$$C\left(t'|t\right) = \begin{cases} c_t^* & \text{if } t = t', \\ \max_{t \in T} c_t^* + k & \text{if not.} \end{cases} \tag{12}$$

Clearly, $C\left(t|t\right) < C\left(t'|t\right)$ for $t' \neq t$, and hence the principal always has an incentive to report truthfully, and thus this contract results in the optimal complete contract with no welfare loss due to the incentive constraints arising from subjective evaluation. ■

This result highlights the fact that under the hypothesis that the budget balancing condition is only weakly satisfied, then the subjective nature of assessments by itself does not imply inefficiency. As long as individuals can agree upon whether performance is acceptable or not, then it is possible to write an efficient contract. In practice one would never expect perfect agreement and hence it is important to know if this result is approximately correct when beliefs are highly, but not perfectly, correlated. To address this question consider a sequence of beliefs, P_{ts}^k with the property $P_{ts}^k > 0$ for all types and that $\lim_{k\to\infty} P_{ts}^k = I_{ts}$. Let $A^k = \left\{c_{ts}^k, w_{ts}^k\right\}_{t,s\in T} \in \Re^{2n^2}$ be an optimal contract given P_{ts}^k, and let $C^k\left(\lambda\right)$ be the associated costs.

Proposition 5 *The optimal contract when there is perfect correlation in beliefs is close to the contract with imperfect correlation. Specifically* $\lim_{k\to\infty} c_{ts}^k = c_{ts}^*$ *and* $\lim_{k\to\infty} C^k\left(\lambda\right) = C^*\left(\lambda\right)$, *the cost function for the standard principal-agent problem.*

This result demonstrates that when beliefs are highly correlated then the optimal consumption is close to the contract predicted in the standard principal agent model, and that the expected social loss will be close to zero. Unfortunately, the social loss when $t \neq s$ (which occurs with low probability) cannot be uniformly bounded, and hence even though beliefs are highly correlated, one cannot conclude that *ex post* the social loss in some states will not be very large. This point is explored further in section II.C.

2.2 II.B No Correlation in Subjective Assessments

Suppose now that the agent's signal entails no information regarding the principal's evaluation, that is $P_{ts} = P_{ts'}$ for all $t, s, s\prime \in T$. In this case making the payments, w_{ts}, depend upon s does not relax the incentive constraints, and hence without loss of generality we may set $w_{ts} = w_t$, for all $s \in T$. The incentive constraint for the principal observing signal t is:

$$\left(w_t - w_{\bar{t}}\right)\gamma_t\left(\lambda\right) \leq 0, \ \forall t, \bar{t} \in T,$$

(recall $\gamma_t\left(\lambda\right) = \lambda\gamma_t^H + \left(1-\lambda\right)\gamma_t^L$), from which one concludes that $w_t = w_{\bar{t}} \equiv w$, for all $t, \bar{t} \in T$. Hence the optimal contract entails the principal facing costs that are independent of his subjective assessment. Notice that the optimal consumption contract with objective performance measures, c^*, is feasible in this case by letting $w = \max_{t\in T} c_t^*$. Since there is a cost to be paid whenever $c_t < w$, this creates an incentive for the principal to compress the variation in payments to the agent. The

interesting result is that rather then reduce the variation for each signal, the optimal contract has $c_t = w$ for all signals except the worst evaluation:

Proposition 6 *Given Assumptions A and B, and suppose there is no correlation in beliefs ($P_{ts} = P_{ts'}$ for all $t, s, s' \in T$) then the optimal contract implementing effort λ based upon the principal's subjective evaluations entails wage payments that do not depend upon the principal's evaluation: $w_t = w + b$ for all $t \in T$, while the agent receives:*

$$c_t = \begin{cases} w + b & t > 1, \\ w & t = 1, \end{cases} \tag{13}$$

where $t = 1$ corresponds to the lowest performance level. Let $\gamma_g^H = \sum_{t=2}^{n} \gamma_t^H$, and similarly for γ_g^L and $\gamma_g(\lambda)$, then w and b are the unique solutions to:

$$u(w + b) = \bar{u} + V(\lambda) + \frac{\gamma_g(\lambda) V'(\lambda)}{\gamma_g^H - \gamma_g^L}, \tag{14}$$

$$u(w) = \bar{u} + V(\lambda) - \frac{\gamma_g(\lambda) V'(\lambda)}{\gamma_g^H - \gamma_g^L}. \tag{15}$$

The cost function is:

$$C^{NC}(\lambda) = w + b = u^{-1}\left(\bar{u} + V(\lambda) + \frac{\gamma_g(\lambda) V'(\lambda)}{\gamma_g^H - \gamma_g^L}\right)$$

This result shows that when there is no correlation between the principal's and agent's beliefs, then the optimal contract pays the same bonus to the agent for all but the very worst signal of performance. The intuition for the result depends upon the fact that the principal wishes to avoid paying the cost ($\max_{t'} c_{t'} - c_t$) when signal t occurs. Since $t = 1$ is most informative regarding low effort, then it is optimal to punish the worker only when this signal is observed. This result is consistent with the observed tendency of supervisors to avoid giving low evaluations to employees. Though it may be normal to view such behavior as a sign of "softness" on the part of the supervisor, this result demonstrates that such pooling of evaluations is part of an optimal contract when there is little correlation between the supervisor's and employee's perceptions of performance.

When such pooling occurs, firms sometimes experiment with schemes that force supervisors to discriminate between employees, for example they may be required to rank order employees from top to bottom.[11] The general mechanism design approach does not exclude the possibility of making pay sensitive to the supervisors true beliefs, but highlights the fact that such pay for performance may come at the cost of increased conflict, a result that is quite consistent with the management literature on the issue. In management terms, if an employee receives a low rating that she feels is unwarranted,

[11] Edward E. Lawler III (2000), page 185, points out that while such rankings are used in practice, it should have no place in most organizations due to the error and randomness that comes into play. He recommends that at most three performance levels be used.

then this would be "demoralizing" and can result in lower output in the future, which in turn results in lost output for the firm. Given that the pooling arises from the lack of correlation in beliefs between the principal and agent, the model is consistent with the emphasis that management texts place upon designing subjective evaluation systems that employees, as well as management, believe result in fair and correct evaluations.[12]

This contract is also optimal when the agent is risk neutral (though it is no longer unique), a case that can be used to illustrate the impact that this contract has for effort. Under a complete contract and a risk neutral agent ($u(w) = w$), there are no agency costs, and hence first best effort satisfies:

$$B = V'\left(\lambda^*\right).$$

With subjective evaluation, one has $C^{NC}(\lambda) = \bar{u} + V(\lambda) + \frac{\gamma_g(\lambda)}{\gamma_g^H - \gamma_g^L} V'(\lambda)$, which combined with the fact $\partial \gamma_b(\lambda) / \partial \lambda = \left(\gamma_g^H - \gamma_g^L\right)$, implies that the first order condition for an interior solution is:

$$B = \frac{\gamma_g\left(\lambda^{NC}\right)}{\gamma_g^H - \gamma_g^L} V''\left(\lambda^{NC}\right).$$

Since beliefs and the second derivative of $V(\lambda)$ can be selected independently from the first derivative of $V(\lambda)$, then it is the case that effort, λ^{NC}, can be less than or greater than the first best effort effort under a complete contract, λ^*. The reason for this result is that increasing effort decreases the probability that a low outcome occurs, and hence the probability that there will be a social loss. If this effect is large enough, then effort under subjective evaluation may be greater than under a complete contract.

2.3 II.C The Effect of Imperfect Correlation

Unlike the case of perfect correlation, the optimal contract may not be continuous when there is no correlation in beliefs, because, as R. Preston McAfee and Philip J. Reny (1992) show for a general bargaining model, if there is a small amount of correlation it may be possible to implement a first best contract. The purpose of this section is to explore in more detail the case of imperfect correlation, and provide conditions under which continuity is restored. The discontinuity problem can be nicely illustrated with the following parameterization of beliefs[13].

