Queen's Economics Department Working Paper No. 908

Incentives, Team Production, Transaction Costs, and the Optimal Contract: Estimates of an Agency Model using Payroll Records

Christopher Ferrall   Bruce Shearer

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

8-1994

Discussion Paper #908

Incentives, Team Production, Transaction Costs, and the Optimal Contract: Estimates of an Agency Model using Payroll Records

by

Christopher Ferrall
Queen's University

and

Bruce Shearer
Laval University

August 1994

Incentives, Team Production, Transactions Costs, and the Optimal Contract: Estimates of an Agency Model using Payroll Records

Christopher Ferrall
Department of Economics
Queen's University
Kingston, Ontario K7L 3N6
CANADA

ferrall@qed.econ.queensu.ca

Bruce Shearer
Department of Economics
Laval University
CRÉFA and CIRANO
Quebec, Quebec G1K 7P4
CANADA

bshearer@ecn.ulaval.ca

August 1994

We acknowledge research support from the Centre for Resource Studies at Queen's University and Ferrall acknowledges support from the Social Science and Humanities Research Council of Canada. For helpful comments and suggestions we thank Dan Bernhardt, Robert Gibbons, Mary Margiotta, Harry Paarsch, Daniel Parent, Craig Riddell, and participants at the conference "Incentives and Contracts: Theory and Evidence" held at the University of Montreal.

Abstract: We apply agency theory to the payroll records of a copper mine that paid a production bonus to teams of workers. As with most incentive pay used by firms, the bonus was simpler in form than the optimal contract that balances incentives, insurance, and free-riding. We explore whether transactions costs help explain this discrepancy. We estimate an agency model for the payroll data using the method of maximum likelihood and find that incentives and free-riding within teams accounted for two-thirds of the bonus system's inefficiency relative to potential full information profits. The remaining one-third of the inefficiency is attributed to the form of the incentive contract as constrained by transactions costs. We discuss alternative explanations and the general empirical content of agency theory.

JEL Classification: L2, D2, J3, C4

Keywords: Principal-Agent Models, Transactions Costs, Performance Pay, Maximum Likelihood Estimation

1. Introduction

A growing literature uses firm-level data to study whether incentives play an important role in the design and performance of contracts. Examples include Jensen and Murphy (1990), Margiotta and Miller (1993), and the papers contained in Ehrenberg (1990) and Blinder (1990). A discrepancy exists between the practice and theory of compensation that requires no empirical analysis to uncover. Firms typically use compensation schemes that are simpler in form than the optimal contract arising from principal-agent models. In agency models, the optimal contract balances two elements of compensation: insurance and incentives. Team production adds the third element of free-riding (Holmstrom 1982). Insurance and incentive concerns lead to an optimal contract that is typically non-linear and that rarely has a closed form (Grossman and Hart 1983, Gibbons 1987). The optimal contract also involves all observable characteristics that are informative about agent behavior (Holmstrom 1979).

Yet compensation schemes based upon mathematically complicated formulas involving all relevant information are typically not observed (Stiglitz 1991). To explain why firms use simple incentive schemes such as piece rates or bonuses, factors other than incomplete information and risk aversion must be introduced into the economic environment. The payment scheme could be supported by implicit arrangements that achieve the same result as the optimal scheme in standard agency theory. Alternatively, transactions costs may limit the degree of complexity of payment schemes. Holmstrom and Milgrom (1987) argue that linear rules are robust to variations in the productive environment. Linear rules may therefore require less costly tinkering than the "optimal" contract of standard agency theory.

This paper considers whether a particular type of transaction cost, namely the cost of implementing payment schemes, can explain why incentive contracts, whether linear or not, often are "too simple." Allowing for implementation costs can assume away the discrepancy between practice and theory unless a structure is imposed on implementation costs based on observed features of contracts. We posit that mathematically complicated contracts are more costly to implement than simpler contracts. Examples of implementation costs include the resources required to communicate the contract to agents, to keep track of the required information, and to compute payments under the contract. We measure empirically the relative importance of incomplete information, implementation costs, and free riding within teams using the payroll records of a copper mine. We also test for the presence of implicit arrangements which the firm may have used to make the explicit pay system as efficient as the optimal agency contract.

During the 1920s, the Britannia Mining and Smelting Company of British Columbia paid teams of workers a production bonus. Teams whose output exceeded a minimum standard received a bonus proportional to output beyond the standard. If y is team output, x is the production standard, and α is the bonus rate, then the bonus equals zero when $y < x$ and $\alpha(y-x)$ when $y \geq x$. We call this contract a *linear bonus*. From the payroll records we observe the payments made under the linear bonus, but we do not observe the values of α and x used by the firm in any pay-period.

We develop, solve, and estimate a principal-agent model that captures essential elements of technology and information inside a mine: worker risk aversion, team production, and asymmetric information between workers and the firm about working conditions. Using Britannia's payroll records, we compute maximum likelihood estimates of the model's parameters. Our estimation strategy is similar to that of Pakes (1986), Rust (1987), Eckstein and Wolpin (1990), and Margiotta and Miller (1993) in the sense that estimating the model requires a nested solution algorithm. There is no closed-form solution for optimal values of α and x, let alone the fully optimal contract. The firm's problem must be solved numerically on each iteration of the estimation procedure. As with Rust's application of dynamic control theory, we use multiple realizations of production shocks within a single firm to identify the model. As with the equilibrium search model estimated by Eckstein and Wolpin, our algorithm solves the problems for two sides of the contract or market. Margiotta and Miller estimate an agency model using data on executive compensation. Their model concerns incentives and dynamics. Our model concerns incentives and transactions costs, which is perhaps a more appropriate focus when studying production workers with no long-term commitments to the firm.

One of our results concerns the empirical content gained by solving the principal's problem numerically. Intuitively, it is unlikely that technology and preference parameters can be disentangled unless the choices of both the principal and agent are modeled. Without separately identifying the parameters of the agency model, little can be inferred about the performance of the bonus system. We provide a proof that formalizes this intuition. By numerically maximizing the firm's profit function, we can identify parameters that determine the cost of incomplete information and indirectly the implementation costs that rationalize the use of the linear bonus system. Our identification results are specific to our model, yet they illustrate that numerical methods are typically necessary for analyzing payroll data in light of principal-agent theory.

Our estimates indicate that there were larger costs associated with incomplete information in the Britannia mine than a more casual analysis might suggest. We estimate that the percentage of full information profits lost under the linear bonus system was between 70 and 80 percent, even though workers received less than 5 percent of their wages in the form of incentive pay. Two thirds of the efficiency loss, or about 50 percent of full information profits, would occur under the optimal contract with incomplete information that ignores implementation costs. The remaining 25 percent of the inefficiency is due to the error in approximating the fully optimal contract under incomplete information with the optimal linear bonus.

In the next section we describe our notion of implementation costs. In section 3 we describe the agency model, and as a baseline we characterize the optimal contract under full information. Then we describe the response of workers to the linear bonus system under asymmetric information, and we derive the firm's objective function for choosing the optimal linear bonus. Section 4 discusses identification of the model with and without modeling firm behavior. Section 5 describes the data and reports estimates of the model. Section 6 uses the estimates to consider implications, limitations, and extensions of the results. Section 7 concludes.

2. Preliminaries: Implementation Costs

Agency theory provides a framework for modeling the costs of incomplete information, but we lack a good model of implementation costs. We avoid the assumption that simple contracts are used because agents or the firm act sub-optimally or that they incur costs to calculate optimal choices. Instead we restrict our notion of implementation costs to the time and effort required to communicate the contract to workers and to the costs of calculating payments under the scheme. Some authors have studied the choice between methods of pay by comparing contracts with simple functional forms. Lazear (1986) and Brown (1990) compare salaries to piece rates, while Lazear and Rosen (1981) and Green and Stokey (1983) compare piece rates to tournaments. The restriction to simple contracts can be based on the assumption that more complicated payment schemes are very costly to implement.

We assume that implementation costs of paying workers according to contracts of the same functional form are equal, and that mathematically more complicated forms are more costly. For instance, the implementation costs of a piece rate scheme do not depend on the piece rate itself. And contracts that contain, say, logarithms are assumed to be more expensive than piece rates, both to communicate to workers and to calculate on a regular basis. Contracts that depend on more variables, such as both past and current performance, require more bookkeeping and are therefore more expensive to implement than contracts based on fewer variables.

Conceptually, we break the firm's choice into two steps:

$$\max_{z \in Z} \left[\max_{g \in G(z)} \pi(g, z) \right] - C(z).$$

Here Z denotes the family of all possible enforceable compensation contracts. An element $z \in Z$ is a form of payment, with specific contracts in z described by a set of parameters g. $G(z)$ denotes the set of possible values of g. For example, if z is the linear bonus scheme used by our firm, then an element of z is defined by the two parameters (α, x), and $G(z) \equiv \{(\alpha, x) : (\alpha, x) \in \mathcal{R}^2_+\}$. $C(z)$ denotes the implementation costs of contracts in z and $\pi(g, z)$ denotes the profits associated with a particular fully specified contract, (g, z). By construction $C(z)$ is a fixed cost when choosing $g \in G(z)$. The profit function $\pi(g, z)$ incorporates the agency relationship with workers, namely incentive compatibility, individual rationality, and free-riding within teams. The firm can be viewed as choosing the optimal parameters $g^\star(z)$ and then choosing the optimal contract form $z^\star$. The classic principal-agent problem, in Grossman and Hart (1983) for example, would be the case $C(z) = 0$ for all forms of payment z. A prior restriction to piece rates would be the case that the cost of implementing any non-linear contract outweighs any gain in the profit function $\pi(g, z)$.

We do not attempt to estimate $C(z)$ directly. Rather, by estimating $g^\star(z^\star)$ given the firm's payment scheme $z^\star$, we can potentially identify the main parameters that determine the profit function $\pi(g, z)$: worker risk aversion, cost of effort, and the distribution of production shocks. We then compare costs of incomplete information with the implementation costs needed to rationalize the choice of $z^\star$. For our application, we compute expected profits under simpler and more complicated compensation schemes than the linear bonus used by the firm. We also approximate numerically the optimal agency contract ignoring implementation costs.

3. Team Production and Incentives in Mining

Several aspects of mining make it an ideal industry to which to apply agency theory. Productivity varies substantially due to working conditions, space constraints make complete monitoring difficult, and the production process is simple. Two primary occupations, miners and muckers, were involved in ore extraction at Britannia. Miners drilled and blasted rock from the face of the tunnel while muckers shoveled the blasted rock (or muck) into ore carts. We assume that each team was composed of one miner and two muckers, because the ratio of shifts worked by muckers to miners varies between 1.5 and 2 in all pay periods. Let the subscript a denote miners and let b denote muckers. Since muckers can only muck rock blasted by the miner, output from a tunnel can be approximated by a Leontief production function,

$$y = \theta \min\{\lambda_a, \lambda_{b,1} + \lambda_{b,2}\} \tag{1}$$

where y is the amount of ore produced, λ_a is the effort of the miner, $\lambda_{b,i}$ is the effort of mucker i, and θ is a random shock to productivity in the tunnel lasting one pay period. Effort can be interpreted as the amount of ore processed by the worker. Below we also account for permanent differences in productivity across areas of the mine as well transitory differences captured by θ.

Workers observe the realization of θ before choosing their level of effort. Each worker has a utility function of the form

$$U(W - \frac{k_i}{2}\lambda^2) \qquad i \in \{a, b\},$$

where W is the wage earned in a period, λ is effort, and $(k_i/2)\lambda^2$ is the quadratic cost of effort for workers in occupation i. U is a von Neumann-Morgenstern utility function satisfying $U' > 0$ and $U'' < 0$; that is, workers are risk averse and they are willing to pay for insurance against production shocks. Since workers observe shocks before choosing effort, the functional form of U does not affect the choice of effort or the firm's problem with full information. U

does, however, determine the firm's choice over incentive contracts, so later we specify its form.

Assumption A1.

(i) θ *is log-normally distributed:* $\ln\theta \sim N(\mu, \sigma^2)$.

(ii) $k_a < k_b$.

We denote the density and cumulative distribution functions of θ as $f(\theta)$ and $F(\theta)$, respectively. From A1.i, $f(\theta) = \frac{1}{\theta\sigma}\phi((\ln\theta - \mu)/\sigma)$ and $F(\theta) = \Phi((\ln\theta - \mu)/\sigma)$ where ϕ and Φ are the standard normal density and distribution functions. We assume that that miners have lower effort costs than muckers (A1.ii), because most workers started as muckers, and because miners were paid a higher base wage than muckers. These facts suggest that workers assigned to mining were more skilled than muckers. Furthermore, efficient production under (1) requires the miner to process twice as much ore as each mucker, so within the model it is optimal to assign the task of mining to higher skilled workers.

To establish a baseline, consider the case when the firm can also observe θ before λ_i is chosen. The optimal full-information contract specifies a wage and effort level for each occupation and each value of θ.

Definition D1. *The optimal full information contract (when* θ *can be observed) is described by two wage functions,* W_a *and* W_b*, and two effort functions,* λ_a *and* λ_b*, that solve:*

$$\max_{W_a, W_b, \lambda_a, \lambda_b} \int_0^\infty \Big(\theta \min\{\lambda_a(\theta), 2\lambda_b(\theta)\} - W_a(\theta) - 2W_b(\theta)\Big) f(\theta) d\theta$$

subject to

$$\text{for} \quad i \in \{a, b\} \qquad \int_0^\infty U\Big(W_i(\theta) - \frac{k_i}{2}\lambda_i(\theta)^2\Big) f(\theta) d\theta = \bar{u}_i$$

where $\bar{u}_i$ *is the reservation utility level for members of occupation* i.

