What They Don't Know Can't Hurt You: The Benefits of Limited Feedback in Organizations

Stephen Hansen*
London School of Economics
JOB MARKET PAPER

February 4, 2009

Abstract

Most firms formally appraise the performance of their employees, but there is widespread evidence that appraisals are not as informative as they could be. Many observers have concluded that organizational dysfunction or human psychology explains limited evaluation. This paper instead provides an economic rationale for why a firm would want to commit to sharing some but not all its private information about workers' output with workers themselves. When a worker has career concerns, feedback generates two effects. First, it creates uncertainty about future effort levels and exposes him to effort risk. Second, it allows him to use current effort to influence the amount of effort the firm expects him to exert in the future, generating coasting incentives. With optimal feedback, no firm identifies the worst performers, and firms in which talent is relatively unimportant do not identify the best performers. The paper also identifies an inefficiency in the provision of information: under reasonable conditions, firms in industries in which talent is important oversupply it. It concludes by relating its findings to observed inter-industry patterns of performance appraisal use.

Keywords: Performance Appraisal, Career Concerns, Incentives, Risk

JEL Codes: D82, D86, M12, M55

---

*Address: London School of Economics, STICERD, Houghton Street, London WC2A 2AE, United Kingdom. Email: `s.hansen@lse.ac.uk`. Web: `http://personal.lse.ac.uk/hansens`. I am exceedingly grateful to my supervisors Gilat Levy and Andrea Prat for their support and encouragement throughout the project. The advice of Luis Garicano has also been invaluable. I also thank the Management and Strategy Department at the Kellogg School of Business for their hospitality while I developed the paper. Special acknowledgements go to Mike Gibbs, who shared two datasets with me, Oliver Denk, who provided extensive comments on earlier drafts, and Selina Gann, who helped clarify my arguments on numerous occasions. Oriana Bandiera, Heski Bar-Isaac, Tim Besley, Alex Bryson, Nathan Foley-Fisher, Yuk-Fai Fong, Tom Hubbard, Jin Li, Michael McMahon, Hannes Mueller, Marco Ottaviani, Canice Prendergast, John Sutton, John Van Reenen, and Mike Whinston also provided helpful suggestions. I appreciate the financial support I have received from the Bank of England throughout my PhD. The usual disclaimer applies.

1 Introduction

Most organizations have performance appraisal systems in place that allow managers to provide feedback to workers at regular intervals. Such arrangements have existed for centuries, and their use has grown to the point that they are now a common feature of the modern workplace.[1] Not only are they ubiquitous, but performance appraisals consume vast amounts of managerial time, both for the human resource professionals that design and administer reviews as well as for the front line managers who actually conduct them.[2]

A typical firm conducts annual or semi-annual performance reviews in which supervisors give numerical ratings to the workers they oversee. Figure 1 displays a rating distribution from a medium-sized service firm in the United States.[3] In this firm, 1 is the rating associated with highest performance, and 5 the worst.

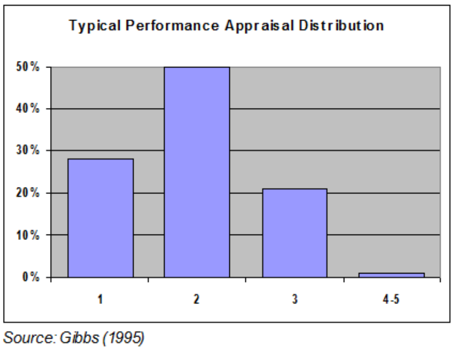

Figure 1: Example of Rating Distribution

In this firm, almost no workers receive negative feedback: 4 and 5 make up just one percent of the sample. Moreover, there is a marked central tendency in the distribution, with fully fifty percent of workers receiving the rating 2. At the very least, one can conclude that managers in this firm do not differentiate between levels of performance as much as the rating scale allows them to. While this distribution comes from only one firm, other firms' rating distributions exhibit similar patterns (Medoff and Abraham 1980, Murphy 1992).

Knowing the true distribution of performance in the firm is impossible, so figure 1 taken by itself does not allow one to reach any definite conclusions about feedback.

[1]Performance appraisal systems were in place by 300 AD in the Chinese state bureaucracy. As of the early 1980's, between seventy four and eighty nine per cent of American businesses used them (Murphy and Cleveland 1991).

[2]The Chief Human Resource Counsel for International Paper recently noted that "...few tasks occupy as much time by human resource professionals as designing, implementing, monitoring, and defending performance appraisal systems" (Murphy and Margulies 2004).

[3]This figure is taken directly from Lazear and Gibbs (2008). The data have previously been analyzed in Baker, Gibbs, and Holmstrom (1994a), Baker, Gibbs, and Holmstrom (1994b), and Gibbs (1995), among others.

However, there is evidence that it reflects managers' hiding private information from workers. Several studies have shown that the ratings that supervisors report to workers are significantly higher and more skewed than the ratings they report to independent researchers (see Murphy and Cleveland 1991, p.79, and references therein). Also, the same patterns emerge when rating categories have labels such as "average" and "below average" (Gibbs 1991).[4] Finally, workers and managers themselves report that managers do not distinguish among workers.[5]

The economics and HRM literature has identified several reasons for why there might be limited feedback in organizations. First, managers are very often not rewarded for providing accurate appraisals, so they might not exert the effort required to assess and document the performance of their workers (Baron and Kreps 1999). Second, managers may exhibit favoritism and bias ratings (Prendergast and Topel 1996). Third, organizational politics may constrain truthtelling.[6] Finally, worker psychology may discourage honesty.[7] In short, many argue that either a prinicpal-agent problem between those wanting to use information (an HR office) and those able to gather it (front-line managers), or else the fallibility of human psychology and relationships, limits information flow in organizations.

These views implicitly assume that limited communication is necessarily the result of organizational dysfunction. However, without a proper understanding of the effects of information disclosure and their relationship to worker motivation and wages, it is impossible to determine what is the "right" level of feedback. The goal of this paper is both to identify the effects of feedback and to derive the optimal feedback policy to which the firm would like to commit. It finds that in a wide variety of situations,

[4] In one particularly stark example, Milkovich, Newman, and Milkovich (2007) report a ten-year study of a thousand-member social service department in which only three of the possible ten thousand ratings were "below average".

[5] In a case study of Merck, Murphy (1992) reports such sentiments as "Tell me this, how in the world can 83 per cent of the people be exceeding job expectations while the company, as a whole, is doing just average?" and "How can I rate my people objectively when the other directors are giving all their people 4s? A 3 isn't acceptable. I wouldn't mind if everyone played by the same rules, but they don't."

[6] For example, Longenecker, Sims, and Gioia (1987) write that

> ...it is likely that political considerations influence executives when they appraise subordinates. Politics in this sense refers to deliberate attempts by individuals to enhance or protect their self-interests when conflicting courses of action are possible. Political action therefore represents a source of bias or inaccuracy in employee appraisal.

[7] For example, Jackman and Strober (2003), writing about performance appraisal, state that workers

> ...hate being criticized, plain and simple. Psychologists have a lot of theories about why people are so sensitive to hearing about their own imperfections. One is that they associate feedback with the critical comments received in their younger years from parents and teachers.

optimal feedback policies never explicitly identify poor performance, but instead pool all poorly performing workers together. Moreover, some firms never explicitly identify good performance, meaning that all workers infer their performance to be in the middle of the distribution. However, the paper also shows that firms would always like to commit to share some of their private information with workers, meaning that providing feedback to workers is in fact a source of value for organizations through its effect on motivation.

Performance appraisal is most relevant in situations in which output is subjective. Moreover, with subjective output, work incentives arise through implicit incentives since output is non-verifiable. The paper therefore examines information disclosure in a principal-agent model in which the agent has career concerns (the paper will hereafter refer to the principal as the firm and the agent as the worker). The worker (he) exerts effort for two periods, after which he earns a fixed reputational reward if his expected talent surpasses a threshold; however, only the firm (it) observes output. However, before the employment relationship begins, the firm can commit to disclosing a set of its posterior beliefs on worker talent to the worker between his first and second period effort choices. There are two main effects of feedback:

1. *Effort risk.* Feedback creates uncertainty about future effort levels. Depending on the feedback that the worker receives he can either find himself working more or less hard in the second period. If he does not receive feedback, then his second period effort level is equal to the average of all possible levels under feedback. Since the worker's preferences over effort levels are given by his (convex) cost function, he would rather work some given amount for certain than the same amount in expectation. So, feedback exposes the worker to effort risk, which increases his disutility of effort.

2. *Coasting incentive.* As in all career concerns models, signal jamming provides effort incentives in the first and second period as the worker seeks to interfere directly with the employer's learning about his ability. Feedback introduces an additional motive for first period effort.[8] Whenever the worker receives feedback, the firm's belief about the amount of effort he will exert in the second period depends on his first period performance since it anticipates he will tailor his effort choice to the probability of promotion. But this implies that the worker can use first period effort to reduce the amount of effort the firm expects him to exert in the second period, which increases the degree to which it attributes second period output to his talent. In summary, the worker wants to use first period effort to trick the firm into thinking he will not work hard—or coast—in the second.

[8] In the model, expected second period effort is independent of the disclosure policy. Hence, an important feature of the model is that the *anticipation* of feedback creates effort incentives.

After identifying these effects of feedback, the paper solves for the optimal disclosure policy. This derives from the firm's profit maximization problem, and depends crucially on the expected future payoff from joining the firm. When it is large, the participation constraint never binds for any disclosure policy, and the firm extracts as much effort as possible from the worker through choosing a disclosure policy that maximizes coasting incentives. It does so by only giving feedback when its posterior on worker talent is sufficiently high. When the expected payoff is low, the worker needs to be compensated to join the firm and his participation constraint always binds. In this case, increasing feedback increases the up-front wage the firm must pay the worker, and the firm only reveals an intermediate range of talent beliefs to the worker.

The paper then endogenizes the wage schedule and the size of the expected future payoff from joining the firm using a model of adverse selection labor market competition. The main result is that a firm offers a high payoff if the productivity of experienced workers in its industry is sufficiently sensitive to talent.

The paper concludes by identifying the industries in which performance appraisal is most widespread using a cross-sectional survey of firms in the United Kingdom, and finds that professional service industries occupy the majority of the top places. While this evidence is consistent with the model, there are other explanations for why such a pattern would arise. The paper therefore distinguishes its rationale for giving feedback from others, and suggests a statistical test that would allow one to identify whether feedback affects motivation when workers have career concerns.

The importance of the paper is that it studies feedback in organizations in a fully micro-founded, rational choice model with standard economic preferences and technology, and yet shows that optimal feedback policies have properties that scholars often view as dysfunctional. Interestingly, it shows that if anything one should worry about certain firms giving *too much* feedback and failing to account for the social cost of their actions in the form of increased effort risk for the worker. In any case, a firm would never want to provide fully informative feedback, and one that did would face de-motivated workers and difficulty in attracting new talent.

Related Literature The paper makes three distinct contributions to the literature. First, it analyzes general disclosure policies on a continuous output space and allows for endogenous compensation. Second, it shows that rating compression and avoidance of negative feedback are compatible with optimal feedback. Third, it examines how the worker's information affects relationships with career concerns.

Several recent papers (Aoyagi 2007, Ederer 2008, Goltsman and Mukherjee 2008) have examined effort maximizing disclosure policies in two period tournaments with two competitors. Aoyagi (2007) relates the optimal disclosure policy to the cost of effort function and finds that if the marginal cost of effort is convex, no disclosure is optimal; if the marginal cost is concave, full disclosure is optimal; and if the marginal cost is linear, all disclosure policies yield equivalent expected effort. Ederer (2008) uses a similar framework, but adds ability into the production function. In this environment, information disclosure can provide effort incentives because it allows a worker to signal his ability to his competitor. When ability and effort are complementary in production and costs are quadratic, full information disclosure is optimal under certain distributions. Goltsman and Mukherjee (2008) restrict production to only taking two values. They find that the optimal disclosure policy reveals no information to the contestants unless both produce a low output in the first period, in which case this outcome is told to both of them. A limitation of these papers is that they do not endogenize the tournament prize, nor the agents' initial compensation. Both ex ante and ex post compensation play an important role in this paper.

Lizzeri, Meyer, and Persico (2002) study optimal disclosure in a two period moral hazard problem in which the principal can offer output contracts. With full disclosure, first period effort is always higher whenever the wage function is non-linear, but expected second period effort costs are higher, like in this paper. When the principal can choose optimal compensation, the second effect dominates and it never reveals information to the worker. This result stands in contrast to the empirical observation that some, albeit limited, feedback appears in most organizations.

MacLeod (2003) looks at how a principal optimally uses its subjective assessment of worker output with static relational contracting. He shows that the provision of incentives entails social waste, and that the waste-minimizing contract pays the agent a fixed wage for all output signals except the one most informative about low output. Fuchs (2007) extends these results to a repeated relationship. He finds that when the principal-agent relationship is finite, the principal gives no feedback to the agent until the last period, and then only if the agent produces a low output in each period. When he is kept in the dark about his performance, the agent attaches a higher weight to being in the bad state, so he keeps working hard to avoid money burning.[9] The focus in both these papers is how to sustain work incentives while keeping social costs low. In this paper, signal jamming incentives induce effort, and do not entail waste. Moreover, the paper assumes the principal commits ex ante to a disclosure policy, so does not impose truth-telling constraints.

The paper also builds on the career concerns literature initiated by Fama (1980)

[9]The result is related to Abreu, Milgrom, and Pearce (1991), who show that delaying the accumulation of information about actions can improve outcomes in repeated games with private monitoring.

and Holmstrom (1999). Several papers have found circumstances under which more information about an agent's behavior harms the principal. For example, in Holmstrom (1999), increasing the precision of the principal's belief on the agent's talent reduces his effort. Dewatripont, Jewitt, and Tirole (1999) provide conditions under which an agent can exert more effort when the prinicpal knows less about his output. In a more recent paper, Kovrijnykh (2007) shows that delaying the release of information about worker performance to the labor market can reduce oversupply of effort in a dynamic career concerns model. When agents care about signalling expertise rather than effort, Prat (2005) has demonstrated how an improvement in incentives arises when the principal does not observe the action the agent takes. In contrast to all these papers, this one shows that limiting the amount of information that the *agent* has about his action can help the principal when the agent has career concerns. Another related paper in the literature is Martinez (2008), who shows that current effort can effect the firm's future beliefs about worker talent. However, he does not explore the relationship of this effect to information disclosure.