Assumption D (Parameterized Beliefs) Suppose P_{ts}^p is such that with probability $1-p$ the agent observes a "no information signal", denoted by $s = 0$, while with probability p she observes the signal $s = t$, where t is the signal observed by the principal.

When $p = 1$ one has the case of perfect correlation, while $p = 0$ corresponds to the case for which the agent has no information regarding performance, and hence

[12] For example Milkovich and Newman (1996) spend most of their chapter on subjective evaluation on the problem of design accurate methods of performance measurement.

[13] I am grateful to Preston McAfee for suggesting this example.

varying p from 1 to 0 continuously varies the degree of correlation between the cases considered in proposition 4 and 6. For this class of beliefs the first best can be achieved for *all* $p > 0$. To see this, let c_t^* be the optimal contract as given in proposition 1, and let $w_{ts} = c_t^*$ for $s = 0$ or $s = t$. It has already been shown that the agent's incentive constraint at an optimal contract is automatically satisfied since consumption does not vary with the agent's report, thus it only remains to structure the payments w_{ts} such that the principal's incentive constraints are satisfied.

If the principal observes t, and the agent is truthful, then the expected payment by the principal is c_t^*, and hence the incentive constraint for the principal when $\hat{t}$ is reported implies:

$$c_t^* \leq (1-p)\, c_{\hat{t}}^* + p \cdot w_{\hat{t}t}.$$

Hence for $\hat{t} \neq t \neq 0$, if one sets $w_{\hat{t}t} = c_t^* + \max\left\{0, \left(c_t^* - c_{\hat{t}}^*\right)/p\right\}$, then this incentive constraint is satisfied. Moreover, the probability of $\hat{t}t$ occurring is zero, and therefore the incentive constraints are satisfied at no cost for all $p > 0$. Observe that this contract requires arbitrarily large payments/penalties as the correlation becomes small, namely whenever $\left(c_t^* - c_{\hat{t}}^*\right) > 0$ then $w_{\hat{t}t} \to \infty$ as $p \to 0$.

As Levin (March 2002) observes, when implementing an optimal contract in a repeated relationship there are likely to be limits on the size of the punishments that can be imposed. Suppose that this limit is given by $S > 0$, and therefore the contract must also satisfy:

$$|w_{ts} - c_t| \leq S, \forall t, s \in T. \tag{16}$$

The addition of this constraint ensures that the set of contracts can be bounded, and is sufficient to ensure the continuity of the optimal contract with respect to beliefs. In this example the punishments became unbounded because some states occur with probability zero. If one supposes that there is some noise, and hence *all* states occur with strictly positive probability, then unbounded *ex post* punishments would also entail unbounded costs *ex ante,* and hence would never be chosen as part of an optimal contract.

Both cases imply that the optimal contract is chosen from a compact set which ensures continuity of the optimal contract with respect to beliefs. These observations are summarized in the following proposition.

Proposition 7 *Suppose assumptions A and B are satisfied, and consider a sequence of beliefs* $P_{ts}^k \to P_{ts}$, *where either (a)* $P_{ts} > 0$ *for all* $ts \in T \times T$ *or (b) condition 16 is satisfied. Then for* $\lambda \in [0, 1)$, *the optimal cost function converges,* $C^k(\lambda) \to C(\lambda)$, *and the limit points of the optimal contract,* c_t^k, *are optimal for the beliefs* P_{ts}.

When P_{ts} corresponds to no-correlation in beliefs and all signals occur with positive probability ($P_{ts} = P_{ts'} > 0$ for all $t, s, s' \in T$), then when beliefs are close to these the optimal contract will have consumption that is approximately constant for all $t > 1$. In fact it is not difficult to show that a somewhat stronger result, namely for some $\bar{k}$, one has for all $k > \bar{k}$ that agent receives the same payment for all signals greater than the lowest possible signal. Beyond this it is difficult to characterize the optimal contract with imperfect correlation in the general case. When beliefs satisfying

assumption D and constraint 16 is imposed, then the pooling of evaluations at the top of the performance distribution holds for the intermediate cases as well.

Proposition 8 *Suppose assumptions A, B and D are satisfied and the amount of loss* ex post *is constrained by S. If S is sufficiently large (but finite) then for every p there is a type* $t(p)$*, such that*

$$\bar{c} = c_n = c_{n-1} .. = c_{t(p)+1} > c_{t(p)} > c_{t(p)-1} > .. > c_1,$$

with the property that for some $\bar{p}$ *sufficiently close to zero* $t(p) = 1$*, for* $\bar{p} \geq p \geq 0$*. Moreover, when correlation is perfect the optimal contract is implemented* ($c_t^1 = c_t^*$)*.*

When the bound on $|w_{ts} - c_t|$ is sufficiently large, then under perfect correlation one obtains the optimal complete contract. As the degree of correlation decreases, then this bound must eventually be binding. The interesting result is that if the bound S is sufficiently large that $b < S$, where b is the bonus pay under the contract with no correlation (see proposition 6), then proposition 7 implies that the contract with imperfect correlation converges to the optimal contract with no correlation, with all wage compression occurring at the top of the distribution. Together these results illustrate that in general when the correlation between the principal's and agent's evaluation of performance is weak, and there are limits upon the *ex post* costs that can be implemented, then the optimal contract entails a pooling of high performance evaluations. This has the consequence that in these cases most people are likely to be judged to be above average!

2.4 II.D Implementing the Optimal Contract

The revelation principal is a technical device that allows one to characterize the optimal contract as a function of the underlying information structure, however it does not describe how one would implement the contract in practice. The purpose of this section is to briefly describe how the optimal contract may be implemented in the context of a repeated agency, and in particular how the social cost necessary for the optimal contract can be rationalized as equilibrium behavior.[14]

Suppose that the optimal contract for the static problem has payments given by $\{w_{ts}, c_t\}$. When beliefs are perfectly correlated and both principal and agent are risk neutral MacLeod and Malcomson (1989) have shown that when the value of the relationship is sufficiently large relative to market alternatives, there exists a budget balancing contract that results in the first best effort by the agent. The equilibrium strategy each period entails the agent selecting effort, followed by the principal paying a bonus if and only if effort is high enough. If the principal does not pay the bonus when the agent has worked hard, then she leaves the relationship.

[14] See Levin (March 2002) for a formal analysis of a repeated agency model with private information with a risk neutral agent. An earlier version of this paper, MacLeod (2001), sketchs the formal repeated game model for the case of a risk averse agent. See also Roy Radner (1985), the first paper on repeated agency with discounting, for a discussion of the relationship between repeated agency problems and repeated game theory.

In order for this contract to work it must be the case that leaving imposes a cost upon the principal, a hypothesis that is reasonable if the principal has made some relationship specific investments.[15] Given this hypothesis, then optimal contract with subjective evaluation can be implemented as follows. After the agent has chosen effort, the principal observes t and makes payment c_t. Under the optimal contract the agent is indifferent between staying or continuing the relationship, therefore it is optimal to quit with probability α_{ts}. If k is the lost to the firm from having the worker leave in the following period, then let this probability satisfy:

$$\alpha_{ts} = (w_{ts} - c_t) / k\delta,$$

where δ is the discount factor that is common to both the principal and agent. If k is sufficiently large that $\alpha_{ts} < 1$, then this ensures that it is optimal for the principal to select c_t, and the strategies of the principal and agent together form a sequential equilibrium.

Observe that one does not need to suppose that the principal has the right to fire the worker at this equilibrium, and hence the use of efficiency wages (the threat of dismissal when there is poor performance) is not a necessary feature for this implementation of the optimal contract. Also notice that one must use mixed strategies for the construction of this equilibrium, a ubiquitous feature of repeated games with asymmetric information (see Michihiro Kandori (2002) for a survey). This suggests that the analysis would be greatly complicated if the principal were also risk averse, an interesting case for future research.

Finally, this construction does not explore the potential for inter-temporal allocation of risk with private information, a case that thus far has not been explored in the literature (see Pearce and Stacchetti (1998) for some progress on this question in the case of public signals). Hence, though the costs necessary for the implementation of the optimal contract can be generated endogenously within a repeated relationship, there is still much work needed to fully understand the nature of repeated agency contracts with private information and risk averse individuals.