Theorem T1. *The optimal full information contract takes the form*

$$\lambda_a(\theta) = \frac{2\theta}{2k_a + k_b}$$
$$\lambda_b(\theta) = \lambda_a(\theta)/2 = \frac{\theta}{2k_a + k_b}$$
$$W_a(\theta) = U^{-1}(\bar{u}_a) + \frac{k_a}{2k_a + k_b}\, y$$
$$W_b(\theta) = U^{-1}(\bar{u}_a) + \frac{k_b}{4(2k_a + k_b)}\, y$$

Proof: All proofs are provided in the Appendix.

With full information, optimal wages consist of a piece rate combined with a base wage or, if the constant term $U^{-1}(\bar{u}_i)$ is negative, a base fee to enter the mine. The contract does not have a constant wage because production shocks affect the productivity of effort, so it is optimal for worker effort to vary with θ. To provide insurance against this variation, wages vary with output. The full insurance wage is linear in output because the cost of effort is quadratic. This is an attractive feature of the model, since we apply it to data generated by a wage contract that is not linear but rather piecewise linear in output. A non-linear contract is only useful if an incentive problem exists. A non-linear contract does not approximate a more complicated full information contract which might be costly to implement.

Corollary C1. *Under Assumption A1, expected output and profits per team under the full information optimal contract equal*

$$E[y] = \int_0^\infty \theta\lambda_a(\theta)f(\theta)d\theta = \frac{1}{k_a + \frac{1}{2}k_b}\exp\{2\mu + 2\sigma^2\}$$
$$E[\pi] = E[y]/2 - U^{-1}(\bar{u}_a) - 2U^{-1}(\bar{u}_b).$$

Corollary C1 shall be used to assess the costs of incomplete information in terms of production efficiency.

Asymmetric Information on θ and the Linear Bonus System

It is reasonable to assume that directly observing worker effort or production conditions in each area of a mine requires significant monitoring costs.

Now consider contracts when the firm only observes team output, y, and therefore can only enforce wage payments that depend on output and occupation. The optimal contract now respects incentive compatibility constraints:

Definition D2. *Under asymmetric information on θ, the optimal team contract solves*

$$\max_{W_a, W_b, \lambda_a, \lambda_b} \int_0^\infty \Big(\theta \min\{\lambda_a(\theta), 2\lambda_b(\theta)\} - W_a(\theta) - 2W_b(\theta)\Big) f(\theta) d\theta$$

subject to

for $i \in \{a, b\}$ $$\lambda_i(\theta) \in \arg\max_{\lambda} \quad U(W_i(\theta) - \frac{k_i}{2}\lambda^2)$$

for $i \in \{a, b\}$ $$\int_0^\infty U\Big(W_i(\theta) - \frac{k_i}{2}\lambda_i(\theta)^2\Big) f(\theta) d\theta = \bar{u}_i$$

The form of the optimal wage contract is unknown, but the solution is not a piece rate as with full information, nor is it the pay system that the firm used. For one thing, the two occupations should be paid different rates because their costs and productivities differ.[1] In section 6 we compute for purposes of comparison a numerical approximation to the optimal contract.

The firm used what we call a linear bonus. Workers in each occupation were paid base wages, denoted β_a and β_b. A team working in some area j of the mine split equally a bonus of the form

$$w_j = \begin{cases} 0 & \text{if } y_j < x_j \\ \alpha_j(y_j - x_j) & \text{if } y_j \geq x_j \end{cases} \tag{2}$$

where x_j is the standard and α_j is the bonus rate for sector j. A team member in occupation i therefore received a total wage equal to $\beta_i + w_j/3$. The production standards and piece rates differed across areas. The standard was also adjusted

[1] Early on, the firm experimented with different bonus rates for different occupations. There appears to have been resistance to this and by 1926 the pay system was changed so that workers split a team bonus equally. We return to this issue in section 6.

for the number of shifts worked in the area during the pay period.[2] It is straightforward to augment the production function in (1) to take into account differences across sectors. That is, let output in sector j take the form

$$y_j = d_j \theta_j \min\{\lambda_a, \lambda_{b,1} + \lambda_{b,2}\} + \nu_j \tag{3}$$

where the d_j and ν_j determine the observable rock quality and other elements of production in sector j. Let d_j and ν_j be the realization of two random variables, d and ν. We make three technical assumptions about the productivity shocks.

Assumption A2.

(i) d, ν *and* θ *are independently distributed;*

(ii) $E[d] = 1$;

(iii) $E[\nu] > 0$.

The firm can set the standards and piece rates to cancel out fixed differences across sectors by choosing values α and x such that

$$\alpha_j = \frac{\alpha}{d_j}$$
$$x_j = d_j x + \nu_j.$$

This system of bonuses and standards equalizes opportunities across areas of the mine prior to the realization of θ. Company reports suggest that balancing outcomes across sectors was important to the firm, perhaps for the following reason: If work in one area is especially difficult, because of bad rock or equipment malfunctions, then the mine cannot easily shut the area down and shift work to other areas. The marginal product of placing extra workers in other tunnels is small due to space constraints, and only by making progress in a bad area can miners reach better rock. Risk averse workers should be sheltered from the risk associated with placement within the mine. Given a system of equalizing bonuses defined by (α, x), we drop for the time being the sector index j from equation (3) and return to the simpler production function (1).

[2] Workers might work in different areas during a pay period, and they were rewarded a share of the bonus in each area in proportion to the number of shifts they had worked in that area.

Effort Levels Within Teams

Figure 1 sketches the output of a team that arises from a Nash equilibrium response to the linear bonus system (2). Each worker chooses his effort level to maximize utility conditional on θ, the parameters of the bonus system (α, x), and the behavior of the other members of his team. A bonus system gives workers no incentive to provide effort when the value of θ falls below some value $\theta^\star$. For $\theta < \theta^\star$, all members of the team set effort and output to zero, because working conditions make it too difficult to earn a bonus. For $\theta > \theta^*$ each worker wants to equate the marginal return to effort to marginal cost. The nature of the production function, however, implies that any effort one occupation supplies above and beyond the effort of the other occupation is wasted. Therefore, in equilibrium, miners will always supply twice the effort level of muckers.

Theorem T2. *Given a bonus system (α, x), there exists a continuum of Nash equilibria effort functions for miners and muckers, denoted $\lambda_a(\theta)$ and $\lambda_b(\theta)$. In each case, $\lambda_b(\theta) = \lambda_a(\theta)/2$. Define $\theta_a^\star = \sqrt{6k_a x/\alpha}$ and $\theta_b^\star = \sqrt{2k_b x/\alpha}$. Then in the unique Pareto efficient Nash equilibrium, the effort function for miners takes the form*

(i) if $k_a \leq k_b < 2k_a$ then
$$\lambda_a(\theta) = \begin{cases} \frac{\alpha\theta}{3k_a} & \text{if } \theta > \theta_a^\star \\ 0 & \text{otherwise} \end{cases}$$

(ii) if $2k_a \leq k_b < 4k_a$ then
$$\lambda_a(\theta) = \begin{cases} \frac{2\alpha\theta}{3k_b} & \text{if } \theta > \sqrt{\frac{3k_b^2 x}{2\alpha(k_b - k_a)}} \\ 0 & \text{otherwise} \end{cases}$$

(iii) if $0 < 4k_a \leq k_b$ then
$$\lambda_a(\theta) = \begin{cases} \frac{2\alpha\theta}{3k_b} & \text{if } \theta > \theta_b^\star \\ 0 & \text{otherwise.} \end{cases}$$

If workers cooperate with each other to maximize total team wages (net of effort costs), then the miner's effort function takes the form

(iv)
$$\lambda_a^c(\theta) = \begin{cases} \frac{2\alpha\theta}{2k_a + k_b} & \text{if } \theta > \theta_c^\star \\ 0 & \text{otherwise} \end{cases}$$

where $\theta_c^\star = \sqrt{\frac{(2k_a+k_b)x}{\alpha}}$.

There are multiple Nash solutions because the effort levels of the two muckers enter y additively, creating a range of θ for which the two muckers may stop shirking simultaneously. The range depends on the effort level of the miners, but the lower bound equals $\theta_b^\star$. For $\theta < \theta_b^\star$ each mucker would set effort to zero in response to any effort level chosen by the miner and the other mucker. In turn, $\theta_a^\star$ is the lowest value θ for which miners choose to shirk even if the muckers are willing to work hard enough to earn a bonus. In each of the cases in Theorem T2, one of the occupations constrains the effort level of the whole team. Which occupation constrains team effort depends on the utility parameters k_a and k_b, leading to the cases T2.(i)–T2.(iii). Define $\theta^\star$ to equal the value of θ at which the team stops shirking in the efficient Nash equilibrium:

$$\theta^\star = \begin{cases} \theta_a^\star & \text{if } k_a \leq k_b < 2k_a \\ \sqrt{\frac{3k_b^2 x}{2\alpha(k_b-k_a)}} & \text{if } 2k_a \leq k_b < 4k_a \\ \theta_b^\star & \text{if } 0 < 4k_a \leq k_b. \end{cases}$$

The equilibrium in which shirking stops at $\theta^\star$ is Pareto efficient because the utility of all team members is highest in this equilibrium, holding constant the bonus system.

The cooperative effort function $\lambda_a^c(\theta)$ given in T2.(iv) is the solution when implicit aspects of compensation are somehow used to overcome the problem of free-riding within teams. The firm or the workers themselves may support the cooperative outcome by punishing workers in the future who act in the non-cooperative fashion. We do not attempt to model the mechanism that might support this cooperative solution, but presumably the team could only be induced to consider the team's share of output under the explicit contract. Later we show that under certain cases this cooperative outcome can be distinguished from the non-cooperative solution. This provides a test for the presence of implicit compensation that supports the explicit bonus system.

The Optimal Linear Bonus

Theorems T1 and T2 hold for any U that is concave and increasing. To solve the firm's problem, we specify U to have constant absolute risk aversion equal to $r > 0$.[3] For occupation $i \in \{a, b\}$,

$$U = -\exp\Big\{-r(\beta_i + w(y)/3 - k_i\lambda_i^2/2)\Big\}. \tag{4}$$

The firm chooses the bonus system (α, x) to maximize expected profit per team, subject to individual rationality and incentive compatibility constraints for each occupation embodied in the Nash effort functions in Theorem T2. Normalizing the price of output to one, expected team profit from the bonus is

$$E[\pi] = E[\text{revenue}] - E[\text{cost}] = \int_0^\infty [\theta\lambda_a(\theta) - w(\theta)]f(\theta)d\theta - \beta_a - 2\beta_b.$$

With U there are no wealth effects in the choice of effort, so the firm can set β_i to solve the individual rationality constraints with equality:

$$\beta_i = \frac{1}{r}\, ln(F(\theta^\star) + H_i(\alpha, x)) - \frac{ln(-\bar{u}_i)}{r}, \tag{5}$$

where

$$H_i(\alpha, x) = \int_{\theta^\star}^\infty \exp\Big\{-r\Big(\frac{\alpha}{3}(\theta\lambda_a(\theta) - x) - \frac{k_i}{2}\lambda_i(\theta)^2\Big)\Big\}f(\theta)d\theta, \quad i \in \{a, b\}$$

is the component of occupation i's base wage that compensates for utility generated when the team provides effort, that is when $\theta > \theta^\star$.

Corollary C2.1. *The optimal linear bonus solves*

$$\begin{aligned}\max_{\alpha, x} \quad & (1-\alpha)\int_{\theta^\star}^\infty \theta\lambda_a(\theta)f(\theta)d\theta \\ & + \alpha x[1 - F(\theta^\star)] - \frac{1}{r}\Big[\ln\Big(F(\theta^\star) + H_a(\alpha, x)\Big) \\ & + 2\ln\Big(F(\theta^\star) + H_b(\alpha, x)\Big)\Big]\end{aligned} \tag{6}$$

[3] From a computational standpoint, exponential utility is perhaps the only feasible functional form when solving the principal's problem. Margiotta and Miller (1993) also maintain this assumption, and most of the simulations of the Grossman and Hart (1983) model performed by Haubrich (1994) use exponential utility.

where $\lambda_a(\theta)$ and $\theta^\star$ correspond to the appropriate case of k_a and k_b in Theorem T2.

Corollary C2.2. *Given A1 and the production function (1), expected output under the linear bonus (α, x) expressed as a percentage of output under the optimal full information contract in T1, equals*

$$1 - t_1 t_2 t_3, \tag{7}$$

where

$$t_1 = \alpha$$
$$t_2 = \begin{cases} \frac{(2k_a+k_b)}{6k_a} & \text{if } k_a \leq k_b < 2k_a \\ \frac{(2k_a+k_b)}{3k_b} & \text{if } 2k_a < k_b. \end{cases}$$
$$t_3 = 1 - \Phi\left(\frac{ln\theta^\star - \mu}{\sigma} - 2\sigma\right).$$

Each of the three components in equation (7) has an economic interpretation as a source of inefficiency under incomplete information. First, output under the linear bonus is scaled by the bonus rate α because the team sets effort in response to its share of output rather than to total output.