The rest of the paper proceeds as follows. The next section presents and motivates the model. The third section solves for the equilibrium and identifies the effects of information disclosure. The fourth section then introduces the profit maximization problem for the firm, derives the optimal feedback policy, discusses its properties, and examines robustness. The fifth section endogenizes the rewards to talent schedule through labor market competition and connects feedback to production technology. The sixth discusses the model in light of observed inter-industry variation in performance appraisal use and discusses possible extensions. Unless otherwise stated, all unproven results in the text are proved in the Appendix.

2 Model

2.1 Setup

There are four time periods $t = 0, 1, 2, 3$. A risk neutral firm F and a risk neutral, liquidity constrained worker meet in period 0 to determine the contract that will define their relationship over periods 1 and 2. Their relationship must last two periods, and neither party can break from the other after period 1. In periods 1 and 2, the worker produces $y_t = \theta + e_t + \varepsilon_t$, where θ is talent, e_t is effort, and $\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right)$ is an output shock uncorrelated across time periods and with θ. Neither the worker nor the firm knows θ at period 0, but they share a common prior distribution $N\left(\overline{\theta}, \sigma_\theta^2\right)$ on it, where $\overline{\theta} > 0$.

The cost to the worker of exerting effort is $g(e_t) = \frac{C}{2}e_t^2$,[10] and he has an outside option of u, which can be thought of as the utility of leisure or the wage he could receive in another industry.

As in standard principal-agent models, the worker privately observes e_t. Unlike in the standard model, though, the firm privately observes y_t. After observing $\{y_\tau\}_{\tau=1}^{t}$, the firm forms belief $\widehat{\theta}_t^F$ on worker talent. After period 2, F continues to employ the worker if $\widehat{\theta}_2^F \geq \theta^*$. The rewards to talent schedule for the worker is given by

$$w\left(\widehat{\theta}_2^F\right) = \begin{cases} \overline{W} \text{ if } \widehat{\theta}_2^F \geq \theta^* \\ \underline{W} \text{ if } \widehat{\theta}_2^F < \theta^* \end{cases}$$

where $\overline{W}$ is the utility the worker receives from continued employment with the firm, and $\underline{W}$ is his best outside option if the firm releases him, so that $W = \overline{W} - \underline{W} > 0$ is the net return from continued employment. The results of the paper depend quite heavily on this particular wage function, so although the model takes it as exogenous for now to simplify the analysis, it later derives it as the outcome of adverse selection labor market competition.

A contract in period 0 consists of two objects. The first is a payment w that F offers the worker to attract him into employment. Because the worker is liquidity constrained, this payment must be non-negative.

Assumption 1 $w \geq 0$.

The second is a disclosure policy that gives the worker information about $\widehat{\theta}_1^F$ before his choice of e_2. In fact, disclosing $\widehat{\theta}_1^F$ is equivalent to disclosing y_1, but since $\widehat{\theta}_1^F$ is more important for the strategic effects in the model, the paper chooses to focus on it.

Definition 1 *A disclosure policy is a mapping ψ such that*

$$\psi\left(\widehat{\theta}_1^F\right) = \begin{cases} \widehat{\theta}_1^F \textit{ if } \widehat{\theta}_1^F \in \Theta \\ \emptyset \textit{ if } \widehat{\theta}_1^F \notin \Theta \end{cases}$$

where the firm discloses $\widehat{\theta}_1^F$ to the worker if and only if $\widehat{\theta}_1^F \in \Theta$.

The structure of a disclosure policy is that the firm communicates to the worker exactly its belief about his talent (in which case $\psi\left(\widehat{\theta}_1^F\right) = \widehat{\theta}_1^F$), or else says nothing at all (in which case $\psi\left(\widehat{\theta}_1^F\right) = \emptyset$).[11] Choosing a disclosure policy is equivalent to choosing

[10] The robustness of the model to this cost structure is discussed in section 4.4

[11] Of course, the outcome $\emptyset$ still has informational content, since the worker learns that the firm's beliefs do not lie in Θ.

Θ, the set of beliefs revealed to the worker. Since disclosure policies that differ at points with measure zero are equivalent, choosing Θ is equivalent to choosing a set of the form $\cup_i (\underline{x}_i, \overline{x}_i)$ such that $(\underline{x}_i, \overline{x}_i) \subseteq \mathbb{R}$ and $\overline{x}_i < \underline{x}_{i+1}$. In other words, Θ is a union of non-overlapping convex subsets of the real number line. While more general specifications of ψ are possible, this definition does allow for flexibility in the sense that the full disclosure policy $\Theta = (-\infty, \infty)$ and the no disclosure policy $\Theta = \emptyset$ are just two extremes of a continuum of admissible forms of ψ.[12]

An important assumption is that F commits to a disclosure policy in period 0 that it cannot overturn after observing y_1. This ensures the most favorable conditions possible for communication between the firm and the worker. If performance appraisals were limited in their informational content because of difficulties the firm had in credibly communicating its private information, then this would constitute a problem for the firm and limit its profit. If on the other hand the firm wants to actively *commit* to limiting information flow, then performance appraisals with limited communication are actually best for the organization.[13]

Definition 2 *The informativeness of a disclosure policy* Θ *is* $\Pr\left(\widehat{\theta}_1^F \in \Theta\right)$.

Informativeness is a measure of how much feedback a disclosure policy provides. Because F discloses $\widehat{\theta}_1^F$ completely or not at all, a natural measure of the quantity of feedback a disclosure policy gives is the probability that $\widehat{\theta}_1^F$ falls in the set of disclosed beliefs.

The paper also distinguishes between two kinds of disclosed interim beliefs.

Definition 3 *Positive feedback is*

$$\left\{\widehat{\theta}_1^F | \widehat{\theta}_1^F \in \Theta, \widehat{\theta}_1^F > \theta^*\right\}$$

and negative feedback is

$$\left\{\widehat{\theta}_1^F | \widehat{\theta}_1^F \in \Theta, \widehat{\theta}_1^F < \theta^*\right\}.$$

When the disclosure policy reveals to the worker that $\widehat{\theta}_1^F > \theta^*$, the worker finds out that earning the reputational reward W is relatively likely, while if he learns that $\widehat{\theta}_1^F < \theta^*$, he learns that it is not.[14] Positive feedback thus constitutes good news about the worker's

[12] Section 4.4 analyzes more general disclosure policies.

[13] The paper later shows that F would indeed like to lie to worker after observing y_1. One possible avenue for deriving Θ as the equilibrium of a strategic communication game between the firm and the worker would be to introduce more periods into their relationship, so that the firm's lying could be found out by the worker and suitably punished.

[14] In equilibrium, knowing $\widehat{\theta}_1^F > \theta^*$ leads the worker to estimate that his probability of earning W is greater than one half, while knowing $\widehat{\theta}_1^F < \theta^*$ means the estimated probability is less than one half.

future prospects of working with F, and negative feedback bad news. These two different kinds of feedback will have important implications for the optimal disclosure policy.

One important assumption is that only F and the worker observe $\psi\left(\widehat{\theta}_1^F\right)$. This essentially rules out a contract that depends explicitly on output, since there is no way for a court to enforce such a contract if it cannot observe any information about y_1. In periods 1 and 2, therefore, only implicit incentives exist for the worker, and the paper assumes that these come from career concerns.

F chooses ψ and w to maximize expected profit, given by $E\left[y_1 + y_2\right] - w$. To summarize, the timing of the game is the following.

1. F offers the worker ψ and w.

2. The worker chooses e_1.

3. The outcome of ψ is revealed to the worker.

4. The worker chooses e_2.

5. If $\widehat{\theta}_2^F \geq \theta^*$ the worker continues to work for F and earns net reward W.

2.2 Motivation for setup

The model provides a natural description for a professional service industry such as management consulting, law, or accounting. Maister (1993) describes the defining characteristics of these industries in detail, and the paper uses his analysis as a benchmark with which to compare the model. Most professional service firms are partnerships with two or three layers. On top, there are the partners themselves, who engage in high-value activities, enjoy high earnings, and have relative job security, while the bottom layer(s) are composed of less experienced workers engaging in lower value tasks. These firms "make few, if any, performance differentials in their compensation of junior staff" (Maister 1993, p.196), and work incentives for juniors derive from the desire for promotion to partner. Moreover, most service firms operate an up-or-out system (either implicitly or explicitly), in which junior workers leave the firm if they have failed to reach partner within a fixed time frame. The assumption that career concerns are the main motivator for workers, and that there is a final period in which reputational rewards accrue to the most talented workers is thus highly realistic.

The paper eventually models the period 3 reward for the worker as the outcome of labor market competition for talent, yet assumes that a particular firm extracts surplus from him in period 0. The motivation for this assumption is that junior staff in service industries "do not join professional firms for *jobs*, but for *careers*" (Maister 1993, p.6).

Entry level positions in service firms give workers valuable experience and training that allow them to move on fairly quickly to prime positions in a wide variety of industries. As long as the number of entry level positions in service industries is smaller than the number of potential entrants, firms have all the bargaining power with junior workers. However, after junior workers serve their apprenticeships, their employers must compete with other firms in other industries to retain top talent.[15]

The fact that outsiders do not observe the outcome of a disclosure policy could arise from several circumstances. First, although many firms provide feedback to their employees, not all firms provide a hard copy of the outcome of performance reviews to their employees. Workers thus may not always be able to prove they did or did not get positive feedback to potential employers. Second, outside firms may not know whether a worker was or was not given feedback in a particular job or company, and so may not even request information on feedback during interviews. Finally, even if outsiders are informed of the outcome of a particular employee's performance review, they cannot always clearly interpret the information. Using performance appraisal information to infer a worker's talent requires information on what job he did, who his supervisor was, and what targets were set for him, and these institutional details are often not known by those outside an organization.

2.3 Definition of equilibrium

Let e_t^F be F's belief on period t worker effort, and e_t^W be the worker's actual period t effort. Let $\widehat{\theta}_t^F = E\left[\theta | \left\{y_\tau, e_\tau^F\right\}_{\tau=1}^t\right]$ be F's belief on worker talent after observing a given history of output realizations, and $\widehat{\theta}_t^W = E\left[\theta | \left\{\psi\left(y_\tau\right), e_\tau^W\right\}_{\tau=1}^t\right]$ be the worker's belief on his talent after observing a history of disclosures. Define $\widehat{\theta}_0^i \equiv \overline{\theta}$ for $i \in \{F, W\}$. Using notation for precisions of the normally distributed variables rather than variances is more convenient in defining equilibrium effort levels, so let $h_\varepsilon \equiv \left(\sigma_\varepsilon^2\right)^{-1}$, $h_\theta \equiv \left(\sigma_\theta^2\right)^{-1}$, and $\lambda_t \equiv \frac{h_\varepsilon}{t h_\varepsilon + h_\theta}$.

[15]In a discussion of the scarcity of young, educated workers, Maister (1993) writes:

> Most professions provide attractive *initial* career opportunities relative to other industries. Law school graduates will probably continue to join law firms, accounting graduates to join accounting firms, and business school graduates will continue to fined consulting and investment banking attractive first jobs.
>
> The real impact of the people crisis will be felt in absorbing the high costs that will result from competition for educated young workers, and continuing to make the professional-firm career path attractive in an environment when mid-level employees will receive numerous "head-hunting" calls (Maister 1993, p.192-3).

Definition 4 *Equilibrium efforts e_1^* and e_2^* are defined by the following conditions, which are defined for $t \in \{1,2\}$:*

$$
\begin{aligned}
e_2^W &= \arg\max_{e_2} -g(e_2) + \Pr\left[\widehat{\theta}_2^F > \theta^*\right] W & (1)\\
e_1^W &= \arg\max_{e_1} -g(e_1) - E\left[g\left(e_2^W\right)\right] + \Pr\left[\widehat{\theta}_2^F > \theta^*\right] W & (2)\\
\widehat{\theta}_t^F &= \lambda_t\left(y_t - e_t^F\right) + (1-\lambda_t)\,\widehat{\theta}_{t-1}^F & (3)\\
\widehat{\theta}_t^W &= E\left[\lambda_t\left(y_t - e_t^W\right) + (1-\lambda_t)\,\widehat{\theta}_{t-1}^W | \psi\left(\widehat{\theta}_1^F\right)\right] & (4)\\
e_t^* &= e_t^W = e_t^F. & (5)
\end{aligned}
$$

These conditions together constitute a Perfect Bayesian Equilibrium. The first two require the worker to maximize utility in all periods, the next two require Bayesian learning on the parts of the worker and the firm about the worker's talent, and the final requires the firm's beliefs about worker effort to be consistent with the worker's effort choices. (3) also shows why disclosing $\widehat{\theta}_1^F$ is equivalent to disclosing y_1: the former is simply a linear function of the latter. On the equilibrium path, F holds the belief about worker talent consistent with his actual effort choices, and in this case $\widehat{\theta}_t^F = \widehat{\theta}_t^W = \widehat{\theta}_t$. The distributions of $\widehat{\theta}_2^F|\widehat{\theta}_1$ and $\widehat{\theta}_1^F$ are important for interpreting equilibrium effort levels, so they are given below.

Lemma 1 *Let $\sigma_2^2 = \frac{\lambda_2}{h_\varepsilon + h_\theta}$ and $\sigma_1^2 = \frac{\lambda_1}{h_\theta}$.*

$$
\begin{aligned}
\widehat{\theta}_2^F|\widehat{\theta}_1 &\sim N\left(\widehat{\theta}_1 + \lambda_2\left(e_2 - e_1^F\right), \sigma_2^2\right)\\
\widehat{\theta}_1^F &\sim N\left(\overline{\theta} + \lambda_1\left(e_1 - e_1^F\right), \sigma_1^2\right).
\end{aligned}
$$

Global concavity of the problems in (1) and (2) is necessary for equilibrium existence. The parametric restriction guaranteeing global concavity for (1) is used in a later proof, so it is explicitly stated as an assumption.

Assumption 2 $\frac{W}{C}\left(\frac{\lambda_2}{\sigma_2}\right)^2 < 1.$

With the elements of the model completely described, the paper turns to solving it.

3 Effects of Feedback

This section solves for the worker's equilibrium effort levels according to Definition 4 via backward induction, and studies how they depend on feedback. The goal is not to study optimal disclosure, but to identify its primary effects. It first considers second period effort along with effort risk, then turns to first period effort along with coasting incentives.