3 III The Effect of Biases upon Pay and Performance

The purpose of this section is to explore the effect of bias on the optimal contract. Here by "optimal" one does not mean socially optimal, but rather the contract that a biased, profit maximizing principal would chose. Bias in decision making can affect labor market outcomes in a number of ways, including through wage levels, the hiring decision and task assignment (see Joseph G. Altonji and Rebecca M. Blank (1999)). There is some work by economists that finds direct evidence of bias in subjective evaluations. For example Claudia Goldin and Cecilia Rouse (2000) find that when evaluators for orchestral positions could not observe the sex of the applicant (a screen was put up shielding the applicant, so only the sound could be heard), the number of women who were hired significantly increased. Lawrence M. Kahn (1991) finds that there is

[15] See Williamson, Wachter and Harris (1975) for a discussion of the importance of such rents for the employment relationship.

evidence of discrimination against French Canadian defense men on hockey teams, a position for which it is difficult to measure performance. There did not appear to be any discrimination against French Canadians when they were in positions such as forwards, where productivity in terms of goals scored can be more easily measured, suggesting that it is the subjective nature of the evaluation that causes the bias to affect compensation.

To explore the effect of this type of bias consider the problem of contracting with two levels of performance: $T = \{A, U\}$, where A denotes acceptable performance, and U denotes unacceptable performance. For simplicity suppose that if the low outcome occurs, both the principal and agent observe U for sure. Given that we have only two levels of performance, risk aversion is not a central factor in the determination of the optimal contract, and therefore it is assumed that the agent is risk neutral. This has the additional benefit of allowing less restrictive assumptions regarding beliefs. Specifically, let γ_{ts} be the probability that the signal pair ts is observed when the good outcome occurs.

Proposition 9 *Suppose that beliefs are positively correlated ($\gamma_{AA}\gamma_{UU} - \gamma_{AU}\gamma_{UA} > 0$). Then the optimal contract with subjective performance implementing effort λ has the form:*

		Agents Report	
		A	*U*
Principal's Report	*A*	$(-b-w, b+w)$	$(-b-w, b+w)$
	U	$(-P-w, w)$	$(-w, w)$

where:

- *The bonus satisfies:* $b = V'(\lambda) / (\gamma_{AA} + \gamma_{AU})$.
- *The penalty satisfies* $P = V'(\lambda) / \gamma_{AA}$
- *The wage satisfies* $w = U^0 + V(\lambda) - \lambda V'(\lambda)$
- *The cost function is* $C(\lambda) = U^0 + V(\lambda) + \lambda \frac{\gamma_{UA}}{\gamma_{AA}} V'(\lambda)$

This proposition illustrates that regardless of the structure of beliefs, the agent's reward depends only upon the principal's evaluation, while the agent's self evaluation is used to provide incentives for the principal to be truthful. Given that the agent's individual rationality constraint is binding the total compensation to the agent is $U^0 + V(\lambda)$, and hence compensation is increasing with effort λ. When signals are contractible, then in this risk neutral setting, there are no agency costs to implementing effort λ and hence the cost of implementing λ is $C^*(\lambda) = U^0 + V(\lambda)$.

When evaluations are subjective, there is an agency cost $\lambda \frac{\gamma_{UA}}{\gamma_{AA}} V'(\lambda)$, whose level is determined by the parameter $\alpha = \frac{\gamma_{UA}}{\gamma_{AA}}$, called the *perceived bias* in the relationship. The perceived bias α is a likelihood ratio representing the principal's belief that performance is acceptable, conditional upon the agent also believing that performance is acceptable. When α is zero the principal always agrees with the agent's self assessment, while when α is infinite there is never any agreement. Let $C(\lambda, \alpha) =$

$U^0 + V(\lambda) + \lambda\alpha V'(\lambda)$ be the cost of effort as a function of the bias, then the optimal level of effort under subjective evaluation solves:

$$B = C_\lambda\left(\lambda^*, \alpha\right).$$

The second order condition for this optimization problem implies $C_{\lambda\lambda}(\lambda^*, \alpha) \geq 0$, and hence one has the following proposition.

Proposition 10 *Suppose that preferences satisfy assumption A, then effort, and hence expected compensation, decreases with an increase in perceived bias:* $\partial\lambda^*/\partial\alpha \leq 0$.

To further explore the effect of beliefs upon performance consider the following parameterization of beliefs that generalizes assumption *D* to allow the agent to have an independent evaluation that is more informative than that of the principal. Let p be the probability that the principal observes A given that a good outcome is observed. If the principal were completely unbiased then $p = 1$, otherwise there is some chance that even though performance is acceptable, the principal feels that the quality is not acceptable. Correlation in beliefs is modelled by letting ρ be the probability that the agent has the same evaluation (or equivalently knows the evaluation) of the principal. With probability $(1 - \rho)$ the agent has her own independent observation of performance, which implies that the which the probability of acceptable performance is q. This parameter can be viewed as the agent's *self-confidence,* namely the probability that she feels her performance is acceptable when she is not able to observe the principal's evaluation.

Given these parameters perceived bias is given by:

$$\alpha(p, q, \rho) = \left(\frac{1-p}{p}\right)\left[\frac{1}{\left(\frac{\rho}{(1-\rho)q}\right)+1}\right].$$

If the principal is unbiased then $p = 1$, and the perceived bias is zero regardless of the agent's beliefs. An increase in bias by the principal always results in an increase in perceived bias, and hence lower performance and pay. This result highlights a potential pitfall in the use of standard labor market data sets to test for discrimination. A common measure of discrimination is the difference in wages between individuals in two different groups having the same skill level. However, as Holzer and Neumark (2000) observe, improving the measure of skill can lower the amount of measured discrimination. The result here demonstrates that even if two individuals have exactly the same skill level, the discriminated against individual will have lower income *and* performance. If skill is identified with performance, this would imply that one would not observe discrimination in the data, even though it may exist in the form of biased evaluations.

This result also complements the work of Stephen Coate and Glen Loury (1993) who show that negative stereotypes may cause individuals to invest less in human capital, and hence to be less productive. Here the argument is more direct - given two identical individuals, if the principal believes that one person is less likely to be productive than another, then that person will work less hard and earn less income. Moreover, if the agent is also more likely to disagree with a biased evaluation, this can lead to more conflict, higher costs, and even lower levels of performance.

When there is some bias then both the correlation in beliefs, and the agent's self confidence affect the level of perceived bias. Specifically an increase in correlation results in a decrease in perceived bias, $\partial\alpha/\partial\rho < 0$, and in the limit $\alpha(p, q, 1) = 0$ when $p > 0$. Hence, even in the presence of bias, if both the principal and agent have *common values* regarding what constitutes acceptable performance, then there is no agency cost to implementing any given level of effort. This is consistent with the advice given in the management literature on the importance of communicating performance expectations clearly to employees.[16]

The optimal contract also has the feature that the agent's compensation depends only upon the principal's signal. This is consistent with the results of the previous section, but with a difference, because there is no presumption that the quality of the principal's signal is better than the agent's. In fact if $q > p$ then the agent's evaluation of performance is a strictly superior measure of actual performance that the principal's, yet it is never optimal to make the agent's compensation dependent upon this information. Again, this result is consistent with the common recommendation in the management literature against the use of self-appraisal.

In summary, the results here further reinforce the importance of self-appraisals in lowering the costs of enforcing the optimal contract, though interestingly what is key is not the quality of ones self appraisal, but the correlation of ones evaluation with ones supervisor. Increasing ones self-confidence, as measured by q, results in an increase in perceived bias, α, and hence lower performance and pay. This has some interesting implications for discrimination law. Normally, an individual would only go through the effort of filing a suit if she or he believed that their employer discriminated against them. This would imply that if employee perception of discrimination increased, while employer bias remains unchanged, then the number of discrimination suites would increase. This possibility is consistent with the findings of John J. Donohue III and Peter Siegleman(1991,2001)) who show that the number of suites filed under title VII of the 1964 Civil Rights Act increased during a period when Donohue and Siegelman had expected both discrimination and the number of such suits should have decreased. The results here illustrate that this observation is consistent with a simple optimal contracting framework, and illustrates the importance of both employer and employee beliefs regarding performance for the determination of compensation.