The second term, t_2, captures the effects of free riding within teams. That is, $1 - t_2$ is the proportion of output lost due to free riding. At any value of θ, each team member chooses effort conditional on θ while ignoring the effect it has on the productivity of the other team members. The resulting loss in output relative to efficient full information output is constant across values of θ, however, the inefficiency is larger when $4k_a \leq k_b$, the case of Theorem T2 when mucker effort constrains team effort. Since muckers have higher costs of effort than miners, free riding is larger when their effort is the binding constraint on team effort. It is straight forward to show that t_2 lies in the interval $[\frac{1}{2}, \frac{2}{3}]$ if $k_a \leq k_b < 2k_a$, and it lies in $[\frac{1}{3}, \frac{1}{2}]$ if $4k_a < k_b$.

The third term in (7) captures the effects of team shirking. That is, $1 - t_3$ is the proportion of output lost due to teams completely shirking and providing zero effort. Under the linear bonus, teams with $\theta < \theta^*$ produce nothing and fail to meet the production standard x, while under the full information optimal contract they produce $1 - t_3 = \Phi\Big(\ln \theta^*/\sigma - 2\sigma\Big)$.

Because Corollary C2.2 is based on the production function (1), it only compares output that is sensitive to incentives. In other words, if output actually takes the form (3), the expression (7) only compares the values of $E[\lambda\theta_j]$. The conditions under which (7) over-estimates the expected productivity lead to:

Corollary C2.3. *Given the production function (3) and assumptions A1 and A2, (7) over estimates the expected output lost under the linear bonus relative to the full information output if:*

$$E[\nu] > \tau Var(d)$$

where

$$\tau = \frac{\frac{2\alpha}{(2k_a+k_b)} t_2 \left[1 - \Phi\left(\frac{\ln\theta^* - \mu}{\sigma} - 2\sigma\right)\right] e^{2\mu+2\sigma^2}}{1 - \alpha t_2 \left[1 - \Phi\left(\frac{\ln\theta^* - \mu}{\sigma} - 2\sigma\right)\right]}.$$

The condition C2.3 holds when the average additive shock dominates the variation in the multiplicative effect d, including the case $Var(d) = 0$.

4. Identification

In this section we discuss estimating the agency model using payroll data generated by the model. We assume the data to consist of bonuses received by a random sample of workers paid under the linear bonus. The agency model contains five parameters of primary interest: $(k_a, k_b, r, \mu, \sigma)$. These parameters determine worker effort levels and the firm's choice of the bonus parameters (α, x). The reservation utilities u_a and u_b are of secondary interest because they determine only the base wages β_a and β_b. First we consider the response of workers to the bonus parameters which they take as given. To the econometrician, (α, x) are unknown (and unrestricted) parameters to be estimated. We call this the unrestricted model. Next we add the restriction that through the maximization of the profit function in Lemma 2 (α, x) are implicit functions of the structural parameters and call this the restricted or structural model. Finally, we consider how data from several pay periods aids in identifying the structural model.

Lemma 3. *Define*

$$\psi = \begin{cases} \frac{9k_a}{\alpha^2} & \textit{if } k_a \leq k_b < 2k_a \\ \frac{9k_b}{2\alpha^2} & \textit{if } k_b \geq 2k_a. \end{cases}$$

and

$$\eta = \begin{cases} 1 & \textit{if } k_a < k_b \leq 2k_a \\ \frac{(k_b - k_a)}{k_a} & \textit{if } 2k_a < k_b \leq 4k_a \\ 3 & \textit{if } 4k_a < k_b. \end{cases}$$

Then, from assumption A1 and the Nash equilibrium effort functions in Theorem T2, the distribution of realized bonuses satisfies the following three properties:

(i) Positive bonuses are bounded away from zero with a lower bound $w(\theta^*) = \frac{\alpha x}{3\eta}$.

(ii) No bonus is received with probability $\Phi\Big[(1/\sigma)\big(1/2 \ln[\psi(w(\theta^*) + \alpha x/3)] - \mu\big)\Big]$.

(iii) For $w > w(\theta^*)$ *the density and cumulative distribution functions equal:*

$$f_w(w) = (2\sqrt{2\pi}\sigma(w + \alpha x/3))^{-1} \exp\Big\{-\frac{1}{8\sigma^2}[\ln(\psi(w + \frac{\alpha x}{3})) - 2\mu]\Big\}$$

$$F_w(w) = \Phi\Big[(1/\sigma)\big(1/2 \ln[\psi(w + \alpha x/3)] - \mu\big)\Big].$$

The parameter η depends on which occupation is the binding constraint on team effort. The value of η is monotonically related to t_2, the free-riding component t_2 of lost output (7). Therefore, η measures the free-riding within teams induced by differences in marginal costs of effort. The bounds on η arise from the assumptions of Leontief technology and quadratic costs of effort. In particular, the Nash equilibrium effort functions are unaffected by the relative values of k_a and k_b when $k_a < k_b \leq 2k_a$ and $4k_a < k_b$. In the density function the parameter ψ scales positive bonuses. Its value determines how effective the bonus rate is in providing team incentives.

All three statements in Lemma 3 follow from inserting the Nash equilibrium effort functions into the bonus equation (2). The economic reason for the first statement is that no team finds it worthwhile to earn small bonuses. Instead, the team sets effort to zero for values of $\theta < \theta^*$. If not for the lower bound on positive bonuses, the distribution of log bonuses would essentially form a censored regression similar to a Tobit model. The structural parameters determine the mean and variance of the disturbance term. Unlike an ordinary Tobit, however, the support of positive bonus also depends upon the model's parameters. Identifying the parameters in this circumstance is a nonstandard problem since the maximum likelihood estimator for the boundary, $w(\theta^*)$, is the smallest positive bonus in the data. Flinn and Heckman (1982), Christensen and Kiefer (1991), and Donald and Paarsch (1993) discuss the properties of boundary estimators.

Theorem T3. *(i) The following four parameters can be identified from the bonus distribution defined in Lemma 3: the standard deviation of production shocks σ, the incentive parameter ψ normalized by $e^{2\mu}$, the minimum bonus $w(\theta^*)$, and the free-riding term η. Using L3.(i), the value αx can be recovered from $w(\theta^*)$ and η, but α and x are not separately identified. (ii) The cooperative solution to team effort is observationally equivalent to the non-cooperative solution when $\eta = 1$.*

Theorem T3.(i) states that the bonus parameters are not separately identified directly from the distribution of bonuses. At best, only their product, αx, is identified. Furthermore, point estimates of k_a and k_b are also not available based solely on the response of workers to the linear bonus. Using Lemma 3, we can determine intervals in which k_a and k_b must lie to be consistent with the estimate η. Combined with Corollary C2, Theorem T3 also shows that the bonus distribution contains limited information about the linear bonus's efficiency relative to the potential full information case. The ratio of expected output under the bonus to full information output is given in (7), and it contains the three terms t_1, t_2, and t_3. The first component, $t_1 = \alpha$, lies between 0 and 1 but is unidentified without imposing profit maximization. The next, t_2, lies in the interval $[\frac{1}{3}, \frac{2}{3}]$ and is also unidentified because it is a function of k_a and k_b. The estimate of η can be used to tighten the bounds on t_2. Only t_3 can be estimated from the distribution of bonuses without maximizing the firm's problem, because t_3 can be calculated from the values listed in T3.

T3.(ii) states that it is possible to test for effective cooperation among team members. In particular, if the restriction $\eta = 1$ can be rejected, then the data provide evidence that the free-riding was not completely eliminated by some implicit aspect of compensation.

By adding the restriction that the firm chooses (α, x) to maximize profits, these parameters become implicit functions (without closed forms) of the structural parameters. From Theorem T3, only four parameters are identified from the bonus distribution defined theoretically by five structural parameters. One normalization that in principle identifies the structural model is to fix the mean of log production shocks μ for one pay period. Since μ does not enter (7), output under the linear bonus relative to full information does not depend on μ, making it a natural parameter to fix during estimation.[4]

Either k_a or k_b enters the firm's objective function only through $H_a(\alpha, x)$ or $H_b(\alpha, x)$ because on the margin team effort is determined by the occupation with the higher relative cost of effort. The other occupation's preferences only affect the level of pay required to compensate for effort. If the value of the cor-

[4] If (α, x) were known, numerical solutions to the firm's problem would still be useful for analyzing the data. Numerical solutions would be required to determine whether (or to impose the condition that) (α, x) maximized profit.

responding $H_i(\alpha, x)$ is insensitive to structural parameters near their estimated values, then k_i has a small effect on the firm's choice of (α, x). Therefore, one of the cost of effort parameters may not be well identified in a given sample.

By using data from several pay periods during which the bonus parameters were changed, we can identify relative movement in mean production shocks. That is, we assume that the preference parameters, (r, k_a, k_b), are constant over time periods and the firm changed the bonus parameters (α, x) in response to changes in the technology parameters (σ, μ) over time as tunnels are extended. By re-solving the firm's problem in each period, a separate value of σ can be estimated for each period and a separate value of μ can be estimated for all but one period. Three parameters vary across periods in the unrestricted model: ψ, $w(\theta^*)$, and σ. Since it depends only upon preferences, η is constant across periods. Each additional period of data therefore adds one degree of freedom in the restricted model that imposes profit-maximizing behavior relative to the unrestricted model based on team behavior alone.

5. Data and Results

Data

We have entered the payroll records of the Britannia mine for the years 1927 and 1928.[5] For each pay period we observe the number of shifts each employee worked, the job he performed (miner or mucker), and the bonus he received. While there were two pay periods per month, bonus rates were changed at most once per month. We therefore combine the data into monthly periods.

Table 1 summarizes the data by month. The ratio of mucker shifts to miner shifts is generally between 1.5 and 2. The proportion of workers receiving

[5] The records are located in the Special Collections section of the University of British Columbia Library. We transferred the records onto microfiche and from the microfiche coded the data into machine-readable form.

a bonus in each month fluctuates between 0.4 and 0.7 with no obvious trend across periods. Workers were paid bonuses according to the number of shifts they worked during the pay period, so the Table reports bonuses per shift. When estimating the model we restrict the sample to workers who worked 25 or more shifts in the month. Base wages per shift were $4.25 for miners and $4.00 for muckers, positive bonuses were on average 4 or 5 percent of base wages. The maximum positive bonus per shift fluctuates a fair amount across periods with outliers in months 7 and 16. On average it is $0.99, or over 23 percent of the base wage. The minimum positive bonus per shift is less than $.02 with outliers in months 1 and 20.

In each period roughly 6 percent of workers received a total bonus equaling $0.50. While the model predicts that positive bonuses are bounded away from zero, it seems unlikely that the theoretical bound was exactly $0.50. Furthermore, in periods where the smallest positive bonus is below $0.50, bonuses of $0.50 still appear regularly. We have found no explanation for this in company records. It appears that either the firm guaranteed a minimum bonus of $0.50 for certain jobs, or it may be that the firm usually rounded smaller bonuses up to $0.50. Our formation of the likelihood function proceeds from the latter explanation.

Likelihood Function

Recall that the distribution of bonuses identifies the free-riding term η and three time-varying parameters: the standard deviation of production shocks σ, the minimum observed bonus $w(\theta^*)$, and the effort coefficient ψ normalized by $e^{2\mu}$. If the firm did not round small bonuses, it would be possible to estimate $w(\theta^*)$ consistently using the the smallest observed positive bonus per shift in each period, denoted w_{min}. However, w_{min} is not a consistent estimate of $w(\theta^*)$ when small bonuses are rounded. With rounding, the probability of receiving a positive bonus is unchanged, but all bonuses between $nw(\theta^*)$ and 0.50 are rounded to 0.50, where n equals the number of shifts worked by an individual during the pay period. The probability that a bonus equals 0.50 equals $F_w(\theta_n^{\star\star}) - F_w(\theta^\star)$, where $\theta_n^{\star\star} = \sqrt{\psi(0.50/n + \alpha x/3)}$. The log-likelihood function for N workers with $(n_1, n_2, \ldots, n_N)$ shifts worked and per-shift bonuses $(w_1, w_2, \ldots, w_N)$ during one pay period equals

$$L(k_a, k_b, r, \mu, \sigma) = \sum_{\{i: w_i = 0\}} \ln(F_w(\theta^\star)) + \sum_{\{i: 0 < n_i w_i \leq 0.50\}} \ln(F_w(\theta_{n_i}^{\star\star}) - F_w(\theta^\star)) + \sum_{\{i: n_i w_i > 0.50\}} \ln(f_w(w_i)). \quad (8)$$

Whether or not rounding occurred, estimates based on maximizing (8) are consistent. If rounding did not occur estimates based on (8) are not efficient. We found parameter estimates based on estimating $w(\theta^\star)$ with w_{min} to be very sensitive to changes in w_{min} across periods so we report estimates based on (8).

To compute L under the restricted model, requires solving (6) numerically for the optimal values of (α, x) given current estimates of $(k_a, k_b, r, \mu, \sigma)$.[6] To reduce the number of estimated parameters, we let the technology parameters, (μ, σ), change at most every two months. We tested this restriction by estimating the non-structural model over all 24 periods, with η free, and allowing σ, μ and ψ to change in each period. We then estimated the model under the restriction that σ, μ and ψ change every two periods. The likelihood-ratio test statistic equals 58.4 with 36 degrees of freedom, which is not significant at the 1% level.