3.1 Second period effort and effort risk

The second period is the last period of the game in which the worker exerts effort. If the game is on the equilibrium path, and the firm holds belief $\widehat{\theta}_1$ on worker talent (notation that was introduced in the previous section for equilibrium beliefs) then, by (3), the firm's second period belief as a function of y_2 is

$$\widehat{\theta}_2^F = \lambda_2 \left(y_2 - e_2^F \right) + (1 - \lambda_2) \widehat{\theta}_1 \tag{6}$$

From (6) one can see the source of effort incentives in the second period. By increasing e_2 the worker can increase $\widehat{\theta}_2^F$ and improve his chance of earning W in period 3. This *signal jamming incentive* appears in all career concerns models of effort supply. In equilibrium, the firm correctly infers that the worker exerts an effort level that equates the marginal benefit of increasing $\widehat{\theta}_2^F$ with the marginal cost of effort. The next result shows how the strength of signal jamming incentives depends on the information that the disclosure policy has revealed to him. Throughout this section and the rest of the paper, the paper will use the notation ϕ to denote the standard normal probability density function.

Proposition 1 *Suppose the game is on the equilibrium path in period 2. Given Assumption 2, e_2^* exists and is unique. If $\psi\left(\widehat{\theta}_1\right) = \widehat{\theta}_1$,*

$$e_2^* \left(\widehat{\theta}_1 \right) = \frac{W}{C} \frac{\lambda_2}{\sigma_2} \phi \left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2} \right) \tag{7}$$

and if $\psi\left(\widehat{\theta}_1\right) \neq \widehat{\theta}_1$,

$$e_2^* = E \left[\frac{W}{C} \frac{\lambda_2}{\sigma_2} \phi \left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2} \right) | \widehat{\theta}_1 \notin \Theta \right]. \tag{8}$$

The easiest way to understand the proposition is to consider figure 2, which plots out second period equilibrium effort as a function of $\widehat{\theta}_1$. The top graph gives equilibrium effort under the maximally informative disclosure policy $\Theta = (-\infty, \infty)$, and the bottom graph shows equilibrium effort under some disclosure policy $\Theta = (\underline{x}, \overline{x})$ where $\underline{x}$ and $\overline{x}$ lie equidistant from θ^*.

Under the full disclosure policy, the worker exerts the most effort when $\widehat{\theta}_1 = \theta^*$.[16] In this situation, whether or not the firm will retain the worker is still highly uncertain, so

[16]This is also the reason it is important for the firm to commit to Θ. Without this commitment, it would always have an incentive to lie to the worker in order to increase his second period effort.

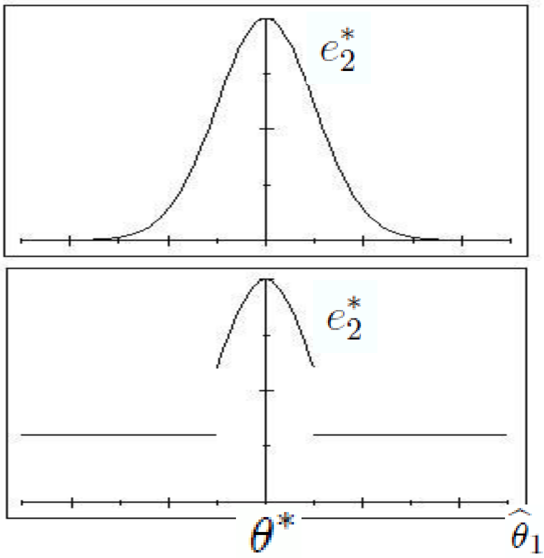

Figure 2: Equilibrium Effort with Full and Partial Disclosure

the worker's future payoff is sensitive to y_2. In contrast, if $\widehat{\theta}_1$ is either very high or very low the worker's period 3 payoff is practically a foregone conclusion: he is almost certain to earn $\overline{W}$ or $\underline{W}$ so his payoff is not sensitive to y_2. In this case, the worker exerts little effort. In terms of period 2 effort, positive and negative feedback exhibit symmetry. As one can see from 2, a worker who learns that his expected talent is $\underline{x}$ works just as hard as worker who learns his expected talent is $\overline{x}$. What matters for effort incentives is how far the feedback is from θ^*, not on which side of θ^* the feedback falls. In this sense, negative feedback is no more demotivating than positive feedback.

The bottom graph in figure 2 shows another important property of a disclosure policy. When the worker does not receive feedback, second period equilibrium effort does not depend directly on first period performance. Instead, he takes the expectation over all possible undisclosed talent types in selecting effort. Although the firm's belief about the worker's talent changes when y_1 changes within Θ^C, the worker's belief about his own talent does not change, so his effort is independent of y_1. This property is important for analyzing first period effort.

Depending on the disclosure policy, the worker can exert more or less effort for any particular $\widehat{\theta}_1$. Under the disclosure policy $\Theta = (\underline{x}, \overline{x})$, a worker for whom $\widehat{\theta}_1$ is either very high or very low works harder than he would under $\Theta = (-\infty, \infty)$ since he does not learn that his talent type is far from θ^*. Conversely, a worker for whom $\widehat{\theta}_1 = \underline{x}$ or $\widehat{\theta}_1 = \overline{x}$ works less harder under $\Theta = (\underline{x}, \overline{x})$ than under $\Theta = (-\infty, \infty)$ since he infers his talent to be further away from θ^* than it actually is. An important question is therefore how a disclosure policy affects expected second period effort over all possible realizations of $\widehat{\theta}_1$.

The next result shows it has no effect at all.

Lemma 2 $E[e_2^*]$ *is independent of* Θ.

The result is a straightforward application of the law of total probability. It arises because with quadratic effort costs the marginal cost of effort is linear, as is the expectation operator. The assumption of quadratic costs is crucial here. Using Jensen's inequality, one can show that if the marginal cost function were convex, $\Theta = \emptyset$ would maximize $E[e_2^*]$, and if it were concave, $\Theta = (-\infty, \infty)$ would. The quadratic cost function is thus the only reasonable cost specification for which there is no effect of information disclosure on $E[e_2^*]$.[17] The reason for making this assumption is to limit the effects of information disclosure to two. With non-quadratic costs there would be a third effect of feedback as well, but effort risk and coasting incentives would remain.

While information disclosure does not impact the expected value of second period effort, it does impact the worker's second period disutility of effort. This is the first major effect of information disclosure.

Lemma 3 *(Effort Risk). Suppose there exist two disclosure policies* Θ *and* Θ' *such that* $\Theta \subset \Theta'$. *Then* $E\left[(e_2^*)^2 | \Theta'\right] > E\left[(e_2^*)^2 | \Theta\right]$.

Mathematically speaking, this result is an application of Jensen's inequality, which in standard form states that $E\left[(e_2^*)^2\right] > (E[e_2^*])^2$. The proof applies Jensen's inequality to $\Theta' \backslash \Theta$, the set of interim beliefs that Θ' discloses but Θ does not. Economically speaking, the result is simply risk aversion, although over effort rather than wealth levels. The worker's preferences over second period effort are given by $-\frac{C}{2}(e_2^*)^2$, a concave function. Since $E[e_2^*]$ is the same on any interval regardless of the disclosure policy, the worker prefers a fixed effort level rather than the same effort level in expectation. Information disclosure thus subjects the worker to effort risk through exposing him to uncertainty about how hard he will have to work in the future. Ceteris paribus, the worker strictly prefers the uninformative disclosure policy $\Theta = \emptyset$. Moreover, since information disclosure increases the worker's expected disutility of second period effort without raising his expected output, $\Theta = \emptyset$ maximizes second period social surplus.

[17] This result echoes those in Aoyagi (2007) and Ederer (2008), who show that with quadratic costs and separability of talent and effort in production, all disclosure policies yield equivalent expected effort in dynamic tournaments.

3.2 First period effort and coasting incentives

To explore the incentives the worker has to provide first period effort, it is again useful to examine (3), which one can alternatively express as

$$\widehat{\theta}_2^F = \lambda_2 \left(y_1 - e_1^F \right) + \lambda_2 \left(y_2 - e_2^F \right) + \frac{h_\theta}{2h_\varepsilon + h_\theta} \overline{\theta}. \tag{9}$$

Just as with second period effort, there are signal jamming incentives in the first period since the worker can increase y_1 through exerting effort. While there is nothing that the worker can do to affect y_2, this section will show that he can affect e_2^F conditional on receiving feedback. e_2^F is relevant for the worker because what matters for his period 3 payoff is not his second period output as such, but the portion of second period output that the firm attributes to his talent. The next result makes more precise the relationship between the disclosure policy and first period effort.

Proposition 2 *For high enough C, first period equilibrium exists and is unique, and is given by*

$$e_1^* (\Theta) = E \left[\frac{W}{C} \frac{\lambda_2}{\sigma_2} \phi (\xi) \right] + E \left[\left(\frac{W^2}{C^2} \frac{\lambda_1 \lambda_2^2}{\sigma_2^3} \right) \xi \phi^2 (\xi) \left| \widehat{\theta}_1 \in \Theta \right. \right] \Pr \left[\widehat{\theta}_1 \in \Theta \right], \tag{10}$$

where

$$\xi = \left(\frac{\widehat{\theta}_1 - \theta^*}{\sigma_2} \right).$$

The two terms in (10) reflect the two sources of first period effort incentives. The first arises from signal jamming. In fact, this expression is equal to second period equilibrium effort with the dislcosure policy $\Theta = \emptyset$. The reason is that in the first period, regardless of the disclosure policy, the worker does not have any information about his talent, so is in exactly the same informational environment as if he were choosing second period effort under $\Theta = \emptyset$.

The second term reflects coasting incentives, and does depend on the disclosure policy. Regardless of the worker's actual first period effort level, the firm will expect the game to be on the equilibrium path in the second. One can therefore use the results of Proposition 1 to derive the expression for e_2^F. As discussed in the previous section, when the worker does not receive feedback, e_2^F is independent of first period performance. So, coasting incentives arise only when the worker receives feedback from the firm. Conditional on receiving feedback, the firm expects the worker to exert efort

$$e_2^F = \frac{W}{C} \frac{\lambda_2}{\sigma_2} \phi \left(\frac{\theta^* - \widehat{\theta}_1^F}{\sigma_2} \right). \tag{11}$$

The worker therefore has an incentive to increase $\widehat{\theta}_1^F$ when $\widehat{\theta}_1^F > \theta^*$ and to decrease $\widehat{\theta}_1^F$ when $\widehat{\theta}_1^F < \theta^*$. In other words, the worker always wants to push $\widehat{\theta}_1^F$ away from θ^*, where e_2^F is highest, and into the tails of the distribution.

Corollary 1 *(Coasting incentives). Removing negative feedback and adding positive feedback to Θ increases e_1^*.*

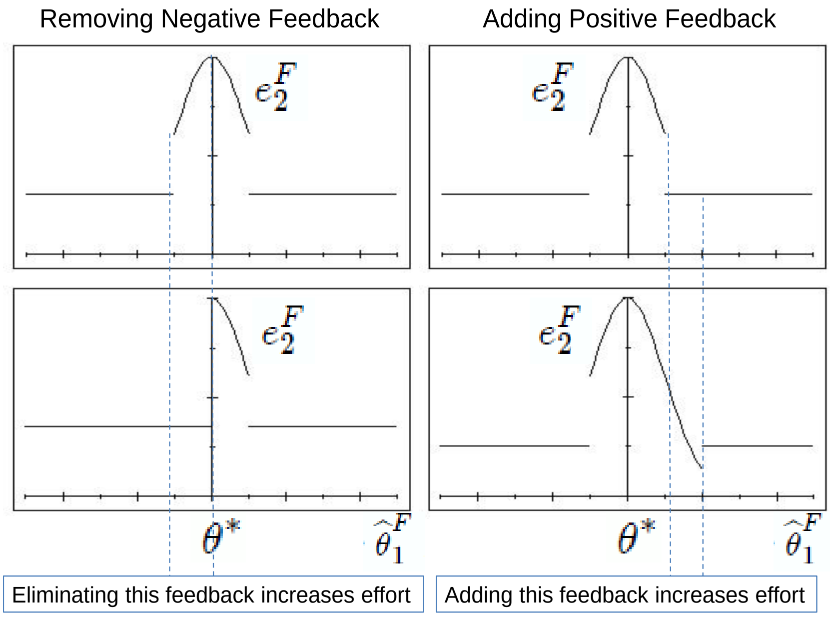

Figure 3: The Effects of Feedback on First Period Effort

Figure 3 plots out the relationship between $\widehat{\theta}_1^F$ and e_2^F under various disclosure policies. The top two graphs are both for the disclosure policy $\Theta = (\underline{x}, \overline{x})$, where $\underline{x}$ and $\overline{x}$ are equidistant from θ^*. For the feedback $(\underline{x}, 0)$ there is a positive relationship between $\widehat{\theta}_1^F$ and e_2^F. So, the worker has an incentive to work less hard in the first period in order to reduce e_2^F and make it easier to signal a hight talent in the second. In this case, coasting incentives counteract signal jamming incentives and discourage first period effort. Eliminating this negative feedback thus increases e_1^*.

Under the disclosure policy $\Theta = (\underline{x}, \overline{x})$, e_2^F is flat for $\widehat{\theta}_1^F \in (\overline{x}, \overline{x}+1)$. If instead the firm offers the disclosure policy $\Theta = (\underline{x}, \overline{x}+1)$, there is a negative relationship between two in this region. Under this disclosure policy, there are extra work incentives because when the worker receives feedback in $(\overline{x}, \overline{x}+1)$, increasing e_1 decreases e_2^F since it pushes $\widehat{\theta}_1^F$ into the upper tail of the distribution where the firm believes the worker exerts less effort. In this case, coasting incentives complement signal jamming incentives and provide an additional motive for exerting effort.

This section has now identified two primary effects of information disclosure. Feedback exposes the worker to effort risk and creates coasting incentives. The next section shows how these combine in an optimal contract.

4 Optimal Disclosure

The trade-off between effort incentives and risk in contract theory dates back at least to Holmstrom (1979). So, in some ways, the present model has analogies with previous work in principal-agent theory, although effort and risk obviously arise for different reasons in this environment. As this section shows, the way the firm optimally resolves them is also quite different. Its problem is to select a contract (w, Θ) to maximize

$$e_1^* - w \tag{12}$$

such that

$$w \geq 0 \tag{13}$$

and

$$w + \Pr\left[\widehat{\theta}_2 \geq \theta^*\right] \overline{W} + \Pr\left[\widehat{\theta}_2 < \theta^*\right] \underline{W} - g\left(e_1^*\right) - E\left[g\left(e_2^*\right)\right] \geq u. \tag{14}$$

(12) is the part of firm profit that the disclosure policy affects. It has a relatively simple form because the retention rule the firm uses is exogenous, and by Lemma 2, Θ does not affect $E\left[e_2^*\right]$. The firm faces two constraints in choosing a contract: (13) is the worker's liquidity constraint and (14) is his participation constraint. The left hand side of (14) reflects the two sources of utility that the worker derives from joining the firm. The first is the initial compensation w and the second is the expected third period payoff net of the first and expected second period effort costs. A useful result is the following.