4 IV Discussion

There has always been a tension between the predictions of the standard principal agent model, with its focus on making pay vary with any measure of performance, and the management literature that emphasizes psychological factors, such as trust and fairness. This paper provides a simple and tractable extension of the principal-agent model to incorporate subjective evaluations, resulting in a model that is able to illuminate the role of individual beliefs regarding performance evaluations in a simple optimal contracting framework. First, it is shown that the degree to which pay can depend upon performance depends upon the degree of correlation in beliefs between the principal

[16] See the discussion in Milkovich and Newman (1996), page 374.

and agent. As Milkovich and Wigdor (1991) observe, the effectiveness of performance contracts can be undermined when employees choose to disagree with the assessment of performance by their employer.[17] In an optimal contract, this lack of correlation in beliefs, or equivalently lack of trust, results in performance appraisals that are compressed at the top, a well know feature of performance appraisal systems in practice.[18] This result is also consistent with the recommendations by some management consultants, such as Lawler III (2000), that the number of levels used for discriminating between individuals in terms of performance be kept low.[19]

When applied to the problem of discrimination, it is found that when there is a downward bias in the evaluations by the principal, this results in higher costs and lower performance. Moreover, if the agent is more likely to disagree with a biased evaluation, this can lead to more conflict, higher costs, and even lower levels of performance. This result may explain why it may be misleading to measure discrimination using standard labor market data sets that look for variations in pay between workers with equal performance. As Holzer and Neumark (2000) observe in their review of the literature, obtaining better measures of employee performance typically results in a less measured discrimination, while the results here suggest that discrimination may affect performance, which in turn affects pay.

Finally, the optimal contract decentralizes decision making regarding compensation in a way that is consistent with a traditional authority relationship, namely the principal evaluates performance and determines compensation. Then the agent responds to this evaluation by either accepting it as fair, or engaging in actions that impose costs upon the firm. Despite the abstract nature of the mechanism design exercise, the result does provide some insights into the optimal structure of the compensation contract. In all the cases we consider, the agent's report has no effect on her compensation, rather it is completely determined by the principal's report. In other words, the mechanism is implemented by having the principal make a compensation decision based on his information, with the agent responding by imposing a cost upon the principal should she believe that the rating is unfair. The issue then is how this "conflict" manifests itself in practice?

4.1 IV.A Conflict in Organizations

Much of the existing literature on subjective evaluation has focused upon the question of whether or not it is possible to construct equilibria, or self-enforcing behaviors that punish deviation from an implicit agreement, behavior that is in effect a costly conflict. This literature demonstrates that there are a wide variety of institutions and behaviors that can fulfill this role. For example Bull (1987) shows that when the principal (firm) may have different unobserved types, then good performance is ensured by the desire to build a reputation for honesty and reliability. Michihiro Kandori and Hitoshi Matsushima (1996) characterize the set of equilibria in a repeated game with private information, and show that one can construct efficient equilibria under the appropriate

[17] Page 69.

[18] Milkovich and Newman (1996), page 364.

[19] See page 185 of Lawler III (2000) where he suggests that at most three levels of performance be used.

conditions, while Levin (March 2002) shows how these techniques may be applied to the repeated contracting model of MacLeod and Malcomson (1989).

These results act as a backdrop to the question of what form the optimal contract will take, given that agents have available actions that can impose costs upon both parties. As proposition 2 shows, it would be impossible to provide incentives based upon subjective evaluations of performance if such costs are not possible. It has been shown that the principal can structure the contract to lower the costs associated with conflict, but when the correlation in beliefs is imperfect, such conflict can never be avoided.

One prevalent theme that runs through the management literature on compensation is the problem of reducing organizational conflict. Implicitly, it is viewed as a bad that exists only because irrational employees have potentially inflated views of their performance. The analysis here suggests that while it is desirable to reduce conflict, when there is the potential for differences in opinion, conflict cannot be completely eliminated. The threat of conflict plays a role in ensuring that the principal has an incentive to treat the agent "fairly", which in this case has a very precise definition: the principal is expected to reward the agent as a function of his true assessment of performance. A nice example of exactly this type of behavior is documented in an article by James Stewart (1993), who describes the case of some traders who left First Bank Boston because they felt that their (objectively large) bonuses were unfairly low! Given that these employees easily found employment elsewhere, this evidence suggests that they were in fact valuable to their original employer, who was forced to pay the cost of recruiting and training a set of replacements.

When employees cannot easily leave, then the problem is more complex, and can be related to the problem of maintaining employee morale. For a firm morale can only be an issue if low morale implies some cost for the firm, either in terms of lower quality performance or increased compensation costs. Economists such as George A. Akerlof (1982) and Truman F. Bewley (1995) have argued that employers take these factors into account by either providing high wages that are viewed as a gift (Akerlof (1982)), or by avoiding lowering wages (Bewley (1997)). In other words workers are viewed as "emotional beings" rather than as economic agents, and as such compensation policies must take these apparently irrational sensibilities into account.

This analysis suggests a more economic interpretation of these behaviors, namely they can be viewed as an integral part of a productive and functional relationship. Formally, if an agent does not agree with a low evaluation by the principal, responding by having low morale is functional because it discourages the principal from providing a low evaluation unless he truly believes it to be deserved. Conversely, the fact that the agent will impose costs upon the principal whenever there is a disagreement regarding performance implies that the principal optimally structures compensation to avoid these costs, and is consistent with Bewley (1995)'s observation that firms wish to avoid conflict with their employees. In the extreme, when there is no correlation between the beliefs of the principal and agent, the agent is only punished when the worst signal of performance is observed, a result that is consistent with the compression of employee evaluations that have been observed in practice.

4.2 IV.B Efficiency Wages and Employment Protection Law

The results also provide some insights into the issue of whether individuals should be given employment protection. Some legal scholars, particularly Richard A. Epstein (1984), have argued that the erosion of the doctrine of employment at will undermines the ability of employers to effectively motivate and utilize their workers. One reason is that with employment protection, employers are not able to use efficiency wages to provide incentives. As Alan B. Krueger and Larry H. Summers (1988) observe, one may be able to explain the existence of inter-industry wage differences using the idea that when performance evaluation is subjective, employers can motivate workers by paying above market clearing wages, and firing workers that fall below a certain standard. Such a contract is known as an "efficiency wage contract."

This contract has the feature that the principal's costs are *independent* of the agent's self-evaluation, yet as is shown in proposition 9, such a contract can be optimal only in the case that the agent's beliefs are independent of the principal's beliefs.[20] If there is any correlation at all, then compensation should depend upon the observations of both the principal and the agent. Moreover, efficiency wages as modelled by Carl Shapiro and Joseph E. Stiglitz (1984), have the feature that the firm faces no costs when firing an employee. Instead the employee bears the cost of an unemployment spell, while the employer immediately hires a replacement from the pool of unemployed workers.

When there is some correlation in beliefs such a contract cannot be efficient. Given that it is the agent who imposes costs upon the principal, then if the principal provides the agent with a low evaluation he will have an incentive to fire her to avoid paying those costs. However, if there is employment protection for the agent, this increases the size of the penalty that an agent can inflict upon the principal, and if such costs are *ex ante* optimal, employment protection may increase, rather than decrease, the overall efficiency of the contract. This is consistent with the observations of Charles J. Goetz and Robert E. Scott (1981) that many employment contracts are characterized by the extensive use of employment protection covenants, such as seniority rules in union contracts and institution of tenure for university professors, while ensuring the right of an individual to leave the employment relationship. This latter right gives the agent an additional action that, in the spirit of the example in Stewart (1993), can punish a principal when she believes he has been unfair.