Unrestricted Estimates

Table 2A presents the estimates for the unrestricted model, allowing η to be estimated freely. The estimate of η of 10.49 lies beyond the region defined

[6] The profit function was maximized using Newton's method. The convergence criterion was very tight since the results must be passed on to the algorithm maximizing L. In particular, all elements of the gradient vector for π had to have absolute values below 1.0e-7. Starting values for L were found using the simplex method, as described in Press *et al.* (1987), then Newton's method was used to achieve convergence. All numerical work was performed in *Gauss*.

by the model. We also estimated the unrestricted (i.e. without profit maximization) model setting $\eta = 3$. These results are presented in Table 2B. Table 2B also reports Chi-squared goodness of fit tests of the unrestricted estimates. The details of the test statistic are given in the Appendix. The distributional assumption appears satisfactory. Only in period 4 (the period in which there are unusually large bonuses) is the fit rejected at the .01 significance level.

The estimates of output lost by pure shirking under the bonus system, $1-t_3$ in (7), are presented in the fourth row of Table 2B. In all periods less than 10 percent of potential full information output was lost due to workers shirking. The percentage of teams shirking is much higher than 10 percent, however, teams earning no bonus received poor draws of θ so their forgone output is well below average productivity. The estimated proportion of output lost due to free riding on other team members, $1-t_2$ in (7), lies in the range $[\frac{1}{2}, \frac{2}{3}]$. Free riding within teams accounts for more lost output than whole teams shirking.

The relatively large base wages paid to workers suggests that the output produced in the absence of any incentives was quite large. Corollary C2A suggests that Table 2B overestimates the percentage of expected output lost under the bonus system. The numbers are merely suggestive, since we are only estimating one part of (7), and Corollary C2A pertains to the whole expression. Measuring the complete expression requires separate identification of α and x, which is accomplished by imposing profit maximization on the bonus parameters.

Structural Estimates

Since η is pushed past its theoretical bound of 3, we estimate the structural model for case T2.(iii), $4k_a < k_b$. Team effort is determined by the muckers so k_a only enters the profit maximization problem through the expected utility constraint for the miners. We found that k_a was poorly identified in the data, so we estimated the model for two extreme cases: (A) $k_a = 0.24k_b$ and (B) $k_a = 0.01k_b$. The results are presented in Table 3. The results are not sensitive to the normalization.

The estimated values of the standard deviation σ are quite close to the values reported in Table 2. The values of μ are generally not significantly different from zero, indicating little movement in the average production shock across periods. Most of the variation in the pay system across periods is through the production standard x. The value of α varies between 0.55 and 0.62, while the value of x varies between 11.01 and 20.29. The risk aversion parameter was estimated on a monthly basis and then converted to a per shift value by dividing by 25, the number of shifts worked in a month.

The goodness of fit test was also performed on the structural model. Since certain parameters are constant across periods, we report the sum of the statistics for each period. There are 71 degrees of freedom–the number of cells in each period minus the number of estimated parameters minus one. The results are significant at the .01 significant level (row 5 of Tables 3A and 3B). The rejection is caused almost entirely by period 4, the same period for which the unrestricted model fit poorly. The test statistic excluding period 4 equals 84.0 in both cases which with 65 degrees of freedom has a p-value of 0.06.

The structural model restricts the cost of effort parameters and the risk aversion parameter to be constant across periods; changes in the bonus distribution are attributed to changes in the profit-maximizing choice of the (α, x) due to varying conditions in the mine. These restrictions are tested by performing a likelihood ratio test. The total likelihood in the unrestricted model is 7310.44. The likelihood for the structural model is 7329.21. The model is estimated over 12 periods, implying 11 degrees of freedom (36 parameters in the unrestricted model, 25 parameters in the structural model). The likelihood ratio of 38 rejects the restrictions placed on the unrestricted model. When period 4 is ignored, the ratio falls to 26 with 10 degrees of freedom but the hypothesis can still be rejected (p-value equals 0.003). The problem appears not to originate with the assumptions of exponential utility and profit maximization required to solve the firm's problem, which still fit the data reasonably well. Rather, the major problem in applying the model to the data is that the unrestricted estimate of η lies above its theoretical range based upon the production function (2) and the quadratic cost of effort. Other convex cost functions do not alter the basic form of the effort function illustrated in Figure 1, but they generate implicit solutions for either θ^* or the effort functions themselves, creating a third level of numerical solutions required to estimate the model (beyond maximizing the profit and likelihood functions). Alternative cost functions were therefore deemed impractical, and the quadratic cost function should be considered a feasible approximation.

The large value of η leads us to reject the hypothesis that the explicit linear bonus was supported by implicit arrangements that eliminated the inefficiency inherent in the form of compensation. Recall from Theorem T3.(ii) that cooperation within teams is empirically equivalent to the non-cooperative case with $\eta = 1$, which is strongly rejected by the data. This result does not rest upon any assumptions about what mechanism was used to enforce cooperation, but rather tests for the presence of cooperation in the distribution of bonuses.

6. Implications

Output Relative to Full Information

Using the parameter estimates in Table 3, we compare the expected output and profits under the linear bonus scheme with alternatives that the firm may have used. Since the structural parameters are estimated conditional on the normalization of k_a, we present two measures of the profits and output available under alternative schemes. The extent to which the normalization affects the estimates depends upon the manner in which k_a enters these expressions.

With estimates of the structural parameters we can calculate all terms in (7), subject to the normalization on k_a. We average across the twelve two-month periods and present the results in Table 4. We estimate the output loss under the linear bonus to be in the range 73 to 81 percent. The major source of output lost was free riding within teams (column 3). This is consistent with the results from the unrestricted model. We estimate the loss of teams with low θ's shirking to be small – on average only 4 percent of full information output is lost in this way, even though the percentage of teams not earning a bonus is on average 35 percent. The pattern in the data that accounts for this is the large variation in positive bonuses. Truncating the lowest 35% of output values is much less costly than losing 35% of average output when production has high variance.

We have already argued that these estimates of productivity lost are likely to be overestimates of the true loss of productivity since we so far we have not taken account of productivity from effort that can be enforced by the level of monitoring used in the mine. To get an idea of how much production was independent of incentives, we draw on some additional information contained in the payroll records. In the late 1920s, Britannia linked the base wage of its workers to the price of copper, which was then fluctuating much more than in the two years our data covers. We combine the size of the copper bonus with three additional assumptions:

Assumption A3.

(i) $Var[d] = 0$;

(ii) $-ln(-\bar{u}_i) = r(PE[\nu]/3 + s_i)$, *where* P *is the price per unit of copper.*

(iii) $s_a > 0, s_b < 0,$ *and* $s_a + 2s_b = 0$.

Assumption A3.(i) strengthens Assumption A2 by forcing all the variation in average productivity across areas of the mine to be additive. A3.(ii) specifies that certainty equivalent income outside the mine is a function of the price of copper and occupation. That is, if $\bar{w}_i$ is occupation i's certainty equivalent income, then

$$\bar{u}_i = -e^{-r\bar{w}_i}.$$

Taking logs gives, $-ln(-\bar{u}_i) = r\bar{w}_i$. We then let $\bar{w}_i = PE[\nu]/3 + s_i$. In effect A3.(ii) assumes that worker productivity at Britannia was typical of the industry. Assumption A3.(iii) allows reservation utility to vary across occupations.

A3.(ii) and (5) imply $\frac{\partial \beta}{\partial P} = E[\nu]/3$ or $E[\nu] = 3\frac{\partial \beta}{\partial P}$. With fluctuating copper prices, the copper bonus adjusts compensation to solve the individual rationality constraints. Under the copper bonus, every \$0.01 increase in the price of copper per ton increased base wages by \$0.25 per-shift. We therefore approximate $\frac{\partial \beta}{\partial P}$ by $\frac{\Delta \beta}{\Delta P} = \frac{\$.25}{\$.01} = 25$ tons per shift, and leads us to estimate $E[\nu] = 75$ tons per shift.[7] The percentage loss in expected output now equals

$$\frac{1 - t_1 t_2 t_3}{1 + 75(k_a + k_b/2)e^{-2(\mu+\sigma^2)}} \tag{9}$$

and decreases to between 58 and 69 percent (column 5 of Table 4). After correcting for productivity that is insensitive to incentives, the estimates of productivity lost relative to full information remain substantial.

Comparisons with Other Incentive Contracts

Comparing output under full information and the linear bonus is only a partial comparison because it does not include the wage costs associated with producing the output. Furthermore, the costs of incomplete information would have increased or decreased if the firm had paid workers using a contract other than the linear bonus. We now compare profits per-team under several different contracts, listed in descending order of expected profits before taking into account the cost of implementing the contract.

---

[7] Production levels recorded in the company reports suggest that this estimate is reasonable. For example, in June 1925 the average tons broken per miner shift was 20.5.

1. Full Information Optimal Contract: Profits under the contract described in Theorem T1, averaged over all pay periods at the estimated parameter values.

2. Incomplete Information Optimal Contract: Profits computed using a discrete approximation to the production shock θ and a flexible-form contract defined by 240 parameters. The parameter values for period 12 were used to form the profit function. The details are described in the Appendix.

3. Incomplete Information Two-Rate Linear Bonus: Profits computed when separate bonus rates α_a and α_b are paid to each occupation. The parameters are chosen to maximize

$$E[\pi] = (1 - \alpha_a - 2\alpha_b) \int_{\theta^*}^{\infty} \theta\lambda_a(\theta)f(\theta)d\theta - \frac{1}{r}[H_a(3\alpha_a, x) + 2H_b(3\alpha_b, x)] \qquad (10)$$

for all 12 periods where $\theta^\star$ is determined by the solution by the efficient Nash equilibrium within the team.

4. Incomplete Information Linear Bonus: Profits computed for the parameters estimates and the contract actually used by the firm.

5. Incomplete Information Simple Piece Rate: Profits computed under a simple piece rate α that maximizes

$$E[\pi] = (1 - \alpha) \int_{0}^{\infty} \theta\lambda_a(\theta)f(\theta)d\theta - \frac{1}{r}[H_a(\alpha, 0) + 2H_b(\alpha, 0)]. \qquad (11)$$

As with the two-rate system, (11) was maximized numerically for each pay period.

In terms of implementation costs as we defined them in section 2, the incomplete information optimal contract is the most costly. Next in order of implementation costs is the two-bonus rate system, then the linear bonus, and finally the simple piece rate and full information optimal contract, which is also a piece rate. To implement the full information contract would presumably require other costly measures including hiring more foremen, collecting more data about local conditions, and so forth. These fall under the costs of having incomplete information on worker action rather than the costs of implementing the contract given the information available. Before comparing profits under these contracts, we state and discuss two results concerning their performance.

Theorem T4. *Under assumptions A1–A3, estimates of expected profits do not depend upon $E[\nu]$, average team productivity without incentives, under full information or any contract that nests the linear bonus system.*

Assumption A4. $\nu \sim N(\mu_\nu, \sigma_\nu^2)$.

Theorem T5. *Under assumptions A1–A4, estimated profits under a piece rate are overestimated by the factor $\frac{1}{6} r\alpha^2\sigma_\nu^2$.*

Theorem T4 says that the correction for the level of team productivity required when comparing output levels is not necessary when comparing profits as long as reservation utilities reflect average productivity and the payment scheme is flexible enough to control for variation in productivity across areas in the mine. In the case of a simple piece rate, the mine cannot use x_j to cancel out fixed productivity differences ν_j across areas of the mine. The productivity differences do not cancel out of expected profits (11). A simple piece rate creates a lottery across areas of the mine, and the expression in T5 is the worker's risk premium associated with this lottery. Without knowing how variable the observable component of productivity was across areas, expected profits under the simple piece rate overstate actual profits.

Table 5 summarizes the performance of the five different contracts. The percentage of full information profits lost under the linear bonus is estimated to be between 72 and 81 percent on average across periods. We estimate a large inefficiency in the incomplete information outcomes even though only about 5% of compensation was in the form of incentive pay. (Standard errors on the estimates of expected profit computed based on the delta method were found to be quite small relative to the magnitude of the estimates.) A difference between casual and structural estimates of the cost of moral hazard has been noted elsewhere, in particular for executives of U. S. corporations. Using the sensitivity of executive pay to shareholder wealth as a measure of the importance of incentives, Jensen and Murphy (1990) find a sensitivity of less than 1 percent. Using estimates of a dynamic agency model, Margiotta and Miller (1993) estimate the effect of moral hazard to be of the same order as total assets controlled by the firm. (See also Haubrich 1994).

The key to both ours results and those of Miller and Margiotta is risk aversion. With risk averse agents, the cost to the principal of incomplete information is not proportional to the amount of variation in pay generated by the optimal incentive scheme. To obtain a measure of the economic impact of incomplete information requires an estimate of risk aversion as well as other aspects of technology and preferences.

We divide the loss in profits associated with the linear bonus into a part attributed to incomplete information and a part attributed to implementation costs using the profits under the approximate optimal contract with incomplete information (row 2 of Table 5). Profits under this contract are about 50% of that with complete information, accounting for two-thirds of the inefficiency of the linear bonus. The remaining one-third is due to the inefficiency in producing incentives under the linear bonus compared to optimal incentive contract.