Lemma 4 $\Pr\left[\widehat{\theta}_2 \geq \theta^*\right] \overline{W} + \Pr\left[\widehat{\theta}_2 < \theta^*\right] \underline{W} - g\left(e_1^*\right) - E\left[g\left(e_2^*\right)\right]$ *is unbounded above.*

One way to demonstrate this is to note that as $\overline{W}$ becomes large, the expected third period wage is unbounded. At the same time, Assumption 2 places a bound on effort costs.

When the expected payoff from joining the firm is large, the worker will accept any contract terms offered him in period 0. When it is smaller, the firm will have to be more disciplined in its choice of contract since the worker will take his outside option if he is not offered enough up-front compensation. The paper thus classifies employment relationships into two types: when the participation constraint does not bind for any contract, there are *high rewards*; when the participation constraint binds when $w = 0$ for any disclosure policy there are *low rewards*.[18] The rest of this section derives the optimal contract in both cases.

4.1 Optimal contract with high rewards

Let (w_H^*, Θ_H^*) denote the optimal contract with high rewards. Its structure is straightforward to derive.

Proposition 3 $w_H^* = 0$ *and* $\Theta_H^* = (\theta^*, \infty)$.

Proof. With high rewards, the worker will accept employment regardless of the up-front wage, so F optimally chooses the lowest possible compensation and sets $w_H^* = 0$. Since the wage is fixed, the firm chooses the disclosure policy to maximize coasting incentives, so the firm chooses $\Theta_H^* = (\theta^*, \infty)$. ∎

With high rewards, the goal of the firm is simply to extract as much effort as possible from the worker, and the instrument available to it for doing so is positive feedback. The fact that positive feedback exposes the worker to effort risk is not relevant for the firm because it does not have to compensate the worker for it since the participation constraint does not bind.

One interesting feature of the result is that it gives a potential explanation of why an organization would want to create a culture of giving only positive feedback. The human resource literature has found that criticism can harm productivity and praise can boost it (Meyer, Kay, and French 1965), but has not pinned down any sharp intuitions explaining why. In this model, coasting incentives explain why a firm that only cares about maximizing worker motivation should always identify good performance and never identify poor performance.

4.2 Optimal contract with low rewards

With low rewards (14) will bind in the optimal contract since the worker must receive some compensation for working for F. Denote by (w_L^*, Θ_L^*) the optimal contract in this situation. F chooses w_L^* to make the participation constraint bind, and chooses Θ_L^* to maximize [19]

$$e_1^* - g\left(e_1^*\right) + E\left[e_2^* - g\left(e_2^*\right)\right]. \tag{15}$$

[18] These two cases are not a complete classification of the possible parameter values of the model. There are cases in which the participation constraint binds for some but not all disclosure policies when $w = 0$. However, the paper does not study them since they add little additional insight.

[19] One obtains this expression through taking (14) as an equality, plugging into (12), and dropping the terms that the disclosure policy does not influence.

Expression (15) is simply social surplus. Θ_L^* maximizes social surplus rather than effort because now F internalizes the effort risk to which the worker becomes exposed when he receives feedback. It does so because it must increase the up-front compensation he must provide to attract the worker into employment. Unlike the case of high rewards, Θ_L^* must now trade-off effort incentives and risk. The efficient first period effort level is given by $Ce_1^{FB} = 1$. The incentive of F to provide feedback depends on $e_1^*(\emptyset)$, the level of first period effort with no feedback and just signal jamming incentives. If $Ce_1^*(\emptyset) < 1$, then giving positive feedback raises first period effort closer to its first best value, raising first period social surplus. However, it also decreases second period social surplus because of effort risk. Thus, the trade-off is between increasing first period effort closer to its first best level and exposing the worker to excess risk. The same trade-off arises if $Ce_1^*(\emptyset) > 1$, but for giving negative feedback. The next result shows how Θ_L^* resolves it.

Proposition 4 (w_L^*, Θ_L^*) *satisfies*

$$\begin{aligned} \text{If } Ce_1^*(\emptyset) &< 1,\ \Theta_L^* = (\theta', \theta''),\ \text{where } \theta^* < \theta' < \theta'' < \infty; \\ \text{If } Ce_1^*(\emptyset) &> 1,\ \Theta_L^* = (\theta', \theta''),\ \text{where } -\infty < \theta' < \theta'' < \theta^*; \\ \text{and if } Ce_1^*(\emptyset) &= 1,\ \Theta_L^* = \emptyset. \end{aligned}$$

and

$$w_L^* = u - \Pr\left[\widehat{\theta}_2 \geq \theta^*\right] \overline{W} - \Pr\left[\widehat{\theta}_2 < \theta^*\right] \underline{W} + g\left(e_1^*\left(\Theta_L^*\right)\right) + E\left[g\left(e_2^*\right) | \Theta = \Theta_L^*\right].$$

One only needs to consider the case $Ce_1^*(\emptyset) < 1$. The argument for $Ce_1^*(\emptyset) > 1$ works in exactly the same fashion, and when $Ce_1^*(\emptyset) = 1$ feedback can only harm social welfare so $\Theta_L^* = \emptyset$ is optimal. Clearly, when $Ce_1^*(\emptyset) < 1$, including feedback anywhere in $(-\infty, \theta^*)$ is not optimal. In order to understand the particular form that Θ_L^* takes, it is useful to consider figure 4, whose origin is set at $(\theta^*, 0)$. The dark line traces out the marginal cost of disclosing talent beliefs $\left(\widehat{\theta}_1, \widehat{\theta}_1 + \Delta\right)$ for an arbitrarily small Δ given a fixed Θ that does not include $(-\infty, \theta^*)$. The lighter line traces out the marginal benefit of disclosing $\left(\widehat{\theta}_1, \widehat{\theta}_1 + \Delta\right)$ given the same Θ.

The most notable feature of the marginal cost curve is that it is single-troughed. To understand why, let $\widetilde{e}_2$ be the effort the worker exerts when he receives no feedback. Because the disclosure policy does not include $(-\infty, \theta^*)$, $\widetilde{e}_2$ lies strictly between 0 and θ^*. The magnitude of the effort risk from disclosing $\left(\widehat{\theta}_1, \widehat{\theta}_1 + \Delta\right)$ is propotional to how far away is $e_2^*\left(\widehat{\theta}_1\right)$ from $\widetilde{e}_2$. For $\widehat{\theta}_1$ large and for $\widehat{\theta}_1$ close to θ^*, effort risk is high since these are circumstances in which $e_2^*\left(\widehat{\theta}_1\right)$ reaches extreme values. On the other hand, there

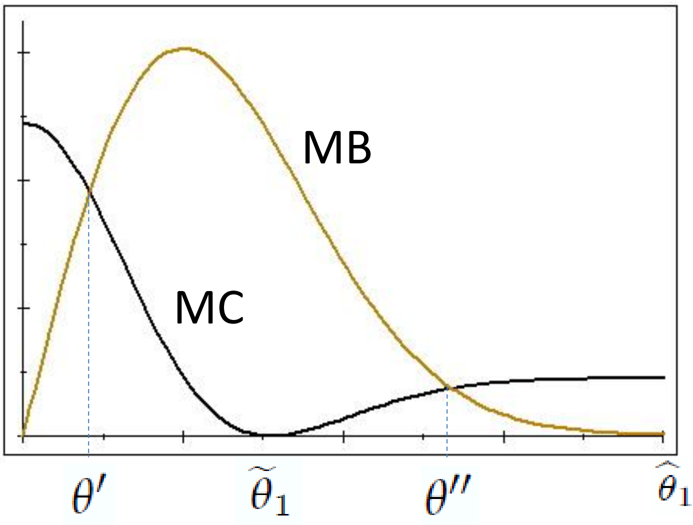

Figure 4: The Marginal Benefit and Cost of Information Disclosure

exists a unique $\widetilde{\theta}_1$ at which $e_2^*\left(\widetilde{\theta}_1\right) = \widetilde{e}_2$. Disclosing beliefs $\left(\widetilde{\theta}_1, \widetilde{\theta}_1 + \Delta\right)$ is therefore riskless, and does not harm social welfare.

The marginal benefit curve is single-peaked. The marginal benefit of including beliefs $\left(\widehat{\theta}_1, \widehat{\theta}_1 + \Delta\right)$ in Θ is proportional to the strength of the coasting incentives that this feedback creates. There are two clear regions where e_2^F does not vary much with e_1. For $\widehat{\theta}_1^F$ close to θ^*, the firm is convinced that the worker will exert high effort to earn the reputational reward W: e_2^F is flat in $\widehat{\theta}_1^F$ in a neighborhood around θ^*. Also, when $\widehat{\theta}_1^F$ is very high, the firm believes the worker will hardly exert effort at all since earning W is nearly guaranteed, and so an increase in e_1 has little effect on e_2^F when $\widehat{\theta}_1^F$ grows large. For these reasons, the marginal benefit curve approaches 0 as $x \to \theta^*$ and $x \to \infty$. It is instead for intermediate talent beliefs that coasting incentives are strongest.

One can see from figure 4 that for any disclosure policy Θ not including $(-\infty, \theta^*)$, there exists an interval (θ', θ'') in which the marginal benefit of disclosing beliefs around $\widehat{\theta}_1 \in (\theta', \theta'')$ exceeds the marginal cost. The marginal cost of disclosing beliefs around any $\widehat{\theta}_1$ that lies outside this interval exceeds the marginal benefit. Therefore it must be the case that the optimal disclosure policy reveals a convex set of beliefs that lie strictly within (θ^*, ∞).

4.3 Features of Optimal Disclosure Policies

This section discusses the properties of Θ_L^* and Θ_H^* in light of the features of real-world rating distributions featured in figure 1. Importantly, both feature *some* information disclosure (except when $Ce_1^*(\emptyset) = 1$ with low rewards). Thus the model provides a rationale for why firms invest in performance appraisal systems at all.

Another important feature of both disclosure policies is that neither provides full information to workers. Thus, observing that a firm does not have very informative rating distributions should in no way lead to the conclusion that its performance appraisal system is somehow dysfunctional.

The disclosure policies are uninformative in a particular way. Neither one explicitly identifies the worst performers; instead, they are always pooled with other types.[20] Moreover, Θ_L^* does not identify top performance. With low rewards, all workers infer their talent to be in the middle of performance distribution. In short, the model shows that leniency and concentration are compatible with optimal feedback.

Of course, these features arise within a specialized model with numerous simplifying assumptions. This section concludes by examining the robustness of the findings to these.

4.4 Robustness

One objection one might raise about Θ_L^* is that workers at the top and bottom of the performance distribution are pooled together, whereas it is perhaps more natural to think that workers know in which end of the distribution their performance lies. However, unless the firm would like to commit to sharing this information with the worker, there is no way for him to know. The next result shows that the firm cannot benefit from disclosing this additional information.

Proposition 5 *Let Θ_1 be the set of all beliefs $\widehat{\theta}_1^F$ and let P be a partition of Θ_1. The profit maximizing disclosure policy satisfying Definition 1 is also profit maximizing within the class of disclosure policies with the form $\psi\left(\widehat{\theta}_1^F\right) \to P_i$ where $\widehat{\theta}_1^F \in P_i$.*

Proof. Suppose ψ is a disclosure policy with the form $\psi\left(\widehat{\theta}_1^F\right) \to P_i$ where $\widehat{\theta}_1^F \in P_i$. From the arguments used in Proposition 2, coasting incentives arise over the set of exactly revealed beliefs $\widetilde{\Theta}_1 = \left\{\widehat{\theta}_1^F | \psi\left(\widehat{\theta}_1^F\right) = \widehat{\theta}_1^F\right\}$. Since e_2^F is flat in $\widehat{\theta}_1^F$ conditional on $\widehat{\theta}_1^F \in P_i \subset \Theta_1 \backslash \widetilde{\Theta}_1$, coasting incentives do not arise over $\Theta_1 \backslash \widetilde{\Theta}_1$.

Now, suppose $\exists \widehat{\theta}_{1i}, \widehat{\theta}_{1j} \in \Theta_1 \backslash \widetilde{\Theta}_1$ such that $\psi\left(\widehat{\theta}_{1i}\right) = P_i$ and $\psi\left(\widehat{\theta}_{1j}\right) = P_j$ where $P_i \neq P_j$. Modifying ψ so that $\psi\left(\widehat{\theta}_1\right) = P_i \cup P_j \; \forall \widehat{\theta}_1 \in P_i \cup P_j$ does not alter coasting incentives, but (weakly) reduces the effort risk to the worker, so cannot reduce profit. ■

Unless information contributes to improving effort incentives, the firm should suppress it because of effort risk. Coasting incentives arise whenever first period effort affects the amount of effort the firm expects in the second period. Unless the disclosure policy reveals

[20] Even when Θ_L^* gives negative feedback, it does not inform the worst performers of their output.

$\widehat{\theta}_1^F$ directly to the worker, e_2^F does not depend on e_1. In order to see this more clearly, one can easily adapt Proposition 1 to show that e_2^F satisfies

$$e_2^F = E\left[\frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^* - \widehat{\theta}_1^F}{\sigma_2}\right)|\widehat{\theta}_1^F \in P_i\right],$$

which does not depend on y_1 when P_i contains a positive measure of beliefs. So, coasting incentives only arise over the range of exactly revealed beliefs. While breaking the range of beliefs that are not exactly revealed to the worker into more than set does not improve first period effort, it does expose the worker to more effort risk since it increases the variance of second period effort. Hence, pooling all the beliefs that are not exactly revealed to the worker into one message cannot make the firm worse off, and the optimal disclosure policies derived in this section are optimal within a general class.

While the assumption of quadratic effort costs makes the exposition of the results clear, examining how the results depend on it is obviously important. Suppose the worker has effort costs given by $g(e_t) = \frac{C}{\beta+1}e_t^{\beta+1}$. As discussed in section 3.1, whenever $\beta \neq 1$ there is a third effect of feedback as well since the disclosure policy affects expected second period effort as well as effort risk and coasting incentives.

Proposition 6 *Let $g(e_t) = \frac{C}{\beta+1}e_t^{\beta+1}$ and suppose that $\beta > 1$. Then the effort maximizing disclosure policy takes the form $\Theta = (\theta', \theta'')$ where $\theta^* < \theta' < \theta'' < \infty$ and the social surplus maximizing disclosure policy takes the form given in Proposition 4.*

When $\beta > 1$ the disclosure policies with high and low rewards both feature avoidance of negative feedback and concentration. Qualitatively, then, the results do not depend on the assumption of quadratic costs. However, the results do depend on convex marginal costs, so are not robust to general cost functions.[21] Nevertheless, since the assumption of non-negative third derivatives on the effort cost function is common in the mechanism design and contract theory literatures, the conditions under which the results arise are not unduly restrictive.

