Models: Managerial remuneration - Risk and incentive tradeoff

1 Introduction

This text presents models and proofs of the classic trade-off between risk and incentives regarding the remuneration of managers by corporate owners. The presentation is rigorous. If you prefer a more intuitive discussion of the problem click here. Otherwise stay put. The models of this text are believed to be good representatives of the formal principal-agency literature of the hidden action type. The analysis follows Tirole [1988, page 51-55] and part 1 in Hart and Holmström [1987]. The text proceeds as follows: Section 2 outlines the typical agency model framework its assumptions and notational apparatus. Then in section 3, the four agency models are presented. Note that footnotes e.g. '[2]' can be viewed by clicking them.

2 The General Framework

Model Problem:

The problem is to study incentive and insurance (risk) issues between a manager [the agent] and some shareholders [the principals]. The shareholders are owners of a firm and they have hired the manager to run it. Results are generated the usual way by changing the underlying assumptions.

General Assumptions:

The following assumptions are part of all the following models unless otherwise specified.

1) Shareholders are risk neutral.[1] For a graphic explanation of the different risk concepts click here.
2) Shareholders objective is to maximize the firms expected profit.
3) There is only one manager.
4) The manager's work effort (e) is unobservable by the shareholders (asymmetric information).
5) Firm profit $\prod(e,\theta)$, manager wage $w(\prod)$, and manager utility $U(w,e)$ are common knowledge.

Neither Tirole nor Hart & Holmström explicitly state the following assumptions. However, it is believed to be important to state them explicitly. If not, one may easily get the wrong impression that the predictions of the models are more robust than they actually are. Experienced economists

[1] From a financial point of view this may seem as a very crude assumption. However, the assumption is only simplifying the mathematics. There is good reason to believe that all the implications of the models holds as long as the risk aversion of the shareholders is significantly less than the risk aversion of the manager. This may be reasonably for two reasons: 1) The shareholders are more capable of diversifying their portfolios [i.e. shareholders face beta risk, the manager face plain variance]. 2) The shareholders can only be expected to have a small fraction of their total wealth in the firm, but the manager may have a large fraction.

know this but outsiders do not.

6) All agents (the manager and the principals) maximize utility. A trivial, tautological assumption. Says that agents want to get more of what they like.
7) All agents are perfectly rational. An unrealistic but strongly simplifying assumption. Says that agents grasp all the information in the model and they have the ability to calculate the optimal decisions instantaneously and costless.
8) Time is static. Everything is as if the world only existed in one timeless period.

The models include a many other assumptions, some made explicit below. However, for expositional reasons they will remain implicit.

Notation and Further Assumptions:

- e is work effort by the manager, and e take on intervals $e \in [e^{min}, e^{max}] = E$.
- θ is uncertainty or state of nature affecting firm profits independent of manager's work effort e. Also $\theta \in \Theta$, where Θ is the distribution space. Θ is assumed to be common knowledge.
- $\prod = \prod(e,\theta)$ is random firm profit. Assume $\prod'_e > 0$. This assumption is crucial in all agency theory. Because $\prod$ depend on θ, $\prod$ is itself stochastic. Assume $\prod \in [\prod^{min}, \prod^{max}] = P$.
- $w(\prod)$ is the manager's wage function or salary scheme. Presumably $w'_{\prod} > 0$ (This will be demonstrated to be the conclusive characteristic of the optimal wage function in the general agency model, case 4 below). In the principal-agent literature $w(\prod)$ is also interpreted as the wage contract that the shareholders design. It is assumed that there exists a wage scheme space W, $w \in W$.
- U_0 is the reservation utility. This is, the expected utility the manager believes he can get in a job elsewhere, less the cost of searching. In other words, the reservation utility is the utility he could get by redeploying his resources at the best alternative use.
- $\Phi(e)$ is the manager's disutility from effort. Assume $\Phi'_e \geq 0$, $\Phi''_{ee} > 0$, $\Phi'(0)=0$, and $\Phi'(\infty)=\infty$.
- $U(w,e) = u[w(\prod(e,\theta))] - \Phi(e)$ is the utility function of the manager. Assume $U'_w > 0$, $U''_{ww} < 0$. This assumption implies that the manager is risk averse. Note that the utility function is separable in income and effort. This simplifies the analysis.
- $f(\prod;e)$ is the density function or mass function for the stochastic firm profit $\prod$ given the effort level {e}.