Now we consider marginal changes in the form of the incentive contract away from the linear bonus, in particular the two-rate and simple piece rate schemes that maximize (10) and (11). Net profits under the two-rate system increase between 16 and 40 percent compared to the single-rate system actually used. Paying miners and muckers different piece rates would have increased efficiency substantially. Profits are smaller under the simple piece rate than under the linear bonus system, since the bonus system nests the piece rate. The average increase in profits under the linear bonus is small. On average, profits increased by a minimum of about 2.3 percent with a production standard. The fact that the estimated difference in profits (ignoring productivity differences across areas of the mine) is small suggests that the main benefit to the firm in introducing the production standard was its ability to cancel out observable productivity differences.

Extensions and Limitations

Our estimates are based on the assumption that workers know the conditions in the area they work before choosing effort. More realistically, workers only receive a signal of productivity when choosing effort. The opposite extreme would be the case of symmetric incomplete information: neither workers nor firms know the value of θ. In this case, team effort is constant relative to the realization of θ. At any set of model parameters, expected profits are smaller when workers do not see and respond to θ than when they do, because there is less information about the production process. In this sense, our estimates of the cost of incomplete information are conservative. However, our estimates of the model's parameters are biased in some unknown way if workers do not observe θ and we assume they do.

To gauge the effect of the bias, we used the parameter estimates for period 12 to re-solve the model when workers also do not observe θ. We calculated the Nash equilibrium within teams, re-maximized the firm's profit function, and generated an artificial data set on bonuses based on 1000 draws of θ. Using this data, we then re-estimated the model under our original assumption that workers know θ. As expected, the parameter estimates were quite different than the estimates we started with, illustrating the bias in assuming the wrong amount of information in the mine. The computed values for expected profits, however, were 54% of the full information result under the (incorrect) assumption that workers know θ, a value similar in magnitude to the 75% figure from the estimates. This suggests that the effect of misspecifying the amount of information available to workers may not have a qualitative effect upon the estimated cost of incomplete information.[8]

The comparison of profits in Table 5 suggests that paying two rates would have cut inefficiency considerably. Britannia's Annual Reports indicate that prior to 1926 the firm experimented with separate piece rates for different members of the team. The reports made at the time indicate that the two-rate system was dropped in favor of the team bonus primarily to improve harmony within teams. While the reduced-form estimates reject the presence of cooperative behavior, we have assumed that teams operate at the Pareto efficient Nash equilibria defined in Theorem T2. Each team member, however, can move the team into another Nash equilibria by simply shirking for some values of θ greater than $\theta^\star$. In effect, this raises the value of $\theta^\star$ appearing in the firm's objective (6). So while we use a non-cooperative solution to the team's problem, we can approximate the notion of cooperation by considering the firm's interest in getting teams to select the efficient Nash equilibrium. Differentiating (6) with respect to $\theta^\star$

$$\frac{\partial E\pi}{\partial\theta^\star} = -f(\theta^\star)\Big((1-\alpha)\theta^\star\lambda_a(\theta^\star) + \alpha x + \frac{1}{r(F(\theta^*) + H_a(\alpha,\theta^*))} + \frac{1}{r(F(\theta^*) + H_b(\alpha,\theta^*))}\Big)$$

$$-\frac{1}{r}\left(\frac{\frac{\partial H_a(\alpha,\theta^*)}{\partial\theta^*}}{F(\theta^*) + H_a(\alpha,\theta^*)} + \frac{\frac{\partial H_b(\alpha,\theta^*)}{\partial\theta^*}}{F(\theta^*) + H_b(\alpha,\theta^*)}\right), \tag{12}$$

where

$$\frac{\partial H_i(\alpha,\theta^*)}{\partial\theta^*} = -\exp\Big\{-r\Big(\frac{\alpha}{3}(\theta^*\lambda_a(\theta^*) - x) - \frac{k_i}{2}\lambda_i^2(\theta^*)\Big)\Big\}\, f(\theta^*) < 0.$$

The second term of (12) is positive. It represents the amount the firm could adjust the base wage to offset change in the amount of shirking. As long as the effect of the first term outweighs the second, profits are maximized at the Nash equilibrium where shirking stops at $\theta = \theta^\star$. Indeed, this holds

[8] It would be preferable to re-estimate the model parameters under the assumption of symmetric incomplete information. This was deemed impractical. It requires three levels of numerical solutions (worker, firm, econometrician) since worker effort must maximize expected utility rather than state-contingent (θ-contingent) utility.

at each set of parameter estimates, because H_a and H_b are not responsive to θ^*. Contracts in which workers receive different bonuses may have caused envy that resulted in break downs in the coordination required to reach the efficient Nash equilibrium. Equation (12) demonstrates that the firm has a vested interest in avoiding these break downs. Workers ultimately receive their reservation utility of $\bar{u}_i$ and occupations were paid different base wages, so this argument requires that utility functions be augmented with some element of envy or peer pressure. At the least, our results suggest that applying models of endogenous cooperative behavior in teams (e.g. Kandel and Lazear 1992) to payroll data may be fruitful.

While we control for heterogeneity in tasks and skills across occupations, skills may differ within occupations as well. In particular, workers may develop skills on the job. Shearer (1994) has matched the payroll data from Britannia with personnel files that include when the worker joined the firm. Using the same framework, he estimates the return to tenure within the firm controlling for the incentives induced by the bonus system.

We specify the environment as a static principal-agent problem. An important issue in the dynamics of incentive pay is the ratchet effect (Gibbons 1987 and Kanemoto and MacLeod 1990). The ratchet effect arises when a firm uses a worker's past performance to determine the parameters of the compensation scheme, and workers recognize this feedback. We ignore the ratchet effect for two reasons. Workers at Britannia changed location within the mine, and as tunnels progress rock conditions evolve over time. Both these facts reduce the extent to which a worker's past performance affects his future compensation even if past performance is used to rate an area relative to other areas. Ickes and Samuelson (1987) argue that worker rotation mitigates the ratchet effect because it reduces the correlation between a worker's performance today and the compensation scheme he expects to face in the future.

Despite opening in late 1800s, Britannia did not begin experimenting with productivity-based pay until 1923. The system remained in effect until 1930. The agency model can help explain the timing of these changes in the compensation system. First, we estimate substantial variation in average conditions over the sample period. Conditions before our sample period may have made monitoring cheaper and random elements of production less important. Second, the firm abandoned the system when world copper prices and labor costs declined in the early 1930s. These trends may have made both direct monitoring of workers using cheaper foremen and termination contracts (Macleod and Malcomson 1989) more cost-effective than incentive contracts. When other costs such as implementation costs are considered, trends in forms of compensation may be better understood.

7. Conclusions

This paper has explored the empirical content of agency theory in a case-study of the Britannia copper mine which used a simple incentive scheme during the 1920s. Firms use simple incentive systems even though agency theory does not sanction them as optimal. We have explored a transactions cost explanation for this discrepancy between practice and theory. We estimate that up to one-third of the loss in profits in the mine was due to such costs limiting the shape of the pay contract. The remaining two-thirds of the loss is associated with incentives and free-riding within teams. We test and reject the possibility that implicit aspects of compensation enforced the outcome under the optimal contract that ignores the cost of implementing complicated contracts.

Our results demonstrate that payroll data are informative about agency theory, in the sense that an agency model serves as the data-generating process for an estimation procedure using payroll data. We also demonstrate the reverse: agency models are informative about payroll data, in the sense that casual estimates of the cost of incomplete information differ considerably from estimates arising from the model itself. The non-standard nature of the principal's objective function with risk averse agents requires numerical solutions, but in return a better understanding is gained of both the data and the theory.

Appendix

Proof of Theorem T1

We first note that maximizing (2) implies $\lambda_a = 2\lambda_b$. With risk averse workers and a risk neutral firm, the optimal wage contract provides complete insurance. Using this condition to invert U for occupation i

$$W_i(\theta) = \frac{k_i}{2}(\lambda_i(\theta))^2 + U_i^{-1}(\bar{u}_i). \tag{1.1}$$

Productive efficiency requires that the marginal product of team effort equal the *sum* of worker marginal effort costs

$$\theta = k_a\lambda_a + \frac{k_b}{2}\lambda_a,$$

or,

$$\lambda_a = \frac{2\theta}{(2k_a + k_b)} \qquad \lambda_b = \frac{\theta}{(2k_a + k_b)}.$$

Subsitituting these expressions back into (1.1)

$$W_a(\theta) = U_a^{-1}(\bar{u}_a) + \frac{k_a}{2k_a + k_b}y \qquad W_b = U_b^{-1} + \frac{k_b}{4(2k_a + k_b)}y. \quad \textbf{QED}$$

Proof of Corollary C1

In general, if $ln(\theta) \sim N(\mu, \sigma^2)$, and Φ is the standard cumulative normal distribution, then

$$\int_z^\infty \theta^2 f(\theta)d\theta = e^{2\mu+2\sigma^2}\left[1 - \Phi\left(\frac{ln(z) - \mu - 2\sigma^2}{\sigma}\right)\right].$$

To see this note that $\theta^2 f'(\theta) = \frac{1}{\sqrt{2\pi}}\frac{\theta}{\sigma}e^{-\frac{1}{2\sigma^2}(ln\theta-\mu)^2}$. Using the change of variables $y = ln\theta - \mu$, it is straightforward to show that

$$\theta^2 f'(\theta) = e^{2\mu+2\sigma^2}\phi(\frac{y - 2\sigma^2}{\sigma}),$$

where ϕ is the standard normal density function. See for example Olkin, Gleser and Derman (1980, p.300). It follows directly that

$$E[y] = \int_0^\infty \theta\lambda_a(\theta)f(\theta)d\theta = \int_0^\infty \frac{1}{k_a + \frac{1}{2}k_b}\theta^2 f(\theta)d(\theta) = \frac{1}{k_a + \frac{1}{2}k_b}e^{2\mu+2\sigma^2}.$$

Therefore

$$E[\pi] = E[y] - E[W_a] - 2E[W_b] = \frac{E[y]}{2} - U_a^{-1}(\bar{u}_a) - 2U_b^{-1}(\bar{u}_b). \quad \textbf{QED}$$

Proof of Theorem T2

(i) Utility of mucker i, conditional on both the miner's effort, λ_a, and mucker $j's$ effort, $\lambda_{b,j}$, is

$$U_{b,i}(\theta|\lambda_a, \lambda_{b,j}) = \frac{\alpha}{3}(\theta min\{\lambda_a, \lambda_{b,i} + \lambda_{b,j}\} - x) - \frac{k_b}{2}\lambda_{b,i}^2.$$

Mucker i equates the marginal cost of his effort to the marginal return at effort level

$$\lambda_{b,i} = \frac{\alpha\theta}{3k_b},$$

which is independent of mucker $j's$ actions. This defines the maximum level of effort that the mucker will supply for each value of θ.

Define $\theta_{b,i}^\star$ to be that value of θ at which mucker i is just indifferent between supplying effort and shirking. Then $\theta_{b,i}^\star$ solves

$$U_{b,i}(\theta_{b,i}^\star|\lambda_a, \lambda_{b,j}) = 0.$$

Since mucker effort enters the production function additively, there are a continuum of possible values $\theta_{b,i}^\star \in [\theta_{b,i,min}^\star, \theta_{b,i,max}^\star]$. Noting that

$$\lambda_{b,j} \in \left[0, \frac{\alpha\theta}{3k_b}\right],$$

$\theta_{b,i,min}^\star$ solves

$$\frac{\alpha}{3}\left(\theta min\{\lambda_a, \lambda_{b,i} + \frac{\alpha\theta}{3k_b}\} - x\right) - \frac{k_b}{2}(\lambda_{b,i})^2 = 0,$$

and $\theta^{\star}_{b,i,max}$ solves

$$\frac{\alpha}{3}\left(\theta min\{\lambda_a, \lambda_{b,i}\} - x\right) - \frac{k_b}{2}(\lambda_{b,i})^2 = 0.$$

It follows directly that for $\theta < \theta_{b,i,min}$, mucker i's dominant strategy is to shirk. Alternatively, for $\theta > \theta^{\star}_{b,i,max}$, mucker i's dominant strategy is to provide positive effort. For $\theta \in (\theta^{\star}_{b,i,min}, \theta^{\star}_{b,i,max})$, mucker i will either provide effort or shirk, depending on his belief over mucker j's actions.

Conditional on λ_a, we can characterize the symmetric best response function for the muckers

$$\lambda_b = \begin{cases} min\{\frac{\alpha\theta}{3k_b}, \frac{\lambda_a}{2}\} & \text{if } \theta > \theta^{\star}_b \text{ and } \lambda_a > x/\theta \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta^{\star}_b$ solves: $U_b(\lambda_b(\theta^{\star}_b)|\lambda_a) = 0$. Similarly, the best response function of the miner can be written

$$\lambda_a = \begin{cases} min\{\frac{\alpha\theta}{3k_a}, 2\lambda_b\} & \text{if } \theta > \theta^{\star}_a \text{ and } 2\lambda_b > x/\theta \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta^{\star}_a$ solves: $U_a(\lambda_a(\theta^{\star}_a)|\lambda_b) = 0$.