The form that Θ_H^* and Θ_L^* take is also dependent on the shape of the period 3 wage profile. While effort risk and coasting incentives are presumably general effects in career concerns models, how they combine in an optimal contract is specific to their relative magnitudes, which derive from the step function rewards to talent schedule. The next section shows how adverse selection labor market competition generates a wage schedule that pays a fixed prize to a worker whose expected talent crosses a threshold. It also provides a microfoundation for high and low rewards.

[21]For example, when $\beta < 1$ the effort maximizing disclosure policies does identify the worst performers.

5 Feedback, Competition, and Technology

The goal of this section is twofold. The first is to endogenize the period 3 wage function assumed in the basic model. This is achieved through adapting the adverse selection labor market competition model of Waldman (1984). The resulting wage schedule pays a fixed prize to workers of high enough talent. The second is to provide a microfoundation for high rewards, which is achieved through specifying a third period production function.

5.1 Rewards to talent

The critical feature for obtaining a micro-founded wage schedule appeared in the basic setup of the model: outsiders cannot observe the outcome of a disclosure policy. This means that after period 2, F has private information (in the form of its equilibrium belief on worker talent $\widehat{\theta}_2$) that is not available to the labor market, which the paper models as a large number of identical firms indexed by $j \in J$. Labor market competition therefore takes place under adverse selection.

Suppose the worker's output in period 3 with F is $y_3^F = \kappa + k\theta + \varepsilon_3$ where $k \geq 1$ and $\kappa > 0$, and his output with a market firm is $y_3^j = k\theta + \varepsilon_3$. The worker's output is therefore potentially more sensitive to his talent in period 3 than in periods 1 and 2, and F earns a rent on every worker type. The first feature captures the fact that more senior workers usually occupy positions of higher responsibility in firms, so that their skill is more important in determining the value they create. The second captures firm specific human capital accumulation. The fact that y_3 does not depend on effort is without loss of generality since final period effort is zero in career concerns models.

After observing y_1 and y_2, F chooses some $w_3^F \in W_3^F = \mathbb{R}^+ \cup \{\emptyset\}$ where $w_3^F = \emptyset$ corresponds to not making the worker a wage offer for period 3 and $w_3^F \neq \emptyset$ to offering him a wage w_3^F for period 3. Outside firms observe w_3^F and each then simultaneously chooses $w_3^j \in \mathbb{R}^+ \cup \{\emptyset\}$. The worker joins the firm that offers him the highest wage, but takes his outside option if no firm makes him a wage offer that exceeds it. All firms incur an arbitrarily small cost δ from making wage offers, which could for example reflect the legal costs of drafting a wage contract.

Denote by Θ_2 the set of all possible realizations of $\widehat{\theta}_2$, and let $\widetilde{w}_3^F : \Theta_2 \rightarrow W_3^F$ be F's strategy in the bidding game. Let $\widehat{w}_3^j$ be market firm j's belief about the strategy that F employs, and let $\widetilde{w}_3^j : W_3^F \rightarrow W_3^j$ be the strategy that each market firm j employs.

Definition 5 *An equilibrium of the labor market competition game is a set of strategies*

$\widetilde{w}_3^{F*} \cup \left\{\widetilde{w}_3^{j*}\right\}_{j=1}^{J}$ *that satisfy*

$$\widetilde{w}_3^{j*} \text{ maximizes } E\left[y_3^j\right] - w_3^j - 1\left(w_3^j \neq \emptyset\right)\delta \;\forall w_3^F \text{ given } \widetilde{w}_3^{F*} \text{ and } \left\{\widetilde{w}_3^{j*}\right\}_{j=1}^{J} \backslash \widetilde{w}_3^{j*}; \tag{16}$$

$$\widetilde{w}_3^{F*} \text{ maximizes } E\left[y_3^F\right] - w_3^F - 1\left(w_3^F \neq \emptyset\right)\delta \;\forall \widehat{\theta}_2 \text{ given } \left\{\widetilde{w}_3^{j*}\right\}_{j=1}^{J}; \tag{17}$$

$$\widehat{w}_3^j = \widetilde{w}_3^{F*}. \tag{18}$$

Condition (16) says that the strategies employed by market firms constitute a Nash Equilibrium; (17) that F maximizes profit given the behavior of the labor market; and (18) that the belief of each market firm about the strategy employed by F is consistent with its actual strategy. These conditions together define a PBE.

The game that market firms play after observing w_3^F is a Bertrand bidding game whose solution is standard and left unproved. While it does not have a unique equilibrium, it does have a unique equilibrium outcome. In equilibrium all market firms must have the same belief on the strategy employed by F, so $\widehat{w}_3^j = \widehat{w}_3$.

Claim 1 *If*

$$w_3^F \neq \emptyset \text{ and } E\left[k\theta | w_3^F, \widehat{w}_3\right] - w_3^F - \delta > u$$

or

$$w_3^F = \emptyset \text{ and } E\left[k\theta | w_3^F = \emptyset, \widehat{w}_3\right] - \delta \geq u$$

then $\max_j w_3^j = E\left[k\theta | w_3^F, \widehat{w}_3\right] - \delta$.

Otherwise, $w_3^j = \emptyset \;\forall j$.

If F makes a positive wage offer to retain the worker but $E\left[k\theta | w_3^F, \widehat{w}_3\right] - w_3^F - \delta > u$, then an outside firm can improve on w_3^F, attract the worker, and earn positive profit. If F does not make the worker a wage offer and $E\left[k\theta | w_3^F = \emptyset, \widehat{w}_3\right] - \delta \geq u$, an outside firm can pay the worker a wage higher than his outside option and also make positive profit. However, Bertrand competition ensures that in either case the worker earns the entire surplus that his expected talent generates in the outside market. The following result describes the equilibrium behavior of F in response to this behavior by the labor market.

Proposition 7 *A pure strategy $\widetilde{w}_3^{F*}$ exists and is unique. It takes the form*

$$\widetilde{w}_3^{F*} = \begin{cases} E\left[k\theta | \widehat{\theta}_2 \geq \theta^*\right] \text{ if } \widehat{\theta}_2 \geq \theta^* \\ \emptyset \text{ if } \widehat{\theta}_2 < \theta^* \end{cases}$$

where θ^* uniquely satisfies

$$\kappa + k\theta^* = E\left[k\theta|\widehat{\theta}_2 \geq \theta^*\right]. \tag{19}$$

In words, the equilibrium strategy $\widetilde{w}_3^{F*}$ calls on F to retain the worker if and only if its belief on his talent exceeds θ^* and to pay all retained worker types a wage independent of $\widehat{\theta}_2$. This strategy combined with the labor market behavior identified in Claim 1 leads to two wages in equilibrium: one to worker-types that F retains and one to worker types that F releases. The resulting rewards to talent schedule thus has the same structure as assumed in the basic model. The reason that F cannot pay any two retained worker-types two different wages is that wages cannot credibly signal the private information. If F paid two different retained types two different wages in equilibrium, and the outside market believed these signals credibly communicated private information, F would have an immediate incentive to "lie" to the market and tell it the worker of higher talent was the one of lower talent through offering it a lower wage. Within the set of retained workers, communication between F and the outside market is impossible.[22]

The reason that only one wage is paid to released workers is because of costly bidding. Without costly bidding, F could still only retain workers above θ^* at wage $E\left[k\theta|\widehat{\theta}_2 \geq \theta^*\right]$; however, it could credibly communicate private information to the market for worker types below θ^*. Providing information for these worker types would be costless and would not affect third period profits. For example, one possible equilibrium strategy for F could specify $\widetilde{w}^{F*} = \widehat{\theta}_2 - \varepsilon$ if $\widehat{\theta}_2 < \theta^*$. The outside market would then hire the worker whenever $k\widehat{\theta}_2 - \delta - u \geq 0$ at wage $k\widehat{\theta}_2 - \delta$. This strategy would lead to a different rewards to talent schedule and change effort incentives in periods 1 and 2. Among the different equilibria of the costless bidding game, the most plausible one would maximize profit. However, solving for the optimal amount of information provision to the outside market is outside the remit of this paper. Assuming costly wage offers not only avoids this problem, but has a realistic interpretation.

5.2 High rewards and technology

The model of labor market competition introduced in this section can also provide a microfoundation for high rewards. The following (unessential) parametric restriction is particularly useful for providing a clean characterization of the expected future rewards the worker enjoys from joining the firm in period 0.

[22]The intuition is similar to that in the cheap talk literature (Crawford and Sobel 1982). Within the set of retained types, the sender of private information (the firm) has no preference concordance at all with the receiver (the outside labor market), so communication is impossible.

Assumption 3 *The parameters of the model are such that* $E\left[k\theta|\widehat{\theta}_2 < \theta^*\right] > u$

The main result is the following.

Lemma 5 *High rewards arise if* k *is high enough.*

Proof. Given Assumption 3, the expected period 3 payoff for the worker is

$$\Pr\left[\widehat{\theta}_2 \geq \theta^*\right] E\left[k\theta|\widehat{\theta}_2 \geq \theta^*\right] + \Pr\left[\widehat{\theta}_2 < \theta^*\right] E\left[k\theta|\widehat{\theta}_2 < \theta^*\right] = kE\left[\theta\right] = k\overline{\theta}.$$

Therefore, as $k \to \infty$, the expected return to employment grows without bound. ■

In period 3, the worker earns his expected output in the market conditional on being retained or released by the firm. The worker's expected period 3 wage is therefore his unconditional expected output in a market firm, or $k\overline{\theta}$. The worker's expected earnings are high whenever the productivity of senior workers in market firms is sensitive to talent. For example, one might imagine a group of consulting firms that works on highly complex and novel problems, and another one that works on more routine ones. If these groups compete within themselves for workers of high ability, beginning employment in the former group would yield higher future rewards since the output of its senior workers is presumably more sensitive to talent.

The previous section showed that in the high reward case, a firm maximizes effort, while in the low reward case it maximizes social surplus. Lemma 5 leads to the conclusion that firms in industries in which future output is sensitive to talent either provide too much positive feedback (if signal jamming incentives alone provide less than first best effort incentives) or not enough negative feedback (if they provide greater than first best effort incentives).

Since the typical concern with signal jamming is too little rather than too much effort provision, it is worth considering this case in more detail. In the model, information has a limited social value. When the firm chooses a disclosure policy to maximize effort, it creates effort incentives that go beyond what a welfare maximizing contract would provide. This is not to say that $e_1^*\left(\Theta_H^*\right) > e_1^{FB} > e_1^*\left(\emptyset\right)$, although in special situations it may be the case that maximizing coasting incentives drives effort above its first best level. Even if $e_1^*\left(\Theta_H^*\right) < e_1^{FB}$, Θ_H^* still motivates the worker too highly. From a social standpoint, there always exist intervals of positive feedback for which the social cost of effort risk outweighs the social gain from effort incentives. Moreover, when $e_1^*\left(\emptyset\right) < e_1^{FB}$, the effort maximizing disclosure policy is too informative since it provides feedback in situations where the social surplus maximizing policy would not.

Lemma 6 *As* $C \to \infty$ *and* $\theta^* \to -\infty$*, the informativeness of* Θ_H^* *approaches* 1*, and the informativeness of the social welfare maximizing disclosure policy approaches* 0*.*

This result shows that the difference between the informativeness of Θ_H^* and the social welfare maximizing disclosure policy can be maximal. When $C \to \infty$ and $\theta^* \to -\infty$, firm H always provides feedback, although a social planner never would. The reason is that when C is high, equilibrium effort in the second period is low. This in turn means that coasting incentives are weak, and from a social standpoint feedback is never justified. However, firm H always offers positive feedback in spite of the fact that it may only add a small amount of extra effort.

Finally, when $e_1^*(\emptyset)$ is greater than e_1^{FB}, an effort maximizing firm uses feedback to generate even higher levels of effort, whereas the social welfare maximizing policy would give negative feedback to dampen effort incentives. In this situation, feedback worsens a pre-existing rat race.[23]

6 Discussion

This paper has argued that professional service firms with up-or-out promotion contracts can benefit from giving feedback because workers respond with higher effort. If the model is correct, then firms in which career concerns are important have a motivation for investing in performance appraisal systems that other firms do not have. The goal of this section is to identify what industries use performance appraisal the most, and to assess how closely they correspond to the kind of industry modelled in the paper. The paper takes data from the Workplace Employment Relations Survey (WERS) from 2004, a cross-sectional survey of 2,295 workplaces in the United Kingdom that collects information on numerous establishment characteristics.[24] The paper considers the subset of private sector firms, since public organizations are likely to have bureaucratically controlled human resource practices. This leaves 1,557 workplaces in the dataset.

The paper divides firms into two groups: those that formally assess the largest non-managerial occupational group and those that do not. It focuses on appraisals for this set of workers since it is the group for which career concerns are probably most important. For each five digit industry, it computes the percentage of firms that formally assess the performance of their core non-managerial employees. In order to generate sufficient inter-industry variation, it drops those industries that have fewer than five firms in the dataset. This leaves ninety five industries, twelve of which have one hundred percent performance appraisal use. These are listed in the table 1, along with the twelve industries with the

[23]Landers, Rebitzer, and Taylor (1996) have found that junior workers in law firms exert inefficiently high levels of effort. This paper shows that human resource practices as well as reputational rewards can contribute to the overprovision of effort.

[24]Kersley, Alpin, Forth, Bryson, Bewley, Dix, and Oxenbridge (2006) gives a detailed description of the dataset.

HIGHEST PA USE	**LOWEST PA USE**
Publishing of journals and periodicals	Other meat and poultry processing
Collection, purification and distribution of water	Manufacture of bread
Retail sale via mail order house	Freight transport by road not elsewhere classified
Non-life insurance	Other transport via railways
Activities auxiliary to insurance and pension funding	Other construction work involving special trades
Other letting of own property	Dispensing chemists
Software consultancy and supply	Independent public houses and bars
Legal activities	Printing not elsewhere classified
Accounting and auditing activities	Forging, pressing, stamping and roll forming of metal
Business and management consultancy activities	Storage and warehousing
Private sector hospital activities	Maintenance and repair of motor vehicles
Non-charitable social work activities	Manufacture of other builders' ware of plastic

Table 1: Performance Appraisal Use by Industry

lowest use of performance appraisal.