A Appendix: Proofs of Propositions

Proof of Proposition 1. For $\lambda = 0$, the optimal contract is clearly $c_t = u^{-1}(\bar{u}) > 0$ for $t \in T$. For $\lambda \in (0,1)$, the convexity of V implies the IC constraint is equivalent to:

$$V'(\lambda) = u(c)^{\top}\left(\gamma^{H} - \gamma^{L}\right). \tag{17}$$

Given the MLRP, assumption B, there always exists a c satisfying 17. Simply let $c_t = \bar{c}$ for $\gamma_t^H - \gamma_t^L \geq 0$ (some of which are strictly positive since the distributions are not

[20] When s is independent of t, then any allocation of the costs P between states UA and UU is part of an optimal contract as long as $P = \left(P_{UA}\gamma_{UA} + P_{UU}\gamma_{UU}\right)/\left(\gamma_{UA} + \gamma_{UU}\right)$.

the same), and $c_t = f(\bar{c})$ when $\gamma_t^H - \gamma_t^L < 0$, where $f(\bar{c})$ is set sufficiently close to zero such that 17 is satisfied with equality. Notice that $f'(\bar{c}) > 0$, and hence one can then choose $\bar{c}$ so that the individual rationality constraint 3 is satisfied with equality. Finally, from Bertsekas (1974), and the fact that $u(\cdot)$ is concave, $\lim_{c\to 0} u(c) = -\infty$ and $\lim_{c\to\infty} u(c) = \infty$, it follows that the variance of solutions to 2 and 3 are bounded, and given the finite support for consumption, the set of feasible consumption contracts satisfying these constraints form a compact set. Hence $C^*(\lambda)$ exists for every $\lambda \in (0, 1)$.

Let μ_0 be the Lagrange multiplier associated with the individual rationality constraint 2, and μ_1 the Lagrange multiplier associated with the first order condition, then from Holmström (1979) the multipliers are strictly positive ($\mu_0, \mu_1 > 0$) and the optimal contract solves:

$$\frac{1}{u'(c_t)} = \mu_0 + \mu_1 \frac{\left(\gamma_t^H - \gamma_t^L\right)}{\gamma_t(\lambda)}.$$

Letting $r_t = \gamma_t^H / \gamma_t^L$ be the likelihood ratio, observe $\frac{(\gamma_t^H - \gamma_t^L)}{\gamma_t(\lambda)} = \frac{r_t - 1}{\lambda r_t + (1-\lambda)}$, from which it follows that the monotone likelihood ratio condition implies that $\frac{\left(\gamma_{t+1}^H - \gamma_{t+1}^L\right)}{\gamma_{t+1}(\lambda)} > \frac{(\gamma_t^H - \gamma_t^L)}{\gamma_t(\lambda)}$. This combined with $\mu_1 > 0$ and the strict concavity of u implies that c_t is strictly increasing in t. ■

Proof of Proposition 3. Notice that given c_{ts}^* the agent's incentive constraint satisfies:

$$V'(\lambda^*) = \sum_{t,s\in T} u\left(c_{ts}^*\right) P_{ts} \left[\gamma_t^H - \gamma_t^L\right] \tag{18}$$

$$= \sum_{t\in T} \left[\gamma_t^H - \gamma_t^L\right] \sum_{s\in T} u\left(c_{ts}^*\right) P_{ts}. \tag{19}$$

Given that the agent is risk averse there exists a $\hat{c}_t \leq \sum_{s\in T} c_{ts}^* P_{ts}$ such that $u\left(\hat{c}_t\right) = \sum_{s\in T} u\left(c_{ts}^*\right) P_{ts}$. Notice that if the agent is paid $\hat{c}_t$ in each state, then her report s does not affect her payoff, and thus constraint 10 is automatically satisfied.

Let us now show that the optimal contract must necessarily have the property that the agent's compensation does not depend upon s. If it did then for some t it must be the case that $\hat{c}_t < \sum_{s\in T} c_{ts}^* P_{ts}$. Let $\Delta_{ts} = c_{ts} - \hat{c}_t$, which from the previous inequality must solve $\sum_{s\in T} \Delta_{ts} P_{ts} \geq 0$, with strict inequality for some t. One can ensure that under this new contract 11 is satisfied by setting:

$$w'_{ts} = w_{ts} - \Delta_{ts}. \tag{20}$$

The difficulty now is that whenever $\sum_{s\in T} \Delta_{ts} P_{ts} > 0$, the wage payment in state t is reduced, and hence the principal's incentive constraint may no longer be binding, which can be restored by increasing the wage payments when t occurs by $\sum_{s\in T} \Delta_{ts} P_{ts}$, so that the wage payment is:

$$\hat{w}_{ts} = w'_{ts} + \sum_{s\in T} \Delta_{ts} P_{ts}. \tag{21}$$

Thus we have shown that given any optimal contract, we can always find another which gives the same payoff to both the principal and agent, and entails $c_{ts} = c_{t\bar{s}}$ for all $t, s, \bar{s} \in T$. If it is the case that $\sum_{s \in T} \Delta_{ts} P_{ts} > 0$ for all t, then the transformation can be made strictly Pareto improving, with the firm offering a lower wage in every state, while leaving the agent no worse off.

Since $w_{ts} \geq c_t \geq 0$, we can define $\Delta_{ts} = w_{ts} - c_t$, and replace the constraint on w_{ts} by $\Delta_{ts} \geq 0$. Notice that Δ_{ts} is a measure of the social loss. In this case the cost of implementing λ, $C^S(\lambda)$ solves:

$$C^S(\lambda) = \min_{\Delta_{ts}, c_t} \sum_{t \in T} c_t \gamma_t(\lambda) + \sum_{t,s \in T} \Delta_{ts} \Gamma_{ts}(\lambda) \tag{22}$$

$$\sum_{t \in T} u(c_t) \gamma_t(\lambda) - V(\lambda) \geq \bar{u}, \tag{23}$$

$$\sum_{t \in T} u(c_t) \left(\gamma_t^H - \gamma_t^L \right) = V'(\lambda), \tag{24}$$

$$(c_t - c_{\bar{t}}) \gamma_t(\lambda) \leq \sum_{s \in T} (\Delta_{\bar{t}s} - \Delta_{ts}) \Gamma_{ts}(\lambda), \ \forall t, \bar{t} \in T, \tag{25}$$

$$\Delta_{ts} \geq 0. \tag{26}$$

From the proof of proposition 1, the set of contracts satisfying 24 and 23 with equality is non-empty and compact. For any consumption contract $c = [c_1, ..., c_n]$, setting $w_{ts} = \max_{t \in T} c_t$, ensures that the incentive constraints for the firm are satisfied, and sets an upper bound for costs, denoted $\bar{C}(\lambda)$. Let $TS(\lambda) = \{ts \in T \times T | \Gamma_{ts}(\lambda) > 0\}$ denote the support of distribution of evaluations. Given that upper bound on costs, then we have a bound, say $\bar{\Delta}$, for all the Δ_{ts}, $ts \in TS(\lambda)$. This implies that for every state $\bar{t}s \notin TS(\lambda)$, one can at no cost set $\Delta_{\bar{t}s}$ sufficiently large that the constraint:

$$(c_t - c_{\bar{t}}) \gamma_t(\lambda) \leq \sum_{s \in T} (\Delta_{\bar{t}s} - \Delta_{ts}) \Gamma_{ts}(\lambda)$$

is not binding whenever $\Gamma_{ts}(\lambda) > 0$. Given these values for $\Delta_{\bar{t}s}$, the remaining values of c_t and Δ_{ts} solve an optimization problem with a continuous objective function, and compact parameter space, and hence a solution exists. ■

Proof of Proposition 5. Given that $P_{ts} = 0$ for $t \neq s$, the constraint set is not compact, and hence Berge's maximum theorem cannot be applied directly here. From the individual rationality constraint and Dimitri P. Bertsekas (1974) it follows that the consumption set is bounded, and hence compact. Therefore there exists a subsequence k', such that $c_{ts}^{k'}$, converges to say c'_{ts}. Since $C^*(\lambda) \leq C^S(\lambda)$ for any set of beliefs, it must be the case that $C^{k'}(\lambda) \geq C^*(\lambda)$. Given that the probability P_{ts}^k approaches zero when $t \neq s$, it is not possible to guarantee that w_{ts}^k is a bounded sequence. Rather, we shall show that costs must be bounded by a contract $\hat{A}^{k'}$ with the property that the associated costs function $\hat{C}^{k'}(\lambda) \to C(\lambda)$. Given the optimality of contract $A^{k'}$ implies $\hat{C}^{k'}(\lambda) \geq C^{k'}(\lambda) \geq C^*(\lambda)$, and we will be done.