Solving for an intersection of best response functions will give a symmetric Nash equilibrium. There are a continuum of these equilibria, due to the continuum of possible values of $\theta^{\star}_b$. We now show that the Nash equilibrium with $\theta^{\star}_b = \theta^{\star}_{b,i,min}$ Pareto dominates all other symmetric Nash equilibria, Let

$$\Theta^{\star}_b = \{\theta | \theta \in [\theta^{\star}_{b,min}, \theta^{\star}_{b,max}]\}$$

Furthermore, let

$$\theta^{\star}_b \in \Theta^{\star}_b \quad \text{and} \quad \theta^{\star}_b - \epsilon \in \Theta^{\star}_b; \quad \epsilon > 0.$$

Finally, define

$$D(\theta^{\star}_b, \epsilon; \theta) = U_b(\lambda_b | \theta^{\star}_b - \epsilon; \theta) - U_b(\lambda_b | \theta^{\star}_b; \theta).$$

$D(\theta^{\star}_b, \epsilon; \theta)$ is the difference in mucker utility between the two Nash equilibria conditional on θ.

We divide states of nature into 3 possible cases and sign $D(\theta_b^\star, \epsilon; \theta)$ in each state.

$$D(\theta_b^\star, \epsilon; \theta) \quad is \quad \begin{cases} = 0 & \text{if } \theta > \theta_b^* \\ > 0 & \text{if } \theta_b^* - \epsilon < \theta \leq \theta_b^* \\ = 0 & \text{if } \theta_b^* - \epsilon > \theta \end{cases}.$$

The first case follows since the value of $\theta^\star$ does not affect the level of effort chosen for $\theta > \theta^\star$. The second case follows directly from the definition of $\theta_b^\star$, and in the third case $D(\theta_b^\star, \epsilon; \theta) = 0$ since shirking will be the optimal strategy for whenever $\theta < \theta_b^\star - \epsilon$.

Since this holds for all values of $\theta^* > \theta_{b,min}^\star$, it follows that the Nash equilibrium with $\theta_b^\star = \theta_{b,min}^\star$, dominates all other Nash equilibria for the muckers. Miners cannot be made worse off by reducing $\theta_b^\star$ since these reductions broaden the range of θ over which the miner can supply effort with positive utility. It therefore follows that the Nash equilibrium with $\theta_b^\star = \theta_{b,min}^\star$ Pareto dominates all other Nash equilibria.

To complete the proof we use the following Lemma.

Lemma. *Conditional on θ, the miner's equilibrium utility level is $\left\{ {greater \atop less} \right\}$ than the utility level of the muckers whenever $k_b \{ {\geq \atop <} \} 4k_a$.*

Proof: First note that in equilibrium miner effort will be twice mucker effort. Since costs are quadratic, it follows that

$$C_b(\lambda_b) = \frac{k_b}{2}(\lambda_b)^2 = \frac{k_b}{2}\left(\frac{\lambda_a}{2}\right)^2 = \frac{k_b}{4k_a} C_a(\lambda_a).$$

Therefore

$$k_b {\geq \atop <} 4k_a \iff C_b(\lambda_b) {\geq \atop <} C_a(\lambda_a) \iff U_b {\leq \atop >} U_a \textbf{ QED}.$$

Consider now case (i) of the Nash equilibrium. If $k_a \leq k_b < 2k_a$, then $\frac{\alpha\theta}{3k_a} < \frac{2\alpha\theta}{3k_b}$. Solving for $U_a(\theta^\star) = 0$ gives $\theta^\star = \sqrt{6k_a x/\alpha}$. $U_b(\theta^\star) > U_a(\theta^\star)$ follows from the Lemma.

Next consider case (ii). If $2k_a < k_b < 4k_a$, then $\frac{2\alpha\theta}{3k_b} < \frac{\alpha\theta}{3k_a}$. It is clear from the Lemma, that the binding minimum level of utility will be the miner's. Solving for $\theta^\star$ from the miner's indirect utility function gives

$$\theta^\star = \sqrt{\frac{3k_b^2 x}{2\alpha(k_b - k_a)}}.$$

Finally consider case (iii). If $k_b > 4k_a$, then $\frac{2\alpha\theta}{3k_b} < \frac{\alpha\theta}{3k_a}$. Solving for $U_b(\theta^\star) = 0$ gives $\theta^\star = \sqrt{2k_b x/\alpha}$. From the Lemma, $U_a(\theta^\star) > U_b(\theta^\star)$ **QED.**

(ii) If team members cooperate with each other, they choose $\lambda_{a,c}$ and $\lambda_{b,c}$ to maximize

$$w_a + 2w_b - c_a(\lambda_{a,c}) - 2c_b(\lambda_{b,c}).$$

As before, the Leontief production function function requires $\lambda_{a,c} = 2\lambda_{b,c}$. Substituting this expression and the form of payment into the objective function gives

$$\alpha(\lambda_{a,c}\theta - x) - \frac{k_a}{2}\lambda_{a,c}^2 - \frac{k_b}{4}\lambda_{a,c}^2.$$

Maximizing this with respect to $\lambda_{a,c}$ gives an interior solution $\lambda_{a,c} = \frac{2\alpha\theta}{2k_a+k_b}$. It is straightforward to solve for the value of θ for which net team compensation is zero, which results in $\theta_c^\star$ in the text. **QED**

Proof of Corollary C2.1

Normalizing the price of output to one, the expected profit per team from the bonus scheme is $E[\pi] = E[\text{revenue}] - E[\text{cost}]$

$$= \int_0^\infty [\theta\lambda_a(\theta) - 3w(\theta)]f(\theta)d\theta - \beta_a - 2\beta_b = \int_{\theta^\star}^\infty \theta\lambda_a(\theta) - \alpha(\theta\lambda_a(\theta) - x)f(\theta)d\theta - \beta_a - 2\beta_b.$$

Using (3), the individual rationality constraints can be written

$$\bar{u}_i = -e^{-r\beta_i}\Big[F(\theta^\star) + H_i(\alpha, x)\Big].$$

Solving for β_i and ignoring that part of β_i that depends on $\bar{u}_i$ gives the expression for expected profits. **QED**

Proof of Corollary C2.2

Substituting the Nash equilibrium effort functions from Theorem 3 into the equation for expected output and using the properties of the log normal distribution gives the following.

$$E[y] = \int_0^\infty \theta\lambda(\theta)f(\theta)d\theta = \begin{cases} \frac{\alpha}{3k_a}e^{2(\mu+\sigma^2)}\left[1-\Phi\left(\frac{ln\theta^*-\mu}{\sigma}-2\sigma\right)\right] & \text{if } k_a \leq k_b < 2k_a; \\ \frac{2\alpha}{3k_b}e^{2(\mu+\sigma^2)}\left[1-\Phi\left(\frac{ln\theta^*-\mu}{\sigma}-2\sigma\right)\right] & \text{if } 2k_a \leq k_b, \end{cases}.$$

Dividing this expression by full information expected output, derived in Corollary C1, gives $t_1 t_2 t_3$ as defined in the text.

Proof of Corollary C2.3

We present the proof for the case $k_a < k_b < 2k_a$; extension to the other cases are straightforward. We estimate the percent difference in expected output between full information and the bonus system by

$$\frac{\frac{1}{(k_a+k_b/2)} - \frac{\alpha}{3k_a}\left[1-\Phi\left(\frac{ln(\theta^*)-\mu}{\sigma}-2\sigma\right)\right]}{\frac{1}{(k_a+k_b/2)}}.$$

If instead the production function was as in (3), let y^b denote the expected output under a linear bonus

$$y^b = E[d]\frac{\alpha}{3k_a}e^{2\mu+2\sigma^2}\left[1-\Phi\left(\frac{ln(\theta^*)}{\sigma}-2\sigma\right)\right] + E[\nu].$$

Under full information, the firm does not use a piece rate to cancel out fixed differences d_j. The marginal benefit of effort to the firm is therefore $d_j\theta_j$, and the sum of the marginal costs to workers is $k_a\lambda_a + k_b/2\lambda_a$. Let y^f denote expected output

$$y^f = \frac{E[d^2]}{(k_a+\frac{k_b}{2})}e^{2\mu+2\sigma^2} + E[\nu].$$

The percentage difference in expected outputs is

$$\frac{\frac{E[d^2]}{(k_a+\frac{k_b}{2})} - E[d]\frac{\alpha}{3k_a}\left[1-\Phi\left(\frac{ln(\theta^*)-\mu}{\sigma}-2\sigma\right)\right]}{\frac{E[d^2]}{(k_a+\frac{k_b}{2})} + E[\nu]\,e^{-2(\mu+\sigma^2)}}.$$

Agebraic manipulation shows that we overestimate the percentage decrease if

$$E[\nu] > \frac{\frac{\alpha}{3k_a}\left[1-\Phi\left(\frac{ln(\theta^*)-\mu}{\sigma}-2\sigma\right)\right]e^{2(\mu+\sigma^2)}}{1-(k_a+\frac{k_b}{2})\frac{\alpha}{3k_a}\left[1-\Phi\left(\frac{ln(\theta^*)-\mu}{\sigma}-2\sigma\right)\right]}(E[d^2]) - E[d]).$$

Under Assumption A2, $E[d] = 1$, so that

$$E[d^2] - E[d] = E[d^2] - (E[d])^2 = Var(d). \quad \textbf{QED}$$

Proof of Lemma L3

Direct substitution of the Nash equilibrium effort functions into the bonus equation (2) gives

$$w(\theta) = \begin{cases} \frac{\theta^2}{\psi} - \frac{\alpha x}{3} & \text{if } \theta > \theta^\star \\ 0 & \text{otherwise} \end{cases}$$

L3.(i) follows from direct substitution of the expression for $\theta^\star$ into the bonus equation. L3.(ii) follows from the fact that

$$Pr(w = 0) = Pr(w < w(\theta^*)) = Pr\left(\theta < \sqrt{\psi(w(\theta^\star) + \frac{\alpha x}{3})}\right),$$

and the log-normality of θ. To derive the density function in L3.(iii) note that since $ln(\theta) \sim N(\mu, \sigma^2)$, $ln(\theta^2) \sim N(2\mu, 4\sigma^2)$ and

$$f_{\theta^2}(\theta^2) = \frac{1}{\theta^2}\frac{1}{\sqrt{8\pi\sigma^2}}exp\left\{\frac{-1}{8\sigma^2}(ln(\theta^2) - 2\mu)^2\right\}$$

Using the change of variables from θ^2 to w gives

$$f_w(w) = \frac{1}{2\sigma(w + \frac{\alpha x}{3})\sqrt{2\pi}}exp\left\{\frac{-1}{8\sigma^2}[ln(\psi(w + \frac{\alpha x}{3})) - 2\mu]^2\right\}.$$

The expression for the cumulative distribution function follows the same steps as in part two. **QED**

Proof Theorem T3

(i) The contribution to the likelihood of a limit observation is

$$\Phi\left(\frac{ln(\sqrt{\psi\frac{\alpha x}{3}(1/\eta + 1)}) - \mu}{\sigma}\right).$$

Similarly, the contribution to the likelihood of a non limit observation is

$$\frac{1}{\sqrt{2\pi}(w + \frac{\alpha x}{3})2\sigma}exp\left\{-\frac{1}{8\sigma^2}[ln(\psi(w_i + \frac{\alpha x}{3})) - 2\mu]^2\right\}.$$

As a function of structural parameters, the lower bound on the wage distribution is $w(\theta^*) = \frac{\alpha x}{3\eta}$. Using results in Donald and Paarsch (1993) on boundary estimators, consistency follows from the fact that $w(\theta^*)$ is monotonic and invertible in $\frac{\alpha x}{3}$, and that the inverted equation

$$\frac{\alpha x}{3} = \eta w_{min}$$

is a smooth function of η and w_{min}, where w_{min} equals the minimum observed positive bonus in the sample.

The parameters ψ and μ are not separately identified, because the contributions of both limit and non limit observations depend only on the ratio $\psi/e^{2\mu}$. For limit observations:

$$\Phi\left(\frac{ln(\{\psi/e^{2\mu}\}\frac{\alpha x}{3}(1/\eta+1))}{\sigma}\right),$$

and similarly for non limit observations:

$$\frac{1}{\sqrt{2\pi}2\sigma(w+\frac{\alpha x}{3})}exp\left\{-\frac{1}{8\sigma^2}[ln(\{\psi/e^{2\mu}\}(w_i+\frac{\alpha x}{3}))]^2\right\}.$$

Replacing $w(\theta^*)$ by its estimate w_{min} we can consider the concentrated likelihood function

$$l = \sum_{w_i=0} ln\left[\Phi\left(\frac{ln(\psi/e^{2\mu}w_{min}(1+\eta))}{2\sigma}\right)\right]$$

$$+\sum_{w_i>0} -ln(\sigma) - ln(w_i+\eta w_{min}) - \frac{1}{8\sigma^2}\left[ln(\psi/e^{2\mu}(w_i+\eta w_{min}))\right]^2.$$

The values $\psi/e^{2\mu}$, η and σ enter l independently and can be consistently estimated by maximizing l.