The majority of industries with universal use of performance appraisal provide professional services, while at the same time, professional services do not appear at all among those industries in which which performance appraisal use is low. Of course, there are many reasons why professional service firms might provide feedback. Performance appraisal might allow a central HR office to more readily identify the most talented individuals and to appropriately reward them. It might also allow workers to identify training needs.[25] A more direct test of the model would be to examine data from a professional service firm that began to provide more feedback to workers, and to examine whether workers began exerting more effort at the same time. Whether such data are currently available is unclear, but the point is that the paper presents a perspective on feedback in organizations that is empirically distinguishable from other stories.

Concluding remarks. This paper began with the basic question: is more feedback always better? The answer is "no" for two different reasons. First, in many situations

[25]However, in the absence of legal restrictions, it is unclear why supervisors could not give feedback on worker performance to an HR office and not the worker. Moreover, firms can provide information to workers about training needs without giving them information about the probabiliy of future promotion (Beer 1987).

a firm would like to commit to limiting the amount of information that workers receive about their output. Full information disclosure would strictly reduce firm profits because negative feedback reduces effort and effort risk causes wages to rise. Second, firms that offer more feedback than others could very well be providing too much information from a social standpoint and lowering surplus. The fact that such firms might be more productive cannot be taken as evidence of their adopting better management practices.

While the paper has shown how one can capture characteristics of real world feedback systems in a model with standard preferences and technology, there is still much to understand. One area for future research would be how to implement the optimal disclosure policy in an organization that could not commit ex ante to disclosing certain beliefs. One avenue to explore would be having the firm share information with a continuum of workers, because through communicating with each other, the workers could detect deviations from the firm's announced disclosure policy. Even if no informative communication could ever arise in equilibrium, the results of the paper imply that this may actually be better for the firm than forcing managers to disclose all information to workers.

A second pertinent extension would be to combine career concerns with other forms of contracting. As the literature review discussed, information disclosure in other contracting environments has quite different effects than in the case of career concerns. One issue likely to arise with multiple periods is that the contracting instrument chosen by the firm would reveal information to the worker in addition to the disclosure policy. For example, the wage paid to the worker under a piece rate contract would indirectly reveal output to the worker.

Finally, information in firms has uses beyond incentive provision. For example, information disclosure potentially allows the worker to learn how to do his job better. Exploring the trade-off between withholding information because of effort risk and disclosing it to build human capital would be another natural extension.

A Omitted Proofs

A.1 Section 2

A.1.1 Lemma 1

Proof. By Bayes' Rule

$$\widehat{\theta}_2^F = \lambda_2 \left(\theta + e_2 + \varepsilon_2 - e_2^F\right) + (1 - \lambda_2) \widehat{\theta}_1$$

and

$$\widehat{\theta}_1^F = \lambda_1 \left(\theta + e_1 + \varepsilon_1 - e_1^F\right) + (1 - \lambda_1) \overline{\theta}$$

Since $\widehat{\theta}_1^F$ is a linear combination of normal random variables, it is itself normal with mean $E\left[\widehat{\theta}_1\right] = \overline{\theta} + \lambda_1 \left(e_1 - e_1^F\right)$ and variance

$$V\left[\widehat{\theta}_1\right] = \lambda_1^2 \left(\frac{h_\theta + h_\varepsilon}{h_\theta h_\varepsilon}\right) = \frac{\lambda_1}{h_\theta}.$$

A standard result in Bayesian statistics (DeGroot 1970) is that $\theta | \widehat{\theta}_1 \sim N\left(\widehat{\theta}_1, \frac{1}{h_\varepsilon + h_\theta}\right)$, so $\widehat{\theta}_2^F | \widehat{\theta}_1$ is normal with mean $\widehat{\theta}_1 + \lambda_2 \left(e_2 - e_2^F\right)$ and variance

$$V\left[\lambda_2 \left(\theta + \varepsilon_2\right)\right] = \lambda_2^2 \left[\frac{1}{h_\varepsilon + h_\theta} + \frac{1}{h_\varepsilon}\right] = \frac{\lambda_2}{h_\varepsilon + h_\theta}.$$

■

A.2 Section 3

One result used in both Propositions 1 and 2 is the following.

Lemma 7 *Suppose there exists a random variable $X \sim N\left(\mu(y), \sigma^2\right)$ where $\mu(y)$ is continuously differentiable and let ϕ be the standard normal density. Then*

$$\frac{\partial}{\partial y} \int_a^\infty f(x)\, dx = \frac{\mu'(y)}{\sigma} \phi\left(\frac{a - \mu(y)}{\sigma}\right).$$

Proof. Using the transformation $v = \frac{x - \mu(y)}{\sigma}$ one can write

$$\frac{\partial}{\partial y} \int_a^\infty f(x)\, dx = \frac{\partial}{\partial y} \int_{\frac{a - \mu(y)}{\sigma}}^\infty \phi(v)\, dv.$$

Applying the Leibnitz Rule for differentiating integrals gives

$$\frac{\mu'(y)}{\sigma} \phi\left(\frac{a - \mu(y)}{\sigma}\right).$$

■

A.2.1 Proposition 1

Proof. First, suppose that $\widehat{\theta}_1 \in \Theta$. The worker's problem can be written as

$$\begin{aligned} & \arg\max_{e_2} -g(e_2) + W \Pr\left[\widehat{\theta}_2^F > \theta^*\right] \\ = \; & \arg\max_{e_2} -\frac{C}{2} e_2^2 + W \int_{\theta^*}^{\infty} f\left(\widehat{\theta}_2^F | \widehat{\theta}_1\right) d\widehat{\theta}_2^F. \end{aligned}$$

Applying Lemma 7, the first order condition for the maximization problem is

$$-Ce_2^W + W\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^* - \widehat{\theta}_1 - \lambda_2\left(e_2^W - e_2^F\right)}{\sigma_2}\right) = 0.$$

For global concavity of the objective function, it must be the case that

$$-C + W\left(\frac{\lambda_2}{\sigma_2}\right)^2 \phi'\left(\frac{\theta^* - \widehat{\theta}_1 - \lambda_2\left(e_2^W - e_2^F\right)}{\sigma_2}\right) < 0$$

holds for all e_2^W and e_2^F. Since $\phi' \leq 1$ this condition is satisfied as long as

$$-C + W\left(\frac{\lambda_2}{\sigma_2}\right)^2 < 0,$$

which holds by Assumption 2. Plugging in the equilibrium condition $e_2^W = e_2^F$ gives the result.

Now, suppose $\widehat{\theta}_1 \notin \Theta$. The worker's maximization problem is

$$\begin{aligned} & \arg\max_{e_2} -\frac{C}{2} e_2^2 + W \Pr\left[\widehat{\theta}_2^F > \theta^*\right] \\ = \; & \arg\max_{e_2} -\frac{C}{2} e_2^2 + W E_{\widehat{\theta}_1}\left\{\Pr\left[\widehat{\theta}_2^F > \theta^* | \widehat{\theta}_1\right] | \widehat{\theta}_1 \notin \Theta\right\}. \end{aligned}$$

Applying the previous steps to the $\Pr\left[\widehat{\theta}_2^F > \theta^* | \widehat{\theta}_1\right]$ term completes the proof. ■

A.2.2 Lemma 2

Proof. Expected second period effort is

$$\begin{aligned} & \int_{\widehat{\theta}_1 \in \Theta} \frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta_2^* - \widehat{\theta}_1}{\sigma_2}\right) f\left(\widehat{\theta}_1\right) d\widehat{\theta}_1 \\ & + E\left[\frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta_2^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \notin \Theta_1\right] \Pr\left[\widehat{\theta}_1 \notin \Theta_1\right] \\ = \; & \int_{-\infty}^{\infty} \frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta_2^* - \widehat{\theta}_1}{\sigma_2}\right) f\left(\widehat{\theta}_1\right) d\widehat{\theta}_1, \end{aligned}$$

which is independent of Θ_1. ■

A.2.3 Lemma 3

Proof.

$$
\begin{aligned}
E\left[(e_2^*)^2 \,|\Theta'\right] &= \Pr\left[\widehat{\theta}_1 \notin \Theta'\right] E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) |\widehat{\theta}_1 \notin \Theta'\right]^2 \\
&\quad + \Pr\left[\widehat{\theta}_1 \in \Theta' \backslash \Theta\right] E\left[\phi^2\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) |\widehat{\theta}_1 \in \Theta' \backslash \Theta\right] \\
&\quad + \Pr\left[\widehat{\theta}_1 \in \Theta\right] E\left[\phi^2\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) |\widehat{\theta}_1 \in \Theta\right] \\
&> \Pr\left[\widehat{\theta}_1 \notin \Theta'\right] E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) |\widehat{\theta}_1 \notin \Theta'\right]^2 \\
&\quad + \Pr\left[\widehat{\theta}_1 \in \Theta' \backslash \Theta\right] E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) |\widehat{\theta}_1 \in \Theta' \backslash \Theta\right]^2 \\
&\quad + \Pr\left[\widehat{\theta}_1 \in \Theta\right] E\left[\phi^2\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) |\widehat{\theta}_1 \in \Theta\right] \\
&> \Pr\left[\widehat{\theta}_1 \notin \Theta\right] E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) |\widehat{\theta}_1 \notin \Theta\right]^2 \\
&\quad + \Pr\left[\widehat{\theta}_1 \in \Theta\right] E\left[\phi^2\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) |\widehat{\theta}_1 \in \Theta\right] \\
&= E\left[(e_2^*)^2 \,|\Theta\right],
\end{aligned}
$$

where the first inequality comes from probability version of Jensen's inequality, and the second from the discrete version. ■

A.2.4 Proposition 2

In order to solve for e_1^* it is necessary to draw on the following claim, which is presented without proof.

Claim 2 *Suppose a uniformly bounded function $h(x,y)$ has a countable number of discontinuity points $\{(x_i, y_i)\}$ that satisfy $x_i = f(y_i)$ where f is continuous. Then*

$$
\frac{\partial}{\partial y}\int_{-\infty}^{\infty} h(x,y)\,dx = \int_{-\infty}^{\infty} \frac{\partial h(x,y)}{\partial y} dx + \sum_i f'(y_i)\left(\lim_{x_i \to f(y_i)^+} h(x,y) - \lim_{x_i \to f(y_i)^-} h(x,y)\right).
$$

Below is the proof of the proposition.

Proof. If $e_1^W \neq e_1^F$ then one can apply Lemma 7 along with equations 3 and 4 to derive

$$e_2^W = \frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^* - \widehat{\theta}_1^W - \lambda_2\left(\left(e_1^W - e_1^F\right) + \left(e_2^W - e_2^F\right)\right)}{\sigma_2}\right)$$

if $\widehat{\theta}_1^F = \widehat{\theta}_1^W + \lambda_1\left(e_1^W - e_1^F\right) \in \Theta$ and

$$e_2^W = E\left[\frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^* - \widehat{\theta}_1^W - \lambda_2\left(\left(e_1^W - e_1^F\right) + \left(e_2^W - e_2^F\right)\right)}{\sigma_2}\right)|\widehat{\theta}_1^W + \lambda_1\left(e_1^W - e_1^F\right) \notin \Theta\right]$$

if $\widehat{\theta}_1^F \notin \Theta$.

Moreover, from Proposition 1,

$$e_2^F = \frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^* - \widehat{\theta}_1^F}{\sigma_2}\right)$$

if $\widehat{\theta}_1^F \in \Theta$ and

$$e_2^F = E\left[\frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \in \Theta\right]$$

if $\widehat{\theta}_1^F \notin \Theta$.

$\widehat{\theta}_2^F|\widehat{\theta}_1^W$ and $e_2^W\left(\widehat{\theta}_1^W\right)$ are discontinuous in $\widehat{\theta}_1^W$, but continuous within the regions $\widehat{\theta}_1^W + \lambda_1\left(e_1^W - e_1^F\right) \in \Theta$ and $\widehat{\theta}_1^W + \lambda_1\left(e_1^W - e_1^F\right) \notin \Theta$. One can separate the objective function into two regions along the following lines

$$\begin{aligned}
& -\frac{C}{2}e_1^2 + \int_{-\infty}^{\infty}\left[\int_{\theta^*}^{\infty} f\left(\widehat{\theta}_2^F|\widehat{\theta}_1^W\right)d\widehat{\theta}_2^F - \frac{C}{2}\left(e_2^W\left(\widehat{\theta}_1^W\right)\right)^2\right] f\left(\widehat{\theta}_1^W\right)d\widehat{\theta}_1^W \\
= \; & -\frac{C}{2}e_1^2 + \int_{\widehat{\theta}_1^W + \lambda_1\left(e_1^W - e_1^F\right)\in\Theta}\left[\int_{\theta^*}^{\infty} f\left(\widehat{\theta}_2^F|\widehat{\theta}_1^W\right)d\widehat{\theta}_2^F - \frac{C}{2}\left(e_2^W\left(\widehat{\theta}_1^W\right)\right)^2\right] f\left(\widehat{\theta}_1^W\right)d\widehat{\theta}_1^W \\
& + \int_{\widehat{\theta}_1^W + \lambda_1\left(e_1^W - e_1^F\right)\notin\Theta}\left[\int_{\theta^*}^{\infty} f\left(\widehat{\theta}_2^F|\widehat{\theta}_1^W\right)d\widehat{\theta}_2^F - \frac{C}{2}\left(e_2^W\left(\widehat{\theta}_1^W\right)\right)^2\right] f\left(\widehat{\theta}_1^W\right)d\widehat{\theta}_1^W.
\end{aligned}$$

One issue that arises when differentiating the above expression with respect to e_1 is that the boundary points separating the sets Θ and $\mathbb{R}\backslash\Theta$ depend on e_1, and these boundary points are precisely where $\left[\int_{\theta^*}^{\infty} f\left(\widehat{\theta}_2^F|\widehat{\theta}_1^W\right)d\widehat{\theta}_2^F - \frac{C}{2}\left(e_2^W\left(\widehat{\theta}_1^W\right)\right)^2\right]$ is discontinuous. Claim 2 resolves this difficulty.