The optimal contract under perfect correlation has a unique consumption contract, c_t^*, with wage payments satisfying $w_{tt}^* = c_t^*$. Select w_{ts}^* when $t \neq s$ to be sufficiently large that the incentive constraint 9 is satisfied with strict inequality. Given

$\lim_{n\to\infty} P_{ts}^{k} = I_{ts}$, this implies that there is an N such that for all $k > N$, 9 is satisfied, and hence the optimal contract is feasible for $k > N$. Let $C^{k'}(\lambda)$ be the corresponding costs. By construction, it has the property that $\lim_{k\to\infty} C^{k'}(\lambda) = C^{*}(\lambda)$, and we are done. ■

Proof of Proposition 6. The discussion preceding the proposition demonstrates $w_t = w_{\bar{t}} \equiv \bar{w}$, for $t, \bar{t} \in T$. The optimal contract is therefore a solution to:

$$C^{NC}(\lambda) \equiv \min_{C\in\Re^{2n^2}} \sum_{t,s\in T} w_{ts}\Gamma_{ts}(\lambda), \tag{27}$$

$$\sum_{t,s\in T} u(c_{ts})\,\Gamma_{ts}(\lambda) - V(\lambda) \geq \bar{u}, \tag{28}$$

$$\lambda \in \arg\max_{\tilde{\lambda}\geq 0} \sum_{t,s\in T} u(c_{ts})\,\Gamma_{ts}\left(\tilde{\lambda}\right) - V\left(\tilde{\lambda}\right), \tag{29}$$

$$\bar{w} - c_t \geq 0,\ t \in T. \tag{30}$$

Let μ_0, μ_1 and β_t be the multipliers for constraints 28, 29 and 30, and hence the Lagrangian for the optimization problem is:

$$L = \bar{w} - \mu_0\left\{\sum_{t\in T} u(c_t)\,\gamma_t(\lambda) - V(\lambda) - \bar{u}\right\} \tag{31}$$

$$-\mu_1\left\{\sum_{t\in T} u(c_t)\left(\gamma_t^{H} - \gamma_t^{L}\right) - V'(\lambda)\right\} - \sum_{t\in T}\beta_t\gamma_t(\lambda)(w - c_t). \tag{32}$$

To simplify the calculation constraint 30 is replaced by $\gamma_t(\lambda)(\bar{w} - c_t) \geq 0$. Now consider a type t such that $w > c_t$, then the complementary slackness condition implies $\beta_t = 0$, which combined with $\partial L/\partial c_t = 0$ implies:

$$\mu_0 = -\mu_1\frac{\left(\gamma_t^{H} - \gamma_t^{L}\right)}{\gamma_t(\lambda)}. \tag{33}$$

In the absence of the incentive constraint, a fixed wage contract would be offered, implying

$$\sum_{t\in T} u(c_t)\left(\gamma_t^{H} - \gamma_t^{L}\right) - V'(\lambda) < 0,$$

and therefore 29 can be replaced by the inequality

$$\sum_{t\in T} u(c_t)\left(\gamma_t^{H} - \gamma_t^{L}\right) - V'(\lambda) \geq 0,$$

and hence $\mu_1 > 0$ whenever $\lambda > 0$. Now from the monotone likelihood condition it follows that 33 can be true for at most one performance level, say t'. For the other performance levels $\bar{w} = c_t$ and there is a $\beta_t \geq 0$ satisfies:

$$\frac{\beta_t}{u'(\bar{w})} = \mu_0 + \mu_1\frac{\left(\gamma_t^{H} - \gamma_t^{L}\right)}{\gamma_t(\lambda)}. \tag{34}$$

This implies that

$$\mu_0 + \mu_1 \frac{\left(\gamma_t^H - \gamma_t^L\right)}{\gamma_t(\lambda)} \geq 0 = \mu_0 + \mu_1 \frac{\left(\gamma_{t'}^H - \gamma_{t'}^L\right)}{\gamma_{t'}(\lambda)},$$

which by the MLRC can only be satisfied if $t' = 1$, the lowest signal.

Therefore the optimal contract takes the form:

$$c_t = \begin{cases} w + b & t > 1, \\ w & t = 1. \end{cases} \tag{35}$$

Using γ_g^k, $k = H, L$, as defined in the statement of the proposition, the incentive constraint implies:

$$u(w+b)\left(\gamma_g^H - \gamma_g^L\right) + u(w)\left(\gamma_1^H - \gamma_1^L\right) = V'(\lambda), \tag{36}$$

from which 14 follows, while the individual rationality constraint implies 15. ■

Proof of Proposition 7. From assumptions A and B, we know that $C^k(\lambda)$ and $C(\lambda)$ are bounded for $\lambda < 1$, which combined with Bertsekas (1974) implies that c_t can be assumed to lie in a compact set. Given that $P_{ts} > 0$, one can choose an N such that $P_{ts}^k \geq \varepsilon > 0$ for all $k > N$, and for all $ts \in T \times T$. This, combined with the fact that consumption must lie in a compact set also implies that w_{ts} can be assumed to lie in a compact set (which is what differentiates this case from the earlier case with perfect correlation). Given that the constraints are continuous in P_{ts}, Berge's maximum theorem implies the statement of the proposition. The addition of *ex post* constraints on the level of costs also implies that w_{ts} can be taken from a compact set, and hence we have continuity in this case as well. ■

Proof of Proposition 8. It has already been shown that at an optimal contract with subjective evaluation c_{ts} is independent of s, and hence the agent's incentive constraints for the revelation of her information are automatically satisfied. Rather than work with w_{ts}, it will be more convenient so choose $\Delta_{ts} = w_{ts} - c_t$, which is constrained to satisfy $S \geq \Delta_{ts} \geq 0$. The principal clearly wishes to set Δ_{ts} to be as small as possibly. This, combined with Assumption D for beliefs will allow a significant simplification of the optimization problem. Consider the problem faced by a principal who observes t and considers reporting t' :

The principal's incentive constraint in this case is:

$$(1-p)(c_t + \Delta_{t0}) + p(c_t + \Delta_{tt}) \leq (1-p)(c_{t'} + \Delta_{t'0}) + p(c_{t'} + \Delta_{t't}).$$

Conditional upon the agent reporting the truth, the cell tt occurs with positive probability only when the agent observes t and reports t. If he observes t', then the agent will report 0 or t', and hence Δ_{tt} occurs only on the left hand side of any incentive constraint, and thus wlog one can set $\Delta_{tt} = 0$. Secondly, the cell $t't$ occurs only when the principal is not truthful, and hence at an optimal contract its value does not affect the expected value of the relationship. Since it only can appear of the right hand side of the incentive constraint, wlog we can set $\Delta_{t't} = S$ whenever $t' \neq t$, and hence the incentive constraint becomes:

$$c_t + (1-p)\Delta_{t0} \leq c_{t'} + (1-p)\Delta_{t'0} + pS. \tag{37}$$

		Agent's Report			
		$1-p$		p - Perfect Correlation	
		0		t	
Principal's Report	t	$c_t + \Delta_{t0}$		$c_t + \Delta_{tt}$	
	$t\prime$	$c_{t'} + \Delta_{t'0}$		$c_t + \Delta_{t't}$	