(ii) Define $\psi_c = \frac{3(2k_a+k_b)}{2\alpha^2}$. Positive wages are then of the form

$$w^+ = \frac{\theta^2}{\psi_c} - \frac{\alpha x}{3}.$$

The minimum observed wage will then be

$$w(\theta_c^*) = \frac{\alpha x}{3}.$$

Thus when workers cooperate with each other the unrestricted distribution of wages in identical to that for the Nash solution when $\eta = 1$. **QED**

The Chi-squared test Statistic

The Chi-squared test is based on the comparison between the predicted and actual proportion of workers receiving different values of bonuses. To conduct the test, the bonus distribution was partitioned on the basis of bonus received and the number of shifts worked. For each period the partitions are

$$W_1 = \{w | w = 0\} \quad W_2 = \{w | 0 < w \le \$.50\} \quad W_3 = \{w | \$.50 < w \le \$1.50\}$$
$$W_4 = \{w | \$1.50 < w \le \$2.50\} \quad W_5 = \{w | \$2.50 < w\}$$

and

$$S_1 = \{s | 25 \le s \le 27\} \quad S_2 = \{s | 27 < s\},$$

creating 10 cells each period. The test statistic is calculated for each period as

$$Q = \sum_{W_j} \sum_{i} \frac{\left(n_{S_i,W_j} - \hat{n}_{S_i,W_j}\right)^2}{\hat{n}_{S_i,W_j}}$$

where

n_{S_i,W_j} equals the observed number of workers with $w \in W_j$ and $shifts \in S_i$;

$\hat{n}_{S_i,W_j}$ equals the predicted number of workers with $w \in W_j$ and $shifts \in S_i$.

$$\begin{aligned}\hat{n}_{S_i,W_j} &= nPr(w \in W_j, s \in S_i) = n \sum_{s \in S_i} Pr(w \in W_j, s) \\ &= n \sum_{s \in S_i} Pr(w \in W_j | s) Pr(s) = \sum_{s \in S_i} Pr(w \in W_j | s) n_s,\end{aligned}$$

where n_s equals the number of workers working s shifts. In each period there are 10 cells but only 8 are estimated freely since the conditional probabilities must sum to one. Furthermore, there are three parameters estimated in each period resulting in 5 degrees of freedom for the test.

Approximation to the Optimal Incomplete Information Contract

The continuous distribution of production shocks is discretized into 80 points of the form:

$$\theta_j = \frac{F^{-1}((j-1)/80) + F^{-1}(j/80)}{2}$$

for $j = 1, 2, \ldots 80$. These are midpoints of intervals defined by percentiles of the true distribution. The probability of θ_j is set to a constant 1/80.

Next, we posit a flexible-form contract defined by 240 values $\{(y_l, \alpha_l^a, \alpha_l^b)\}_{l=1}^{80}$ and two base wages β_a and β_b such that the wage paid to workers in occupation i is a step function:

$$W_i(y) = \beta_i + \alpha_l^i y_l \text{ when } y_l \leq y < y_{l+1}$$

where $y_{81} = \infty$. The values of y_l are points at which wage payments jump to new values, and α_l^i are the shares of output paid to workers in occupation i at step l. Given that the base wages can be chosen to meet the individual rationality constraints, three of the contract's parameters are normalized: $y_1 = \alpha_1^a = \alpha_1^b = 0$. This contract uses the fact that with 80 values of θ there are at most 80 different values of output in equilibrium. As the number of discrete values is increased, the distribution of θ converges to the actual distribution and the step-function converges to a completely flexible wage contract.

In equilibrium, teams only compute the utility of producing values of output conditional on the draw θ_j. The team effort required to produce y_l given θ_j is $\lambda_{jl} = y_l/\theta_j$. The utility to miners in doing this is

$$U_{a,j,l} = \alpha_l^a y_l - \frac{k_a}{2}\lambda_{jl}^2.$$

Since muckers each process half the output their utility is

$$U_{b,j,l} = \alpha_l^b y_l - \frac{k_b}{2}\left(\frac{\lambda_{jl}}{2}\right)^2.$$

The efficient Nash equilibrium within the team is defined as the greatest value of y_l such that both occupations receive higher utility from producing y_l than any lower output:

$$l^*(j) = \max\{l: \text{ for } i = a, b \;\; U_{i,j,l} \geq U_{i,j,m}, \;\; m = 1, 2, \ldots, l\}.$$

This definition of equilibrium uses the symmetry of the muckers and the perfect complementary of effort across occupations, since each occupation's effort determines the maximum amount of output. In equilibrium, the probability of y_l being produced equals

$$p_l = \sum_{j=1}^{80} \frac{1}{80} I_{\{l=l^*(j)\}}$$

where $I_{\{A\}}$ is the indicator function for event A.

The firm's objective function can now be written:

$$\max_{\{(y_l, \alpha_l^a, \alpha_l^b)\}_{l=2}^{80}} \quad \sum_{l=1}^{80} p_l(1 - \alpha_a - 2\alpha_b) y_l - \frac{1}{r} \ln\left(\sum_{j=1}^{80} \frac{1}{80} \exp\{-rU_{a,j,l^*(j)}\}\right)$$

$$+2\ln\left(\sum_{j=1}^{80} \frac{1}{80} \exp\{-rU_{b,j,l^*(j)}\}\right).$$

This function was maximized using the NMSIMP algorithm described in Press et al. (1987). Experiments were done raising the number of points from 80 to several thousands. These experiments confirmed that expected profits under the full information contract converged in the number of points to the theoretical value given in Corollary C1. For 80 points, which was near the limit of computational possibilities for maximizing the firm's profit function, the difference in full information profits was less than 5%.

Proof of Theorem T4

(i) *Bonus System:*

The bonus system uses x_j to cancel out fixed differences ν_j, ie. $x_j = x + \nu_j$. Expected output is $E[y] = E[\theta\lambda_a(\theta)] + E[\nu]$ and expected profits per team are

$$E[\pi] = E[\theta\lambda_a(\theta)] + E[\nu] - E[w_a(\theta) + 2w_b(\theta)] - \beta_a - 2\beta_b,$$

Under assumptions A1–A3

$$\beta_i = \frac{ln(F(\theta^*) + H_i(\alpha, x))}{r} + \frac{E[\nu]}{3} + s_i \qquad \text{for } i \in a, b$$

and expected profits reduce to

$$E[\pi] = E[\theta\lambda_a(\theta)] - E[w_a(\theta) + 2w_b(\theta)] - \frac{ln(F(\theta^*) + H_a(\alpha, x))}{r} - 2\frac{ln(F(\theta^*) + H_a(\alpha, x))}{r}.$$

(ii) *Full Information*

Under full information θ, ν and λ are observable. The firm chooses effort, $\lambda_i(\theta, \nu)$ and wages, $w_i(\theta, \nu)$ to maximize expected profits subject to the workers expected utility constraint. Expected profits per team are

$$\int_\theta \int_\nu \theta\lambda(\theta, \nu) + \nu - w_a(\theta, \nu) - 2w_b(\theta, \nu) f(\theta) g(\nu) d\theta d\nu.$$

Expected utility for occupation i is

$$\int_\theta \int_\nu u(w_i(\theta, \nu) - c_i(\lambda(\theta, \nu)) f(\theta) g(\nu) d\theta d\nu.$$

As in Theorem T1 the optimal contract implies that $\lambda_a = 2\lambda_b$, and is characterized by the optimal risk sharing and efficiency conditions. Because the marginal benefit of worker effort is independent of fixed effects, ν does not affect the efficiency of effort condition. That is

$$\lambda_a = \frac{\theta}{k_a + k_b/2} \text{ and } \lambda_b = \frac{\lambda_a}{2}.$$

Optimal risk sharing implies full insurance for the worker, that is

$$U(w_i(\theta, \nu) - c_i(\lambda_i(\theta))) = \bar{u}_i \quad \text{or} \quad w_i(\theta, \nu) = c_i(\lambda_i(\theta)) + U^{-1}(\bar{u}_i).$$

Assumptions A1–A3 imply

$$w_a = \frac{2k_a\theta^2}{(2k_a + k_b)^2} + \frac{E[\nu]}{3} + s_a$$

$$w_b = \frac{k_b\theta^2}{2(2k_a + k_b)^2} + \frac{E[\nu]}{3} + s_b$$

and expected profits are independent of fixed effects ν

$$E[\pi] = \int_\theta \frac{\theta^2}{2k_a + k_b} f(\theta) d\theta. \quad \textbf{QED}$$

Proof of Theorem T5

Expected profits of the firm are

$$\int_\nu \int_\theta [\theta\lambda(\theta) + \nu](1-\alpha) f(\theta) g(\nu) d\theta d\nu - \beta_a - 2\beta_b,$$

where β_a and β_b solve the participation constraints of the miners and muckers respectively. To solve for β_a

$$\int_\nu \int_\theta -exp\Big\{-r[\beta_a + \frac{\alpha}{3}(\theta\lambda_a + \nu) - \frac{k_a}{2}\lambda_a^2]\Big\} f(\theta) g(\nu) d\theta d\nu = \bar{u}_a.$$

Using the independence of θ and ν gives

$$\beta_a = \frac{1}{r}\Big\{ ln\Big[\int_\nu e^{-r\frac{\alpha}{3}\nu} g(\nu) d\nu\Big] + ln\Big[\int_\theta e^{-r(\frac{\alpha}{3}\theta\lambda_a - \frac{k_a}{2}\lambda_a^2)} f(\theta) d\theta\Big] - ln(-\bar{u}_a)\Big\}$$

Given the normality of ν we use the result

$$E[e^{-k\nu}] = e^{-k\mu_\nu + \frac{1}{2}k^2\sigma_\nu^2}.$$

Furthermore, using

$$\int_\theta e^{-r(\frac{\alpha}{3}\theta\lambda_a - \frac{k_a}{2}\lambda_a^2)} f(\theta) d\theta = H_a(\alpha, 0)$$

$$\text{and } -ln(-\bar{u}_i) = r(E[\nu]/3 + s_a)$$

gives

$$\beta_a = \frac{ln[H_a(\alpha,0)]}{r} + \frac{1}{3}(1-\alpha)E[\nu] + \frac{r}{2}\frac{\alpha^2}{9}\sigma_\nu^2 + s_a.$$

The same procedure for β_b gives

$$\beta_b = \frac{ln[H_b(\alpha,0)]}{r} + \frac{1}{3}(1-\alpha)E[\nu] + \frac{r}{2}\frac{\alpha^2}{9}\sigma_\nu^2 + s_b.$$

Actual expected profits are

$$E[\pi] = \int_\theta \theta\lambda(\theta)(1-\alpha) f(\theta) d\theta - \frac{ln[H_a(\alpha,0)]}{r} - \frac{2ln[H_b(\alpha,0)]]}{r} - 3\frac{r}{2}\frac{\alpha^2}{9}\sigma_\nu^2.$$

but we estimate profits under the piece rate as

$$E[\pi] = \int_\theta \theta\lambda(\theta)(1-\alpha) f(\theta) d\theta - \frac{ln[H_a(\alpha,0)]}{r} - \frac{2ln[H_b(\alpha,0)]]}{r}.$$

Clearly this over–estimates profits under the piece rate by the amount $\frac{1}{6}r\alpha^2\sigma_\nu^2$.
QED

REFERENCES

Blinder, A. (ed.) (1990). *Paying for Productivity: A Look at the Evidence.* Washington: The Brookings Institution.

Britannia Mining and Smelting Company Limited. Records. Provincial Archives of British Columbia. Add. Mss. 1221. Victoria, British Columbia.

Brown, C. (1990). "Firms' Choice of Method of Pay," *Industrial and Labor Relations Review*, 43, February, 165s-182s.

Christensen, B. J. and Kiefer, N. M. (1991). "The Exact Likelihood Function for an Empirical Job Search Model," *Econometric Theory*, 7, 464-486.

Donald, S. and Paarsch, H. (1993). "Piecewise Pseudo-Maximum Likelihood Estimation in Empirical Models of Auctions," *International Economic Review*, 34, 125-148.

Ehrenberg, R. (ed.) (1990). *Industrial & Labor Relations Review, Special Issue: Do Compensation Policies Matter?*, 43.

Eckstein, Z. and Wolpin, K. I. (1990). "Estimating an Equilibrium Search Model using Panel Data on Individuals," *Econometrica*, 58, 4, 783-808.

Flinn, C. and Heckman, J. J. (1982). "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics* 18, 115-168.

Gibbons, R. (1987). "Piece-Rate Incentive Schemes," *Journal of Labor Economics,* 5, October, 413-29.

Green, J. and Stokey, N. (1983). "Comparison of Tournaments and Contracts," *Journal of Political Economy* 91, June, 349-64.

Grossman, S. J. and Hart, O. D. (1983). "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, January, 7-46.

Haubrick, J. G. (1994). "Risk Aversion, Performance Pay, and the Principal-Agent Problem," *Journal of Political Economy*, 102, 2, 258-276.

Holmstrom, B. (1979). "Moral Hazard and Observability," *Bell Journal of Economics*, 19, Winter, 74-91.

Holmstrom, B. (1982). "Moral Hazard in Teams," *Bell Journal of Economics*, 13, Spring, 324-40.

Holmstrom, B. and Milgrom, P. (1987). "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*55, March, 303-328.

Ickes, B. W., and Samuelson, L. (1987). "Job Transfers and Incentives in Complex Organizations: Thwarting the Ratchet Effect," *The Rand Journal of Economics*, 18, 275-286.

Jensen, M. C. and Murphy, K. J. (1990). "Performance Pay and Top-Management Incentives," *Journal of Political Economy*, 98, 225-264.

Kandel, E. and Lazear, E. P. (1992). "Peer Pressure and Partnerships," *Journal of Political Economy*, 100, 4, 801-817.

Kanemoto, Y. and MacLeod, W. B. (1992). "The Ratchet Effect and the Market for Secondhand Workers," *Journal of Labor Economics*, 10, 1, 85-98.

Lazear, E. P. (1986). "Salaries and Piece Rates," *Journal of Business*, 59, July, 405-431.