The Model:

Shareholders expected payoff: $\pi^{Sha} = E_\theta[\prod(e,\theta) - w(\prod(e,\theta))] = \int_{\{\prod \in P\}} [\prod(e,\theta) - w(\prod(e,\theta))] f(\prod;e) d\prod$

Managers expected payoff: $\pi^{Man} = E_\theta[U(w(\prod(e,\theta)),e)] = \int_{\{\prod \in P\}} u(w(\prod(e,\theta)) f(\prod;e) d\prod - \Phi(e)$

The shareholders must design a wage contract w*(.) and pick a effort level e* that maximizes their expected payoff:

$$\{w^*(.), e^*\} \equiv \underset{\{w(.)\in W,\, e\in E\}}{\operatorname{argmax}}\ \pi^{Sha} = \underset{\{w(.)\in W,\, e\in E\}}{\operatorname{argmax}} \int [\prod(e,\theta) - w(\prod(e,\theta))]f(\prod;e)d\prod \qquad (1)$$

subject to,

$$\underset{\{e\in E\}}{\operatorname{argmax}}\ \pi^{Man} = \underset{\{e\in E\}}{\operatorname{argmax}} \int u(w(\prod(e,\theta)))f(\prod;e)d\prod - \Phi(e) \geq U_0,\ \text{and} \qquad (IR,2)$$

$$\int_{\{\prod\in sP\}} u(w(\prod(e^*,\theta)))f(\prod;e^*)d\prod - \Phi(e^*) \geq \int_{\{\prod\in P\}} u(w(\prod(e,\theta)))f(\prod;e)d\prod - \Phi(e),\ \forall\ e\in[e^{min}, e^{max}] \qquad (IC,3)$$

The first constraint is called the **individual rationality constraint** (IR) or individual participation constraint. It states that given the wage contract {w(.)} a rational agent will require the utility from the effort level {e} that maximizes his utility to be at least equal to his reservation utility $\{U_0\}$. This constraint is comparable to the volunteer trade assumption in the perfect market economy. Standing alone it cannot prevent the problem from reaching first best. The second restriction is the **incentive compatibility** (IC) constraint. It states that given the observable wage contract {w(.)}, the unobservable level of effort chosen by the principal {e*} must maximize the utility of the agent over all possible levels of effort $\{\forall e\in[e^{min}, e^{max}]\}$ for the agent to have an incentive to actual conduct that level of effort. Without going into detail this restriction embed the moral hazard issues of the model. The above model provides a general framework for the corporate principal agent problem. The following studies this framework in four different cases.

3 The Basic Models

Case 1: The Full Information Case

Modified Assumptions

1) The shareholders observe effort {e}. This is, full information or perfect monitoring. Replaces assumption 4 above.
2) The manager is risk averse. That is, $U''_{ww} < 0$.

When the shareholders can observe the manager's effort the problem becomes much more simplified since they can impose any effort level {e} by threatening with severe punishment should the effort order not be obeyed. The incentive compatibility constraint (IC) is therefore not relevant. For simplicity assume e is solved. Then the wage contract becomes the only relevant decision variable. The problem is therefore that the shareholders must design a wage contract w*(.) that maximizes their expected payoff:

$$\{w^*(.)\} \equiv \underset{\{w(.)\in W\}}{\operatorname{argmax}} \int [\prod(e,\theta) - w(\prod(e,\theta))]f(\prod;e)d\prod \qquad (4)$$

subject to,

$$\underset{\{e\in E\}}{\operatorname{argmax}} \int u(w(\prod(e,\theta)))f(\prod;e)d\prod - \Phi(e) \geq U_0 \qquad (IR,5)$$

The individual rationality constraint must be binding. If not the shareholders would have incentive to lower the salary {w(∏)} until it becomes binding to save on their own cost of management services. The Lagrangian may now be formed.

$$L = \int_{\{\prod\in P\}} \{[\prod(e,\theta) - w(\prod(e,\theta))]f(\prod(e,\theta);e) + \lambda[u(w(\prod(e,\theta))) - \Phi(e) - U_0]f(\prod(e,\theta);e)\}d\prod$$

F.O.C. with respect to w: $f(\prod(e,\theta);e) + \lambda u'_w(w(\prod(e,\theta)))f(\prod(e,\theta);e) = 0$

<=>

$$u'_w(w(\prod(e,\theta))) = 1/\lambda \qquad (6)$$

Q.E.D.