Table 1: The State Contingent Contract

Define the new variable $\bar{c} = \max_{t \in T} c_t$, that will use as a new choice variable for the principal that must satisfy the constraint $\bar{c} \geq c_t$. Again, since the principle wishes to make Δ_{t0} as small as possible, then it is clear that $\Delta_{t0} = 0$ it $c_t = \bar{c}$. Hence not only must $\bar{c} \leq c_{t'} + (1-p)\,\Delta_{t'0} + pS$ hold for all t', if these inequalities do hold then the inequalities 37 will also be satisfied. Together, this implies that the optimal contract is the solution to the following program:

$$C^p(\lambda) \equiv \min_{c_t, \bar{c}, \Delta_{t0}} \sum_{t \in T} (c_t + (1-p)\,\Delta_{t0})\,\gamma_t(\lambda), \tag{38a}$$

$$\sum_{t \in T} u(c_t)\,\gamma_t(\lambda) - V(\lambda) \geq \bar{u}, \tag{38b}$$

$$\sum_{t \in T} u(c_t)\left(\gamma_t^H(\lambda) - \gamma_t^L(\lambda)\right) \geq V'(\lambda), \tag{38c}$$

$$c_t + (1-p)\,\Delta_{t0} + pS \geq \bar{c},\ \forall t \in T, \tag{38d}$$

$$\bar{c} \geq c_t,\ \forall t \in T, \tag{38e}$$

$$S \geq \Delta_{t0} \geq 0. \tag{38f}$$

Let μ_0 and μ_1 be the Lagrange multipliers for the first two constraints, and β_t and α_t be the multipliers for the next two constraints. The hypothesis of the proposition implies that the last inequality, $S \geq \Delta_{t0}$ is satisfied strictly. If $\Delta_{t0} > 0$, this implies that $\beta_t = \gamma_t(\lambda) > 0$. This has two implications. First, if at the first best contract $c_t^* + pS > \bar{c}$ then this constraint is not binding, and hence $\Delta_{t0} = 0$. Under the assumption on S, this condition holds when p is sufficiently close 1.

As $p \to 0$ constraint 38d must eventually be binding for some t, in which case $\beta_t = \gamma_t(\lambda)$, and $\alpha_t = 0$, and hence in this case the first order condition for c_t^p is:

$$\gamma_t(\lambda) - \mu_0 u'(c_t)\,\gamma_t(\lambda) - \mu_1 u'(c_t)\left(\gamma_t^H(\lambda) - \gamma_t^L(\lambda)\right) - \beta_t = 0, \text{ or}$$

$$\mu_0 = -\mu_1 \frac{\gamma_t^H(\lambda) - \gamma_t^L(\lambda)}{\gamma_t(\lambda)}.$$

From the monotone likelihood ratio property and the arguments in the proof of proposition 6 it follows that this equality can only hold for $t = 1$. For the other types, either

constraint 38d is not binding, in which case:

$$\frac{1}{u'(c_t)} = \mu_0 + \mu_1 \frac{\gamma_t^H(\lambda) - \gamma_t^L(\lambda)}{\gamma_t(\lambda)},$$

or it is the case that $c_t = \bar{c}$, where

$$\frac{1+\alpha_t}{u'(c_t)} = \mu_0 + \mu_1 \frac{\gamma_t^H(\lambda) - \gamma_t^L(\lambda)}{\gamma_t(\lambda)}.$$

Since $\alpha_t \geq 0$, the MLRP implies the existence of the function $t(p)$ in proposition. The existence of $\bar{p}$ follows from the fact that for some $\bar{p} > 0$ it must be the case that 38d is binding for type 2, and will continue to be binding for $p \leq \bar{p}$. ■

Proof of proposition 9. Proof proceeds by first expressing the optimization problem as a linear programing problem, and then showing that the contract in proposition satisfies the appropriate duality conditions. Let the contract be denoted by $\psi = \{w_{ts}, c_{ts}\}_{t,s\in\{A,U\}}$, and observe that without loss of generality we can set $c_{UU} = 0$

and $w_{UU} = 0$ due to the linearity of the constraints are linear. Let Ψ^0 denote the set of contracts with $c_{UU} = 0$ and $w_{UU} = 0$. These values can then be rescaled once a solution is found to satisfy the IR constraint. Suppose that the expected payment to the agent when there is a success must be at least $\bar{c}$, which also fixes λ via the IC constraint:

$$\bar{c} = V'(\lambda).$$

Thus the optimization problem becomes:

$$\min_{\psi\in\Psi^0} \lambda\left(w_{AA}\gamma_{AA} + w_{AU}\gamma_{AU} + w_{UA}\gamma_{UA}\right) \tag{39}$$

subject to

$$c_{AA}\gamma_{AA} + c_{AU}\gamma_{AU} + c_{UA}\gamma_{UA} \geq \bar{c} \tag{40}$$

$$w_{kA}\gamma_{kA} + w_{kU}\gamma_{kU} \leq w_{lA}\gamma_{kA} + w_{lA}\gamma_{kA}, \text{ for } k,l \in \{A,U\} \tag{41}$$

$$c_{Ak}\gamma_{Ak} + c_{Uk}\gamma_{Uk} \geq c_{Al}\gamma_{Ak} + c_{Al}\gamma_{Ak}, \text{ for } k,l \in \{A,U\} \tag{42}$$

$$w_{ij} + c_{ij} \leq 0,\ i,j \in \{A,U\} \tag{43}$$

For $\lambda > 0$ this problem can be restated as a linear programing problem of the form:

$$\max_{y\in\Re^6} a'y \tag{44}$$

$$\text{subject to : } Ay \leq \hat{c} \tag{45}$$

by letting the choice variable and parameters be

$$y = [w_{AA}, w_{AU}, w_{UA}, c_{AA}, c_{AU}, c_{UA}]^\top \tag{46}$$

$$a = \left[-\gamma_{AA}, -\gamma_{AU}, -\gamma_{UA}, 0, 0, 0\right]^\top \tag{47}$$

$$\hat{c} = [-\bar{c}, 0, 0, 0, 0, 0, 0, 0]^\top \tag{48}$$

and the matrix A is given by:

$$A = \begin{bmatrix} 0 & 0 & 0 & -\gamma_{AA} & -\gamma_{AU} & -\gamma_{UA} \\ \gamma_{AA} & \gamma_{AU} & -\gamma_{AA} & 0 & 0 & 0 \\ -\gamma_{UA} & -\gamma_{UU} & \gamma_{UA} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\gamma_{AA} & \gamma_{AA} & -\gamma_{UA} \\ 0 & 0 & 0 & \gamma_{AU} & -\gamma_{AU} & \gamma_{UU} \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}. \tag{49}$$

The first row is the negative of the payoff to the agent when there is a success, a constraint that is always binding. The second and third rows are the incentive constraints for the high and low type principals, while the fourth and fifth rows are the similar incentive constraints for the agent. The last three rows are the budget constraints.

We solve this problem by showing that the contract given in the theorem solves this linear programing problem, that is $y^* = [b, b, P, b, b, 0]$. By the complementary slackness theorem this is optimal if there is an $x \in \Re^8_+$ such that $A^\top x = a$ and $x_i v_i = 0$ for every i, where $v = c - Ay^*$. Notice that for y^* the incentive constraints for the agent are satisfied automatically, and hence $v_4 = v_5 = 0$. The principal will set total compensation as low as possible hence $b = \bar{c}/\gamma_{A*}$, where $\gamma_{A*} = \gamma_{AA} + \gamma_{AU}$ is the probability that the principal has a high signal. This implies that $v_1 = 0$. The penalty P is to provide an incentive to the principal to reveal that she has observed a good signal. Since it involves a social cost, then it will be made as small as possible to ensure that the principal's incentive constraint is binding, or

$$-b \geq -P\gamma_{AA}/\gamma_{A*}, \text{ implying} \tag{50}$$

$$P = \gamma_{A*} b/\gamma_{AA} = \bar{c}/\gamma_{AA}. \tag{51}$$

This implies that $v_2 = 0$. Notice that $\gamma_{AA}\gamma_{UU} - \gamma_{UA}\gamma_{AU} > 0$ implies that the principal's second constraint is automatically satisfied. The first of the two budget constraints is satisfied with equality, and hence $v_6 = v_7 = 0$. Therefore we need to find an $x^* = [x_1, x_2, 0, x_4, x_5, x_6, x_7, 0]^\top \geq 0$ satisfying $A^\top x = a$. When $\gamma_{AU} > 0$ the latter has a unique solution given by:

$$\begin{aligned} x_1 &= \tfrac{\gamma_{UA}+\gamma_{AU}}{\gamma_{AU}} \\ x_2 &= \tfrac{1}{\gamma_{AA}}\gamma_{UA} \\ x_4 &= \gamma_{UA}\gamma_{AU}\tfrac{\gamma_{UA}+\gamma_{AA}}{\gamma_{AA}\gamma_{UU}-\gamma_{AU}\gamma_{UA}} \\ x_5 &= \gamma_{UA}\tfrac{\gamma_{AA}+\gamma_{UA}}{(\gamma_{AA}\gamma_{UU}-\gamma_{AU}\gamma_{UA})} \\ x_6 &= \gamma_{AA} + \gamma_{UA} \\ x_7 &= \gamma_{AU}\tfrac{\gamma_{UA}+\gamma_{AA}}{\gamma_{AA}} \end{aligned}, \tag{52}$$

all of which are strictly positive under the hypothesis that $\gamma_{AA}\gamma_{UU} - \gamma_{AU}\gamma_{UA} > 0$.