Using Lemma 7, one finds that the first order condition for e_1^W is

$$
\begin{aligned}
Ce_1^W &= \int_{\widehat{\theta}_1^W+\lambda_1\left(e_1^W-e_1^F\right)\in\Theta} \left[\begin{array}{c} W\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^*-\widehat{\theta}_1^W-\lambda_2\left(\left(e_1^W-e_1^F\right)+\left(e_2^W-e_2^F\right)\right)}{\sigma_2}\right)\times \\ \left(1+\left(\frac{\partial e_2^W\left(\widehat{\theta}_1^W\right)}{\partial e_1}\right)_{e_1=e_1^W}-\left(\frac{\partial e_2^F\left(\widehat{\theta}_1^W\right)}{\partial e_1}\right)_{e_1=e_1^W}\right) \\ -Ce_2^W\left(\widehat{\theta}_1^W\right)\left(\frac{\partial e_2^W\left(\widehat{\theta}_1^W\right)}{\partial e_1}\right)_{e_1=e_1^W} \end{array} \right] f\left(\widehat{\theta}_1^W\right)d\widehat{\theta}_1^W \\
&+\int_{\widehat{\theta}_1^W+\lambda_1\left(e_1^W-e_1^F\right)\notin\Theta} \left[\begin{array}{c} W\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^*-\widehat{\theta}_1^W-\lambda_2\left(\left(e_1^W-e_1^F\right)+\left(e_2^W-e_2^F\right)\right)}{\sigma_2}\right)\times \\ \left(1+\left(\frac{\partial e_2^W\left(\widehat{\theta}_1^W\right)}{\partial e_1}\right)_{e_1=e_1^W}-\left(\frac{\partial e_2^F\left(\widehat{\theta}_1^W\right)}{\partial e_1}\right)_{e_1=e_1^W}\right) \\ -Ce_2^W\left(\widehat{\theta}_1^W\right)\left(\frac{\partial e_2^W\left(\widehat{\theta}_1^W\right)}{\partial e_1}\right)_{e_1=e_1^W} \end{array} \right] f\left(\widehat{\theta}_1^W\right)d\widehat{\theta}_1^W \\
&-\lambda_1\sum_i \left(\begin{array}{c} \lim_{\widehat{\theta}_1^W\to\left[a_i-\lambda_1\left(e_1^W-e_1^F\right)\right]^+}\Pr\left[\left(\widehat{\theta}_2^F|\widehat{\theta}_1^W\right)>\theta^*|e_1=e_1^W\right] \\ -\lim_{\widehat{\theta}_1^W\to\left[a_i-\lambda_1\left(e_1^W-e_1^F\right)\right]^-}\Pr\left[\left(\widehat{\theta}_2^F|\widehat{\theta}_1^W\right)>\theta^*|e_1=e_1^W\right] \end{array} \right)
\end{aligned}
$$

where $\{a_i\}$ is the set of all finite points $\underline{x}_i$ and $\overline{x}_i$. One can verify that for high enough C the problem is globally concave, but the details are omitted for the sake of space. The above expression simplifies to

$$
\begin{aligned}
Ce_1^W &= \int_{\widehat{\theta}_1^W+\lambda_1\left(e_1^W-e_1^F\right)\in\Theta} \left[\begin{array}{c} W\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^*-\widehat{\theta}_1^W-\lambda_2\left(\left(e_1^W-e_1^F\right)+\left(e_2^W-e_2^F\right)\right)}{\sigma_2}\right)\times \\ \left(1-\left(\frac{\partial e_2^F\left(\widehat{\theta}_1^W\right)}{\partial e_1}\right)_{e_1=e_1^W}\right) \end{array} \right] f\left(\widehat{\theta}_1^W\right)d\widehat{\theta}_1^W \\
&+\int_{\widehat{\theta}_1^W+\lambda_1\left(e_1^W-e_1^F\right)\notin\Theta} W\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^*-\widehat{\theta}_1^W-\lambda_2\left(\left(e_1^W-e_1^F\right)+\left(e_2^W-e_2^F\right)\right)}{\sigma_2}\right) f\left(\widehat{\theta}_1^W\right)d\widehat{\theta}_1^W \\
&-\lambda_1\sum_i \left(\begin{array}{c} \lim_{\widehat{\theta}_1^W\to\left[a_i-\lambda_1\left(e_1^W-e_1^F\right)\right]^+}\Pr\left[\left(\widehat{\theta}_2^F|\widehat{\theta}_1^W\right)>\theta^*|e_1=e_1^W\right] \\ -\lim_{\widehat{\theta}_1^W\to\left[a_i-\lambda_1\left(e_1^W-e_1^F\right)\right]^-}\Pr\left[\left(\widehat{\theta}_2^F|\widehat{\theta}_1^W\right)>\theta^*|e_1=e_1^W\right] \end{array} \right).
\end{aligned}
$$

Conditional on $\widehat{\theta}_1^W+\lambda_1\left(e_1^W-e_1^F\right)\in\Theta$,

$$
\begin{aligned}
\left(\frac{\partial e_2^F\left(\widehat{\theta}_1^W\right)}{\partial e_1}\right)_{e_1=e_1^W} &= \left(\frac{\partial\left[\frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^*-\widehat{\theta}_1^W-\lambda_1\left(e_1-e_1^F\right)}{\sigma_2}\right)\right]}{\partial e_1}\right)_{e_1=e_1^W} \\
&= \frac{W}{C}\frac{\lambda_1\lambda_2}{\sigma_2^2}\phi'\left(\frac{\theta^*-\widehat{\theta}_1^W-\lambda_1\left(e_1^W-e_1^F\right)}{\sigma_2}\right).
\end{aligned}
$$

Plugging in the equilibrium condition $e_1^W = e_1^F$ into the first order condition yields

$$\begin{aligned} Ce_1^W &= \int_{\widehat{\theta}_1 \in \Theta} \left[\frac{W}{C} \frac{\lambda_2}{\sigma_2} \phi \left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2} \right) \left(1 - \frac{W}{C} \frac{\lambda_1 \lambda_2}{\sigma_2^2} \phi' \left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2} \right) \right) \right] f\left(\widehat{\theta}_1\right) d\widehat{\theta}_1 \\ &\quad + \int_{\widehat{\theta}_1 \notin \Theta} \left[\frac{W}{C} \frac{\lambda_2}{\sigma_2} \phi \left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2} \right) \right] f\left(\widehat{\theta}_1\right) d\widehat{\theta}_1. \end{aligned}$$

The result follows from the fact that ϕ satisfies $\phi'(x) = -x\phi(x)$. ■

A.3 Section 4

A.3.1 Lemma 4

Proof. One can write

$$\begin{aligned} & \Pr\left[\widehat{\theta}_2 \geq \theta^*\right] \overline{W} + \Pr\left[\widehat{\theta}_2 < \theta^*\right] \underline{W} - g\left(e_1^*\right) - E\left[g\left(e_2^*\right)\right] \\ > \;& \min\left\{ \Pr\left[\widehat{\theta}_2 \geq \theta^*\right], \Pr\left[\widehat{\theta}_2 < \theta^*\right] \right\} W - \tfrac{W^2}{2C} \tfrac{\lambda_2^2}{\sigma_2^2} k_1(\Theta) - \tfrac{W^2}{2C} \tfrac{\lambda_2^2}{\sigma_2^2} k_2(\Theta) \end{aligned} \tag{20}$$

where $k_1(\Theta)$ and $k_2(\Theta)$ are bounded above for all Θ. Moreover, by Assumption 2, $\frac{W^2}{2C}\frac{\lambda_2^2}{\sigma_2^2}$ is bounded above by $\frac{1}{2}$. Therefore, as $W \to \infty$ (20) grows arbitrarily large. ■

A.3.2 Proposition 4

Proof. Expected second period effort costs under an arbitrary Θ are

$$\frac{W^2}{2C}\left(\frac{\lambda_2}{\sigma_2}\right)^2 \left\{ \begin{array}{c} \Pr\left[\widehat{\theta}_1 \notin \Theta\right] \left(E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \notin \Theta\right] \right)^2 \\ + \Pr\left[\widehat{\theta}_1 \in \Theta\right] E\left[\phi^2\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \in \Theta\right] \end{array} \right\}$$

The derivative with respect to $\underline{x}_i$ of the expression is equal to

$$\begin{aligned} & \frac{W^2}{2C}\left(\frac{\lambda_2}{\sigma_2}\right)^2 \left\{ \begin{array}{c} 2E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \notin \Theta\right] \phi\left(\frac{\theta^* - \underline{x}_i}{\sigma_2}\right) f\left(\underline{x}_i\right) \\ - \left(E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \notin \Theta\right] \right)^2 f\left(\underline{x}_i\right) - \phi^2\left(\frac{\theta^* - \underline{x}_i}{\sigma_2}\right) f\left(\underline{x}_i\right) \end{array} \right\} \\ = \;& -f\left(\underline{x}_i\right) \frac{W^2}{2C}\left(\frac{\lambda_2}{\sigma_2}\right)^2 \left[E\left(\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \notin \Theta \right) - \phi\left(\frac{\theta^* - \underline{x}_i}{\sigma_2}\right) \right]^2, \end{aligned}$$

with a corresponding expression for the derivative with respect to $\overline{x}_i$.

The derivative of first period social surplus $e_1^* - \frac{C}{2}\left(e_1^*\right)^2$ with respect to $\underline{x}_i$ is

$$-\left(\frac{W}{C}\right)^2 \frac{\lambda_1\left(\lambda_2\right)^2}{\left(\sigma_2\right)^3} \left(\frac{\underline{x}_i - \theta^*}{\sigma_2}\right) \phi^2\left(\frac{\theta^* - \underline{x}_i}{\sigma_2}\right) f\left(\underline{x}_i\right) \xi$$

where

$$\xi = \left[1 - Ce_1^*(\Theta)\right]$$

with a similar expression for the derivative with respect to $\overline{x}_i$.

A necessary condition for Θ_L^* to maximize social surplus is that it satisfy the following three sets of conditions:

$$\begin{aligned} & \frac{\lambda_1}{C\sigma_2}\left(\frac{x-\theta^*}{\sigma_2}\right)\phi^2\left(\frac{\theta^*-x}{\sigma_2}\right)\xi\left(\Theta_L^*\right) \\ = & \left[E\left(\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \notin \Theta_L^*\right)-\phi\left(\frac{\theta^*-x}{\sigma_2}\right)\right]^2 \end{aligned} \tag{21}$$

$\forall x \in \{\underline{x}_i, \overline{x}_i\}$;

$$\begin{aligned} & \frac{\lambda_1}{C\sigma_2}\left(\frac{x-\theta^*}{\sigma_2}\right)\phi^2\left(\frac{\theta^*-x}{\sigma_2}\right)\xi\left(\Theta_L^*\right) \\ \geq & \left[E\left(\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \notin \Theta_L^*\right)-\phi\left(\frac{\theta^*-x}{\sigma_2}\right)\right]^2 \end{aligned} \tag{22}$$

$\forall\ x \in \Theta_L^*$; and

$$\begin{aligned} & \frac{\lambda_1}{C\sigma_2}\left(\frac{x-\theta^*}{\sigma_2}\right)\phi^2\left(\frac{\theta^*-x}{\sigma_2}\right)\xi\left(\Theta_L^*\right) \\ \leq & \left[E\left(\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \notin \Theta_L^*\right)-\phi\left(\frac{\theta^*-x}{\sigma_2}\right)\right]^2 \end{aligned} \tag{23}$$

$\forall\ x \notin \Theta_L^*$.

Conditions (21) say that the marginal gain (loss) in second period welfare equals the marginal loss (gain) in first period welfare at all points $\underline{x}_i$ ($\overline{x}_i$). Next, conditions (22) say that for all disclosed beliefs, the marginal gain from disclosing information must exceed the marginal cost. If there were some x' that did not satisfy this condition, then the disclosure policy $\Theta_H^* \backslash (x', x'+\Delta)$ would strictly dominate Θ_L^* for small enough Δ. Finally, conditions (23) say that for all undisclosed beliefs, the marginal cost of disclosure must exceed the marginal gain.

Suppose that

$$1 > W\frac{\lambda_2}{\sigma_2}\int_{-\infty}^{\infty}\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)f\left(\widehat{\theta}_1\right)d\widehat{\theta}_1$$

so that $Ce_1(\emptyset) < 1$ (the proof works identically if the opposite is true). Clearly $(-\infty, \theta^*) \cap \Theta_L^* = \emptyset$. Now, for an arbitrary Θ not containing $(-\infty, \theta^*)$ there exists a unique point x' such that

$$E\left(\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \notin \Theta\right)-\phi\left(\frac{\theta^*-x'}{\sigma_2}\right)=0$$

since

$$\phi\left(0\right) > E\left(\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \notin \Theta\right)$$

and

$$\lim_{x \to \infty} \phi\left(\frac{\theta^* - x}{\sigma_2}\right) = 0.$$

Clearly

$$\left[E\left(\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \notin \Theta\right) - \phi\left(\frac{\theta^* - x}{\sigma_2}\right)\right]^2$$

is decreasing in x for $x \in (\theta^*, x')$ and increasing in x for (x', ∞).

Now, for an arbitrary Θ

$$\frac{\lambda_1}{C\sigma_2}\left(\frac{x - \theta^*}{\sigma_2}\right)\phi^2\left(\frac{\theta^* - x}{\sigma_2}\right)\xi\left(\Theta\right)$$

equals 0 when $x = \theta^*$, is strictly increasing in x for $x \in (\theta^*, x'')$, strictly decreasing in x for $x \in (x'', \infty)$, and tends to 0 as $x \to \infty$. From these observations one can conclude that for any Θ not containing $(-\infty, \theta^*)$, the points x that satisfy

$$\begin{aligned} & \frac{\lambda_1}{C\sigma_2}\left(\frac{x - \theta^*}{\sigma_2}\right)\phi^2\left(\frac{\theta^* - x}{\sigma_2}\right)\xi\left(\Theta\right) \\ \geq & \left[E\left(\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right) | \widehat{\theta}_1 \notin \Theta\right) - \phi\left(\frac{\theta^* - x}{\sigma_2}\right)\right]^2 \end{aligned} \tag{24}$$

form the set $(x_1, x_2) \subset (\theta^*, \infty)$. For this reason, any Θ that satisfies (21), (22), and (23) must be convex and lie strictly in (θ^*, ∞). To prove that one can always find a Θ that satisfies (21), (22), and (23), consider some $\Theta = (y_1, y_2)$ and let $\widetilde{x}_2$ be the highest x_2 at which (24) holds with equality for all such Θ. Define

$$A \equiv \left\{(a_1, a_2) | \theta^* \leq a_1 \leq a_2 \leq \widetilde{x}_2, (a_1, a_2) \in \mathbb{R}^2\right\}.$$

Consider the mapping $h : A \to A$ that gives the points x_1 and x_2 at which (24) holds with equality as a function of the points y_1 and y_2. Because h is continuous and A is a convex, compact subset of $\mathbb{R}^2$, by Brouwer's Fixed Point Theorem there is some point $(y_1', y_2') \in A$ such that $h\left(y_1', y_2'\right) = (y_1', y_2')$. Hence there exists some $\Theta = (y_1, y_2) \subset (\theta^*, \infty)$ that satisfies (21), (22), and (23). ■