So the optimal wage structure is independent of profit. In other words, the agent is given a constant wage in order to be fully insured. The optimal choice of effort level may be found by taking the derivative with respect to e. Analysis is omitted. The intuition is that if the shareholders are risk neutral and the manager is risk averse, and then all the risk should be borne by the risk neutral part because he can bear the risk costless. Furthermore, it does not create any incentive problem because in this model version the effort level is observable. Threatening with severe punishment can impose the efficient level of effort. This full information solution is a first best solution. It does not deviate from what would be obtained in the perfect market economy model. The "mechanism" that ensures first best is not the market mechanism but the existence of an omniscient planner [the principal or the shareholders]. He has all the information necessary to calculate the first best solution, and the powers [perfect monitoring and punishment abilities] to ensure that it is actually carried out. The assumption of symmetric information on effort level does not fit many real world situations. It would probably not be wrong to assume that most principal-agent relations of the shareholder-manager type imply a significant degree of asymmetric information. This is the issue in the following three models.

Case 2: Asymmetric information, and risk neutral manager.

Modified Assumptions

1) The shareholders cannot observe effort {e}. Back to assumption 4 above.
2) The manager is risk neutral. This is, $U'_w > 0$, and $U''_{ww} = 0$.
3) The utility function below is now assumed to be measuring money equivalents. This implies that utility can be measured cardinally. The reason for this assumption is that we need to make the manager indifferent between receiving profit or utility.

Modified Notation and Further Assumptions

- $U(w,e) = w(\prod(e,\theta)) - \Phi(e)$ is the utility function of the manager. This utility function is clearly risk neutral since $U''_{ww} = 0$. This is new!

Now the manager's utility is: $U(w,e) = w(\prod(e,\theta)) - \Phi(e)$

<=>

$$w(\prod(e,\theta)) = U(w,e) + \Phi(e) \tag{7}$$

With this manager-utility function, the IR constraint becomes:

$$\int w(\prod(e,\theta))f(\prod;e)d\prod - \Phi(e) \geq U_0 \quad \{e\in E\}$$

and it must be binding. Otherwise the manager could just decrease the wage until it becomes binding. The IR restriction may therefore be written:

$$\int U(w,e)f(\prod;e)d\prod = U_0 \tag{IR,8}$$

$\{e \in E\}$

Then the shareholder's payoff function becomes:

$$\pi^{Sha} = E_{\theta}[\prod(e,\theta) - w(\prod(e,\theta))] = \int_{\{\prod \in P\}} [\prod(e,\theta) - w(\prod(e,\theta))]f(\prod;e)d\prod$$

(by substitution of (7) above)

$$= \int_{\{\prod \in P\}} [\prod(e,\theta) - U(w,e) - \Phi(e)]f(\prod;e)d\prod$$

(by substitution of (8) above)

$$= \int_{\{\prod \in P\}} [\prod(e,\theta) - U_0 - \Phi(e)]f(\prod;e)d\prod$$

$$= \int_{\{\prod \in P\}} [\prod(e,\theta)]f(\prod;e)d\prod - U_0 - \Phi(e)$$

Solving this with respect to the optimal effort decision {e*} yields the same first best outcome as under the full information case 1 above, since only the IR restriction has been used. This is, for given w:

$$\{e^*\} = \underset{\{e \in E\}}{\operatorname{argmax}}[\int [\prod(e,\theta)]f(\prod;e)d\prod - U_0 - \Phi(e)] \tag{9}$$

The following demonstrates that a wage contract making the agent the residual claimant of the firm's profit gives the manager the full incentive to maximize the principals' profit. This will result in a solution identical to the first best, full information model. The agent may be made a residual claimant on the firms profit if he is offered to lease the firm at an annual price of:[2]

$$p = \underset{\{e \in E\}}{\operatorname{argmax}}[\int \prod(e,\theta)f(\prod;e)d\prod - U_0 - \Phi(e)] \tag{10}$$

Note that this is the value that maximizes the shareholder's value under full information. The question is now; does the manager accept this lease? The answer is YES! His payoff function when he has to pay the lease p and has the right to the firms profit is:

$$\pi^{Man} = \int_{\{\prod \in P\}} [\prod(e,\theta)f(\prod;e) - \Phi(e) - p]d\prod \qquad \text{(Note! Use assumption 3 above)}$$

[2] A lease is a more correct formulation than the sale proposed by Tirole [1988, page 52]. If the firm was sold it should be done at the present value of the annual profit flows. Mathematically it would complicate matters. However, more subtle it would impose stronger demands on the principals' and the agent's rationality. They would have to make unbiased estimates of the entire future profit flow from the firm. With a one-year lease they could sign a new lease each year modified according to the latest information. This seems a lot more likely. This is my own idea, but I would not be surprised if it is made elsewhere.