If $\gamma_{AU} = 0$ the optimal contract has the same form, except that when the agent has a high evaluation he has strict incentives to reveal his information, implying that $v_4 \neq 0$, and hence we need to allow $x_4 \geq 0$. In addition since the principal receives zero from

the cell AU, this implies that both of her incentive constraints are binding, and hence we must now allow $x_3 \geq 0$. Thus we must find $x^* = [x_1, x_2, x_3, 0, x_5, x_6, x_7, 0]^\top \geq 0$ satisfying $A^\top x = a$, with $\gamma_{AU} = 0$. The solution is:

$$\begin{array}{c} x_1 = \frac{\gamma_{UA}+\gamma_{AA}}{\gamma_{AA}} \\ x_2 = \frac{1}{\gamma_{AA}}\gamma_{UA} \\ x_3 = 0 \\ x_5 = \frac{1}{\gamma_{UU}}\gamma_{UA}\frac{\gamma_{UA}+\gamma_{AA}}{\gamma_{AA}} \\ x_6 = \gamma_{UA} + \gamma_{AA} \\ x_7 = 0 \end{array} . \tag{53}$$

This demonstrates that the optimal contract takes the form of a bonus to the agent whenever the principal has a high signal. The only role played by the agent's signal is to provide incentives for truthful revelation by the principal through the imposition of the penalty P. The incentive constraint for the agent's effort satisfies $V'(\lambda) = \gamma_{A*}b$, yielding the bonus equation. From equation 51 one gets the equation for P. The individual rationality constraint implies that:

$$w + \lambda\gamma_{A*}b - V(\lambda) = U^0,$$

from which we obtain the expression for the wage. The cost function is given by the wage costs plus the expected cost from the imposition of P :

$$C(\lambda) = U^0 + V(\lambda) + \lambda\gamma_{UA}P,$$

yielding the final expression in the proposition. ■

References

Akerlof, George A., "Labor Exchange as Partial Gift Exchange," *Quarterly Journal of Economics*, November 1982, *97* (4), 543–69.

Altonji, Joseph G. and Rebecca M. Blank, "Race and Gender in the Labor Market," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics Volume 3C*, Amsterdam: North-Holland, 1999, pp. 3143–3260.

Baker, George, Robert Gibbons, and Kevin J. Murphy, "Subjective Performance Measures in Optimal Incentive Contracts," *The Quarterly Journal of Economics*, November 1994, *109* (439), 1125–1156.

Bertsekas, Dimitri P., "Necessary and Sufficient Conditions for Existence of an Optimal Portfolio," *Journal of Economic Theory*, 1974, *8*, 235–247.

Bewley, Truman F., "A Depressed Labor Market as Explained by Participants," *American Economic Review*, May 1995, *85* (2), 250–254.

Bull, Clive, "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 1987, *102*, 147–159.

Coate, Stephen and Glen Loury, "Will Affirmative-Action Policies Eliminate Negative Stereotypes?," *American Economic Review*, 1993, *85* (5), 1220–1240.

Donohue III, John J. and Peter Siegelman, "The Changing Nature of Employment Discrimination Litigation," *Stanford Law Review*, May 1991, *43*, 983–1033.

____ , "The Evolution of Employment Discrimination Law in the 1990s: An Empirical Investigation," May 2001. Mimeo Stanford Law School.

Epstein, Richard A., "In Defense of the Contract at Will," *The University of Chicago Law Review*, 1984, *51* (947), 947–982.

Goetz, Charles J. and Robert E. Scott, "Principles of Relational Contracts," *Virginia Law Review*, 1981, *67*, 1089–1119.

Goldin, Claudia and Cecilia Rouse, "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians," *American Economic Review*, September 2000, *90* (4), 715–41.

Green, Jerry R. and Jean-Jacques Laffont, *Incentives in Public Decision Making*, Amsterdam: North-Holland, 1979.

Grossman, Sanford J. and Oliver D. Hart, "An Analysis of the Principal-Agent Problem," *Econometrica*, January 1983, *51*, 7–46.

Harris, Milton and Arthur Raviv, "Optimal Incentive Contracts with Imperfect Information," *Journal of Economic Theory*, 1979, *20*, 231–259.

Holmström, Bengt, "Moral Hazard and Observability," *Bell Journal of Economics*, 1979, *10*, 74–91.

____ and Paul Milgrom, "Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 1991, *7*, 24–52.

Holzer, Harry and David Neumark, "Assessing Affirmative Action," *Journal of Economic Literature*, September 2000, *38*, 483–568.

Kahn, Lawrence M., "Discrimination in Professional Sports: A Survey of the Literature," *Industrial and Labor Relations Review*, April 1991, *44* (3), 395–418.

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

____ and Hitoshi Matsushima, "Private Observation, Communication and Collusion," *Econometrica*, May 1998, *66* (3), 627–652.

Krueger, Alan B. and Lawrence H. Summers, "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, March 1988, *56* (2), 259–294.

Levin, Jonathan, "Relational Incentive Contacts," March 2002. forthcoming American Economic Review.

MacLeod, W. Bentley, "On Optimal Contracting with Subjective Evaluation," May 2001. CLEO Discussion Paper, 01-11, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=276122.

____ and Daniel Parent, "Jobs Characteristics and the Form of Compensation," *Research in Labor Economics*, 1999, *18*, 177–242.

____ and James M. Malcomson, "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, March 1989, *57* (2), 447–480.

McAfee, R. Preston and Philip J. Reny, "Correlated Information and Mechanism Design," *Econometrica*, March 1992, *60* (2), 395–421.

Milkovich, George T. and Alexandra K. Wigdor, *Pay for Performance: Evaluating Performance and Appraisal Merit Pay*, Washington, D.C., U.S.A.: National Academy Press, 1991.

____ and Jeffrey M. Newman, *Compensation*, Chicago, IL: Irwin, 1996.

Moore, John, "Implementation, Contracts, and Renegotiation in Environments with Complete Information," in Jean-Jacques Laffont, ed., *Advances in Economic Theory: Sixth World Congress*, Vol. I, Cambridge, UK: Cambridge University Press, 1992, chapter 5, pp. 182–282.

Lawler III, Edward E., *Rewarding Excellence*, San Francisco: Jossey-Bass Inc., 2000.

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

Prendergast, Canice, "A Theory of "Yes Men"," *American Economic Review*, September 1993, *83* (4), 757–770.

____ , "The Provision of Incentives in Firms," *Journal of Economic Literature*, March 1999, *36* (1), 7–63.

____ **and Robert H. Topel**, "Favoritism in Organizations," *Journal of Political Economy*, 1996, *104* (5), 958–978.

Radner, Roy, "Repeated Principal-Agent Games with Dsicounting," *Econometrica*, September 1985, *53* (5), 1173–1198.

Shapiro, Carl and Joseph E. Stiglitz, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, June 1984, *74* (3), 433–444.

Stewart, James B., "Taking the Dare," *The New Yorker*, 1993, pp. 433–444. July 23.

Williamson, Oliver E., Michael L. Wachter, and Jeffrey E. Harris, "Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange," *Bell Journal of Economics*, Spring 1975, *6* (1), 250–278.