Lazear, E. P. and Rosen, S. (1981). "Rank-Order Tournaments as Optimal Labor Contracts," *Journal of Political Economy*, 89, October, 841-64.

MacLeod, W. B. and Malcomson, J (1989). "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57, March, 447-480.

Margiotta, M. M. and Miller, R. A. (1993). "Managerial Compensation and the Cost of Moral Hazard," GSIA Working Paper 1993-19, November, Carnegie Mellon.

Olkin, I., Gleser, L., and Derman, C. (1980). *Probability Models and Applications.* New York: Macmillan.

Pakes, A. (1986). "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54, July, 755-784.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1987). *Numerical Recipes: The Art of Scientific Computing,* New York, Cambridge.

Rust, J. (1987). "Optimal Replacement of GMC Bus Engines: An Empirical

Model of Harold Zurcher," *Econometrica*, 55, 999-1034.

Shearer, B. (1994). "Piece Rates, Principal Agent Models and Productivity Profiles: Parametric and Semi-Parametric Evidence," Department of Economics Working Paper 9406, Laval University.

Stiglitz, J. E. (1991). "Symposium on Organizations and Economics," *Journal of Economic Perspectives*, 5, 2, 15-24.

FIGURE 1:
NASH EQUILIBRIUM EFFORT AND OUTPUT

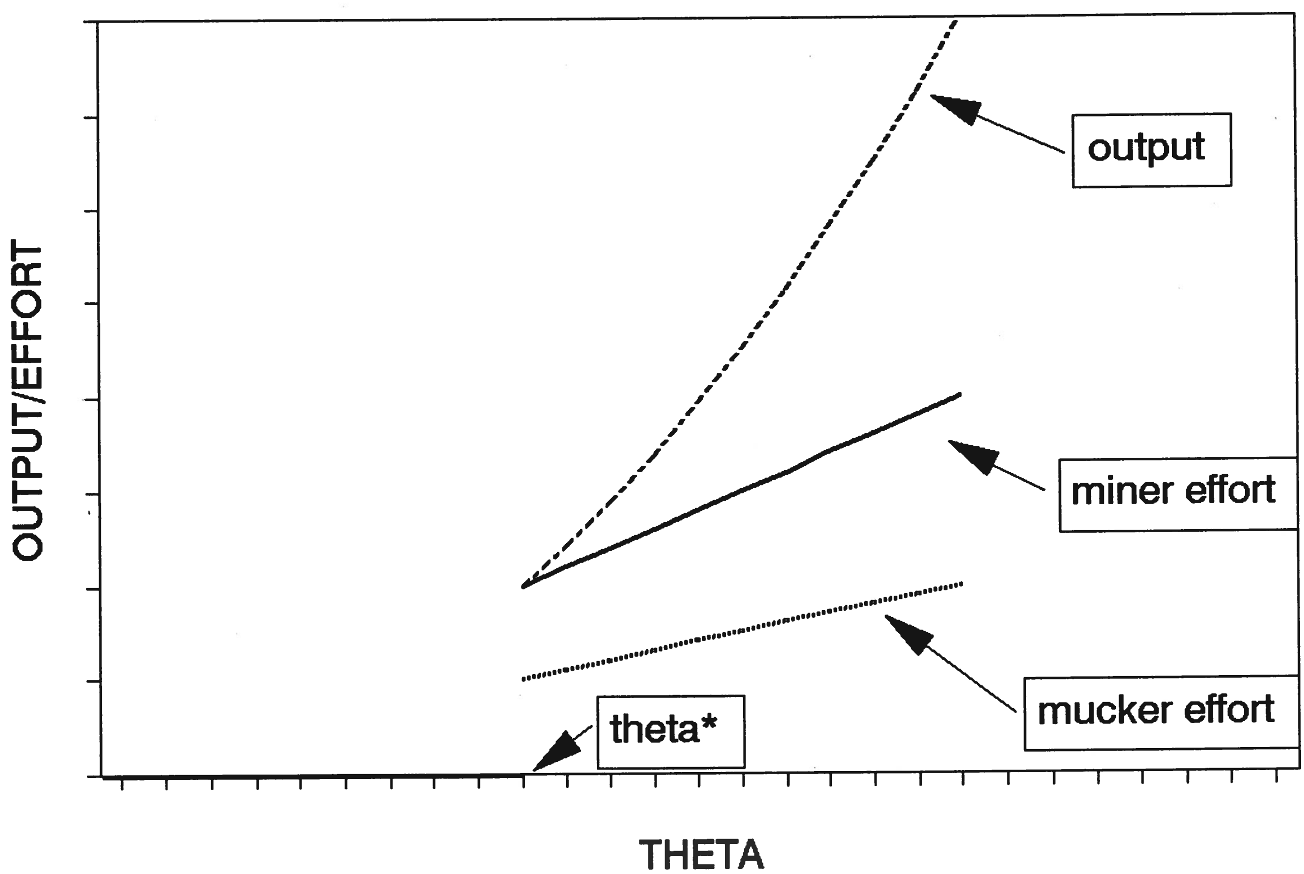

TABLE 1.
BONUS PAYMENTS TO MINERS AND MUCKERS: 1927-1928

	SHIFTS WORKED			TOTAL BONUS		POSITIVE BONUSES PER SHIFT			
MONTH	MINER TOTAL	MUCKER TOTAL	RATIO	PROP RECD	PROP= $0.50	AVG	STDEV	MIN	MAX
1	1422	2389	1.68	0.48	0.04	0.16	0.19	0.01	1.04
2	1349	2241	1.66	0.59	0.08	0.09	0.10	0.02	0.46
3	1462	2236	1.53	0.71	0.06	0.12	0.11	0.02	0.51
4	991	1769	1.79	0.67	0.04	0.12	0.14	0.02	0.60
5	1120	1852	1.65	0.70	0.10	0.13	0.15	0.02	0.59
6	1433	2098	1.46	0.72	0.05	0.13	0.14	0.02	0.64
7	1212	2091	1.73	0.79	0.06	0.22	0.42	0.02	2.38
8	1127	1996	1.77	0.82	0.08	0.12	0.12	0.02	0.52
9	1169	1961	1.68	0.68	0.05	0.19	0.23	0.02	0.95
10	1414	1874	1.33	0.74	0.06	0.11	0.12	0.02	0.61
11	1082	1824	1.69	0.56	0.04	0.18	0.20	0.02	0.74
12	946	1500	1.59	0.53	0.04	0.19	0.17	0.02	0.63
13	1103	1886	1.71	0.61	0.02	0.19	0.21	0.02	0.84
14	838	1833	2.19	0.81	0.04	0.23	0.28	0.02	1.40
15	872	1749	2.01	0.68	0.08	0.20	0.24	0.02	1.43
16	759	1503	1.98	0.68	0.09	0.19	0.39	0.02	2.06
17	923	1616	1.75	0.59	0.07	0.17	0.22	0.02	1.02
18	850	1823	2.14	0.61	0.09	0.14	0.16	0.02	0.83
19	569	1327	2.33	0.67	0.07	0.17	0.23	0.02	1.03
20	790	1492	1.89	0.61	0.01	0.17	0.19	0.00	0.70
21	779	1648	2.12	0.52	0.06	0.23	0.31	0.02	1.15
22	1090	1793	1.64	0.65	0.07	0.20	0.27	0.02	1.03
23	1288	1882	1.46	0.59	0.06	0.22	0.25	0.02	1.52
24	885	1490	1.68	0.64	0.08	0.23	0.28	0.02	1.14
AVG	1061	1828	1.77	0.65	0.06	0.17	0.21	0.02	0.99

Notes: Month 1 = January, 1927. Bonues are expressed in dollars, and are based on workers with 25 or more shifts in each month.

TABLE 2.
MAXIMUM LIKELIHOOD ESTIMATES OF UNRESTRICTED MODEL

TABLE 2(A): eta estimated

	eta
	10.49
	(0.00)

	PERIOD:	1	2	3	4	5	6	7	8	9	10	11	12
(1)	psi*	0.13	0.08	0.10	0.12	0.10	0.08	0.07	0.12	0.14	0.17	0.16	0.11
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
(2)	alphx/3	6.54	7.63	5.80	3.13	5.40	9.88	8.55	4.59	5.05	3.27	4.17	5.78
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
(3)	sigma	0.46	0.36	0.42	0.53	0.46	0.46	0.45	0.59	0.56	0.69	0.71	0.62
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	ln like	-7297											

TABLE 2(B): eta = 3.0

		1	2	3	4	5	6	7	8	9	10	11	12
(1)	psi*	0.22	0.13	0.14	0.13	0.13	0.16	0.10	0.16	0.21	0.23	0.22	0.17
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
(2)	alphx/3	2.99	3.51	2.85	2.18	3.06	3.65	3.85	2.45	2.52	1.76	2.37	2.64
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
(3)	sigma	0.60	0.47	0.52	0.57	0.54	0.67	0.57	0.69	0.68	0.80	0.81	0.78
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	ln like	-7310											
(4)	1-t3	0.06	0.09	0.08	0.06	0.07	0.04	0.06	0.04	0.04	0.02	0.02	0.03
(5)	-ln like	657.5	672.9	677.8	738.2	696.8	528.8	668.4	549.1	510.1	455.3	548.0	607.7
(6)	CHI^2	5.239	11.49	7.663	30.11 *	4.743	7.252	3.13	4.531	4.157	10.42	4.675	7.141

Notes: p-values for t-statistic are in brackets;
chi^2 refers to the chi squared test statistic for fitting the bonus distribution. psi* is the estimate of psi/exp(2*mu).
psi is defined in Lemma 3; t3 is defined in Corollary 2.
* indicates value of test statistic is significant at 1% level.

TABLE 3.
MAXIMUM LIKELIHOOD ESTIMATES OF STRUCTURAL PARAMETERS

Table 3(A): k_a = .24k_b

k_b	r
0.01	0.10
(0.00)	(0.00)

	Period	1	2	3	4	5	6	7	8	9	10	11	12
(1)	sigma	0.78	0.50	0.53	0.52	0.54	0.73	0.56	0.66	0.73	0.74	0.80	0.74
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
(2)	mu	-0.12	0.04	0.00	0.03	0.04	0.05	0.16	0.01	-0.07	-0.12	-0.08	0.00
		(0.20)	(0.60)	(0.97)	(0.72)	(0.57)	(0.57)	(0.02)	(0.95)	(0.43)	(0.23)	(0.44)	--
(3)	alpha	0.55	0.62	0.61	0.61	0.61	0.56	0.60	0.57	0.56	0.56	0.55	0.56
(4)	x	11.62	14.94	13.45	14.33	14.95	17.38	20.28	14.22	12.28	11.04	13.51	15.42
(5)	Chi^2	119.5 *											
	ln like	-7329											

Table 3(B): k_a = .01k_b

k_b	r
0.01	0.10
(0.00)	(0.00)

	Period	1	2	3	4	5	6	7	8	9	10	11	12
(1)	sigma	0.78	0.50	0.54	0.52	0.54	0.73	0.56	0.66	0.73	0.74	0.80	0.74
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
(2)	mu	-0.12	0.04	0.00	0.03	0.04	0.05	0.16	0.01	-0.07	-0.12	-0.08	0.00
		(0.20)	(0.59)	(0.97)	(0.72)	(0.57)	(0.57)	(0.02)	(0.95)	(0.43)	(0.23)	(0.43)	--
(3)	alpha	0.55	0.62	0.60	0.61	0.61	0.56	0.60	0.57	0.56	0.56	0.55	0.56
(4)	x	11.66	14.87	13.38	14.27	14.92	17.46	20.29	14.23	12.29	11.01	13.54	15.46
(5)	Chi^2	119.4 *											
	ln like	-7329											

Note: p-values for t statistics in brackets. * indicates test significant at 1% level.
r was estimated on a monthly basis and then converted to a per shift value by dividing by 25 (shifts).

TABLE 4.

EXPECTED OUTPUT UNDER LINEAR BONUS RELATIVE TO FULL INFORMATION

	TOTAL (1-t1*t2*t3)	SHIRKING (1-t3)	FREE RIDING (1-t2)	RETURN TO EFFORT (1-t1)	TOTAL ADJUSTED FOR E[v]
Case A (k_a=0.24k_b)	0.73	0.04	0.51	0.42	0.58
Case B (k_a=0.01k_b)	0.81	0.04	0.66	0.42	0.69

Note: Output under full information optimal contract equals 1.
Values are averages over twelve two-month periods. t1, t2, and t3 defined in Corollary 2. Adjustment in final column defined in Equation (9).

TABLE 5.

EXPECTED PROFITS UNDER ALTERNATIVE COMPENSATION SCHEMES

ENVIRONMENT/METHOD	PROFITS CASE A		PROFITS CASE B	
	ACTUAL	RELATIVE	ACTUAL	RELATIVE
(1) F.I. Optimal Contract	134.96	1.00	197.68	1.00
(2) I.I. Optimal Contract	72.80	0.54	78.24	0.40
(3) I.I. Two Bonus Rates	44.14	0.33	53.52	0.27
(4) I.I. Linear Bonus	37.86	0.28	38.11	0.19
(5) I.I. Piece Rate	37.01	0.27	37.31	0.19

Note: F.I.=Full Information, I.I.=Incomplete Information. Values are averages over twelve two month periods, except row (2) is computed using period 12 estimates. Case A and B defined in Table 3.