A.3.3 Proposition 6

Proof. Second period equilibrium effort conditional on receiving feedback $\widehat{\theta}_2$ is

$$e_2^* = \left(\frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)\right)^{\frac{1}{\beta}}$$

and conditional on not receiving feedback is

$$e_2^* = E\left[\frac{W}{C}\frac{\lambda_2}{\sigma_2}\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \notin \Theta\right]^{\frac{1}{\beta}}.$$

First period equilibrium effort is

$$e_1^* = \left(\frac{W}{C}\frac{\lambda_2}{\sigma_2}\right)^{\frac{1}{\beta}}\left[\frac{\int_{-\infty}^{\infty}\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)f\left(\widehat{\theta}_1\right)d\widehat{\theta}_1}{\frac{1}{\beta}\frac{W}{C}\frac{\lambda_1\lambda_2}{\sigma_2^2}\int_{\widehat{\theta}_1\in\Theta}\left(\frac{\widehat{\theta}_1-\theta^*}{\sigma_2}\right)\left[\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)\right]^{\frac{\beta+1}{\beta}}f\left(\widehat{\theta}_1\right)d\widehat{\theta}_1}\right]^{\frac{1}{\beta}}$$

Including an interval of disclosed beliefs $\left(\widehat{\theta}_1, \widehat{\theta}_1 + \Delta\right)$ for small Δ in a disclosure policy Θ decreases second period effort by an amount approximately proportional to

$$-\left(\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)\right)^{\frac{1}{\beta}} + \frac{1}{\beta}\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \notin \Theta\right]^{\frac{1-\beta}{\beta}}$$
$$+\left(\frac{\beta-1}{\beta}\right)E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \notin \Theta\right]^{\frac{1}{\beta}}. \quad (25)$$

Whenever $\beta > 1$ this expression is positive. Since giving negative feedback reduces first period effort and expected second period effort, the effort maximizing disclosure policy will not contain any negative feedback. In the range $\widehat{\theta}_1 \in (\theta^*, \infty)$, one can verify that 25 is single troughed and achieves a minimum value of 0 at point $\widetilde{\theta}_1$ satisfying

$$\phi\left(\frac{\theta^* - \widetilde{\theta}_1}{\sigma_2}\right) = E\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)|\widehat{\theta}_1 \notin \Theta\right].$$

Including an interval of disclosed beliefs $\left(\widehat{\theta}_1, \widehat{\theta}_1 + \Delta\right)$ for small Δ in a disclosure policy Θ increases first period effort by an amount approximately proportional to

$$\left(\frac{\widehat{\theta}_1 - \theta^*}{\sigma_2}\right)\left[\phi\left(\frac{\theta^* - \widehat{\theta}_1}{\sigma_2}\right)\right]^{\frac{\beta+1}{\beta}}\left[\frac{\int_{-\infty}^{\infty}\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)f\left(\widehat{\theta}_1\right)d\widehat{\theta}_1}{\frac{1}{\beta}\frac{W}{C}\frac{\lambda_1\lambda_2}{\sigma_2^2}\int_{\widehat{\theta}_1\in\Theta}\left(\frac{\widehat{\theta}_1-\theta^*}{\sigma_2}\right)\left[\phi\left(\frac{\theta^*-\widehat{\theta}_1}{\sigma_2}\right)\right]^{\frac{\beta+1}{\beta}}f\left(\widehat{\theta}_1\right)d\widehat{\theta}_1}\right]^{\frac{1-\beta}{\beta}}$$

One can easily show that this expression is single peaked in $\widehat{\theta}_1$ and that it approaches 0 for $\widehat{\theta}_1 \to \theta^*$ and $\widehat{\theta}_1 \to \infty$. Thus one can apply exactly the same argument as in the proof of Proposition 4 to show that the effort maximizing disclosure policy when $\beta > 1$ will take the form $\Theta = (\theta', \theta'')$ where $\theta^* < \theta' < \theta'' < \infty$.

Similar arguments show why the form of the social welfare maximizing disclosure policy again takes the form of Proposition 4. ■

A.4 Section 5

A.4.1 Proposition 7

Proof. Conjecture an equilibrium pure strategy $\widetilde{w}^{F*}$, and let $\Theta_2' = \left\{\widehat{\theta}_2 | \widetilde{w}^{F*}\left(\widehat{\theta}_2\right) \neq \emptyset\right\}$. Now, suppose that in equilibrium $\exists \widehat{\theta}_2', \widehat{\theta}_2'' \in \Theta_2'$ such that $\widehat{\theta}_2' \neq \widehat{\theta}_2''$ and, without loss of generality, $\widetilde{w}^{F*}\left(\widehat{\theta}_2'\right) < \widetilde{w}^{F*}\left(\widehat{\theta}_2''\right)$. Since $\widetilde{w}^{F*}$ is an equilibrium it must be the case that

$$\kappa + k\widehat{\theta}_2 - \widetilde{w}^{F*}\left(\widehat{\theta}_2\right) - \delta \geq 0 \text{ for } \widehat{\theta}_2 = \widehat{\theta}_2', \widehat{\theta}_2'', \text{ and}$$

$$\widetilde{w}^{F*}\left(\widehat{\theta}_2\right) \geq E\left[k\theta | \widetilde{w}^{F*}\left(\widehat{\theta}_2\right), \widehat{w} = \widetilde{w}^{F*}\right] - \delta \text{ for } \widehat{\theta}_2 = \widehat{\theta}_2', \widehat{\theta}_2''.$$

The first condition says that the worker yields F non-negative profits. The second condition says that the worker is paid at least his outside option, since if he were not some market firm would bid him away and F would have expended δ for no gain. So, regardless of whether the worker is paid $\widetilde{w}^{F*}\left(\widehat{\theta}_2'\right)$ or $\widetilde{w}^{F*}\left(\widehat{\theta}_2''\right)$, the outside market does not bid for him. But since

$$\kappa + k\widehat{\theta}_2'' - \widetilde{w}^{F*}\left(\widehat{\theta}_2'\right) - \delta > \kappa + k\widehat{\theta}_2'' - \widetilde{w}^{F*}\left(\widehat{\theta}_2''\right) - \delta \geq 0,$$

F strictly prefers to offer a worker of type $\widehat{\theta}_2''$ a wage $\widetilde{w}^{F*}\left(\widehat{\theta}_2'\right)$. So in any pure strategy equilibrium $\widetilde{w}^{F*}\left(\widehat{\theta}_2\right)$ is constant $\forall \widehat{\theta}_2 \in \Theta_2'$. That is, all workers offered a wage contract by F are offered the same wage.

$\widetilde{w}^{F*}$ must then satisfy

$$\kappa + k\theta^* - \widetilde{w}^{F*}\left(\theta^*\right) - \delta = 0, \text{ where } \theta^* = \inf \Theta_2', \text{ and}$$

$$\widetilde{w}^{F*}\left(\widehat{\theta}_2\right) = E\left[k\theta | \widehat{\theta}_2 \geq \theta^*\right] - \delta \; \forall \widehat{\theta}_2 \in \Theta'.$$

The first condition says that the minimum retained type must yield F zero profits, since otherwise it profitably retain a type $\theta^* - \varepsilon$ for small enough ε or else not make a wage offer to type θ^* and save δ. If θ^* gives F zero profit, all types $\widehat{\theta}_2 > \theta^*$ yield positive profit, so all such types are retained. In equilibrium outside firms must correctly infer this retention rule. To maximize profit, F should exactly offer the worker his expected output with any outside firm. Plugging the first condition above into the second gives the following equation for the equilibrium threshold:

$$\kappa + k\theta^* = E\left[k\theta | \widehat{\theta}_2 \geq \theta^*\right].$$

It remains to be shown that θ^* exists and is unique. Two helpful results are that if $X \sim N\left(\mu, \sigma^2\right)$ then $E\left[X | X \geq a\right] = \mu + \sigma\gamma\left(\frac{a-\mu}{\sigma}\right)$ where γ is the normal hazard rate, and $\gamma' \in (0,1)$ over the entire domain of γ (Greene 2003, p.759). Therefore, the derivative of the left hand side of the above equation is bigger than the derivative of the right hand side. So, if a solution exists it is unique. As $\theta^* \to -\infty$ the left hand side of the above equation is clearly less than the right hand side.

One can also show that

$$\lim_{\theta^* \to \infty} \theta^* - E\left[\theta | \widehat{\theta}_2 \geq \theta^*\right] = 0$$

so that

$$\lim_{\theta^* \to \infty} \kappa + k\theta^* - E\left[k\theta | \widehat{\theta}_2 \geq \theta^*\right] > 0.$$

This completes the proof. ■

A.4.2 Lemma 6

Proof. By definition, Θ_L^* is the social surplus maximizing disclosure policy. The equations that defines $\underline{x}$ and $\overline{x}$ under Θ_L^* are 21 evaluated at $x = \underline{x}$ and $x = \overline{x}$. The cost of effort parameter C does not play a role in defining θ^*, so one can vary it without changing the right hand side of 21. As $C \to \infty$ the left hand side of 21 approaches 0 for any Θ and x since $\frac{\lambda_1}{C\sigma_2}$ becomes arbitrarily small and $Ce_1^*\left(\Theta_L^*\right)$ is bounded above for all C. This implies that $\underline{x} - \overline{x} \to 0$ since 21 has a unique solution for x is the left hand side is 0. ■

References

Abreu, D., P. Milgrom, and D. Pearce (1991): “Information and Timing in Repeated Partnerships,” *Econometrica*, 59(6), 1713–33.

Aoyagi, M. (2007): “Information Feedback in a Dynamic Tournament,” mimeo, The Institute of Social and Economic Research, Osaka University.

Baker, G., M. Gibbs, and B. Holmstrom (1994a): “The Internal Economics of the Firm: Evidence from Personnel Data,” *The Quarterly Journal of Economics*, 109(4), 881–919.

——— (1994b): “The Wage Policy of a Firm,” *The Quarterly Journal of Economics*, 109(4), 921–55.

Baron, J. N., and D. M. Kreps (1999): *Strategic Human Resources: Frameworks for General Managers*. John Wiley, New York, New York.

Beer, M. (1987): “Performance Appraisal,” in *Handbook of Organizational Behavior*, ed. by J. W. Lorsch, pp. 286–300. Prentice Hall, Englewood Cliffs, New Jersey.

Crawford, V. P., and J. Sobel (1982): “Strategic Information Transmission,” *Econometrica*, 50(6), 1431–51.

DeGroot, M. H. (1970): *Optimal Statistical Decisions*. McGraw-Hill, New York, New York.

Dewatripont, M., I. Jewitt, and J. Tirole (1999): “The Economics of Career Concerns, Part I: Comparing Information Structures,” *Review of Economic Studies*, 66(1), 183–98.

Ederer, F. (2008): “Feedback and Motivation in Dynamic Tournaments,” mimeo, MIT.

Fama, E. F. (1980): “Agency Problems and the Theory of the Firm,” *Journal of Political Economy*, 88(2), 288–307.

Fuchs, W. (2007): “Contracting with Repeated Moral Hazard and Private Evaluations,” *American Economic Review*, 97(4), 1432–1448.

Gibbs, M. (1991): “An Economic Approach to Process in Pay and Performance Appraisals,” mimeo, University of Chicago GSB.

——— (1995): “Incentive Compensation in a Corporate Hierarchy,” *Journal of Accounting and Economics*, 19(2-3), 247–277.

Goltsman, M., and A. Mukherjee (2008): "Information Disclosure in Multistage Tournaments," mimeo, University of Western Ontario.

Greene, W. H. (2003): *Econometric Analysis*. Prentice Hall, Upper Saddle River, New Jersey.

Holmstrom, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10(1), 74–91.

——— (1999): "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66(1), 169–82.

Jackman, J. M., and M. H. Strober (2003): "Fear of Feedback," *Harvard Business Review*, 81(4), 101–07.

Kersley, B., C. Alpin, J. Forth, A. Bryson, H. Bewley, G. Dix, and S. Oxenbridge (2006): *Inside the Workplace: Findings from the 2004 Workplace Employment Relations Survey*. Routledge, Abingdon, Oxfordshire.

Kovrijnykh, A. (2007): "Career Uncertainty and Dynamic Incentives," mimeo, University of Chicago.

Landers, R. M., J. B. Rebitzer, and L. J. Taylor (1996): "Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms," *American Economic Review*, 86(3), 329–48.

Lazear, E. P., and M. Gibbs (2008): *Personnel Economics in Practice*. Wiley, Hoboken, NJ, second edn.

Lizzeri, A., M. A. Meyer, and N. Persico (2002): "The Incentive Effects of Interim Performance Evaluations," Penn CARESS Working Papers 02-09.

Longenecker, C. ., H. P. Sims, and D. A. Gioia (1987): "Behind the Mask: The Politics of Employee Appraisal," *The Academy of Management Executive*, 1(3), 183–93.

MacLeod, W. B. (2003): "Optimal Contracting with Subjective Evaluation," *American Economic Review*, 93(1), 216–240.

Maister, D. H. (1993): *Managing the Professional Service Firm*. Free Press Paperbacks, New York, New York.

Martinez, L. (2008): "Reputation, Career Concerns, and Job Assignments," mimeo, Federal Reserve Bank of Richmond.

Medoff, J. L., and K. G. Abraham (1980): "Experience, Performance, and Earnings," *The Quarterly Journal of Economics*, 95(4), 703–36.

Meyer, H. H., E. Kay, and J. R. P. French (1965): "Split Roles in Performance Appraisal," *Harvard Business Review*, 43(1), 123–29.

Milkovich, G. T., J. M. Newman, and C. Milkovich (2007): *Compensation*. McGraw-Hill, New York, New York, ninth edn.

Murphy, K. J. (1992): "Performance Measurement and Appraisal: Motivating Managers to Identify and Reward Performance," in *Performance Measurement, Evaluation, and Incentives*, ed. by W. J. Bruns, pp. 37–62, Boston, Massachusetts.

Murphy, K. M., and J. N. Cleveland (1991): *Performance Appraisal: An Organizational Persepctive*. Allyn and Bacon, Needham Heights, Massachusetts.

Murphy, T. H., and J. Margulies (2004): "Performance Appraisals," Presentation, ABA Labor and Employment Law Section, Equal Employment Opportunity Committee, Mid-Winter Meeting.

Prat, A. (2005): "The Wrong Kind of Transparency," *American Economic Review*, 95(3), 862–877.

Prendergast, C., and R. H. Topel (1996): "Favoritism in Organizations," *Journal of Political Economy*, 104(5), 958–78.

Waldman, M. (1984): "Job Assignments, Signalling, and Efficiency," *RAND Journal of Economics*, 15(2), 255–267.