<--firm profit--><-work disu.-><- lease->

$$= \int_{\{\prod \in P\}} [\prod(e,\theta)f(\prod;e) - \Phi(e) - \{\prod(e,\theta)f(\prod;e) - U_0 - \Phi(e)\}]d\prod = U_0 \quad (11)$$

Q.E.D.

Since the manager receives his reservation utility he accepts. The shareholders and the manager get exactly the same payoffs as under full information. However, the situation is very different. In this model the agent must bear the entire risk in order to have the full incentive to maximize firm profit. This strategy was not optimal in case 1, because the manager was risk averse. Neither was it necessary for inducement of incentives, because of the special powers of the principal [perfect monitoring and punishment]. In this case, our manager is risk neutral and is therefore capable of bearing the risk costlessly. This is fortunate, because the principals have no other ways [imperfect monitoring and punishment powers] to induce the optimal solution than to make the agent the residual claimant of the firm's profit.

The nice feature with the model above is that it induces optimal resource allocation without having an omniscient principal dictating the actions. In this respect, it is like the market mechanism. Unfortunately, it assumes risk neutral agents. This might be reasonable if the agent is rich relative to the size of the firm profit. However, this is not a fact of real world life. Among the private enterprises most GDP value is made in big corporations with profits that are many times higher than their manager's salaries. The next two models analyze more realistic situations with risk averse managers.

Case 3: Asymmetric information, and an infinitely risk averse manager.

Modified Assumptions

1) The manager is infinitely risk averse. That is, $U'_w > 0$, and $U''_{ww} <<< 0$.
2) The shareholders cannot observe effort {e}. This is asymmetric information.
3) For simplicity assume that the distribution of the profit $\prod = \prod(e,\theta)$ is such that $\prod \in [\prod^{min},\prod^{max}]=P$, and that the distribution borders are independent of the effort level. We still have that $\prod'_e >0$, but this is so only when we look at the mean value of the profit.

This case turns out to be the simplest in terms of mathematics. An implication of infinitely risk aversion is that the manager will avoid risk at any cost in mean return. However, the agent is capable to make a choice between two risky payoffs. To be specific, the agent always prefers a random wage w_1 and a level of effort e_1 to another pair (w_2,e_2) if $\min(w_1) > \min(w_2)$ or $\min(w_1) = \min(w_2)$ and $e_1 < e_2$. Note that in the infinite risk averse case the mean value is not a relevant decision variable. Now, look at the manager's payoff function replicated below for convenience and consider assumption 3 above:

Manager's expected payoff: $\pi^{Man} = E[U(w(\prod(e,\theta)),e)] = \int_{\theta} u(w(\prod(e,\theta))f(\prod;e)d\prod - \Phi(e)$ $\{\prod \in P\}$

The wage {w} depends on profit {∏} that again depends on effort {e}. However, according to assumption 3, the wage only depends on profit with respect to the mean value, not with respect to the border values $\{[\prod^{min}, \prod^{max}]\}$. So the minimum wage is independent of work effort. Furthermore, because our utility function {U} only attaches value to the minimum wage, the effort level {e} only affects the managers payoff through the disutility term Φ(e). Since $\Phi'_e > 0$ the manager will optimize by choosing e^{min} no matter how much his expected salary would increase if he cared to work more than e^{min}. Given this fact the shareholders cannot do better than to give the manager his reservation utility evaluated at e^{min}. The resulting payoffs becomes:

Shareholder's expected payoff: $\pi^{Sha} = [\prod(e^{min},\theta) - w(\prod(e^{min},\theta))]$

Manager's expected payoff: $\pi^{Man} = u(w(\prod(e^{min},\theta))) - \Phi(e^{min}) = U_0^{min}$

Q.E.D.

In this model, the manager gets a constant wage equal to his reservation utility evaluated at e^{min}, and the shareholders get a profit that is clearly below first best e*. Naturally, this conclusion depends as any other model on the nature of the assumptions. Especially assumption 3 which could be called the assumption of unmoving support. This is, unmoving borders. It would probably be more realistic to assume a probability distribution where $\prod^{min}$ is strictly increasing with e $[\prod^{min}{}'_e(e) > 0]$. If the principal knows this relation and we further assume that the principals have punishment powers then the case reverts back to one as the full information case. The following explains why.

Any effort level could be obtained by threatening with a punishment. This is so because to each effort level {e} there correspond a unique density function $\{f(\prod(e,\theta);e)\}$ with a unique minimum bound of the distribution tale $\{\prod^{min}\}$. If the effort level e_0 is wanted there will be a unique border, $\prod^{min}_0$. However, if this border turns out to be violated then the principals know that the agent has lied about his effort level. Even if it was very unlikely that a bad state {θ} turned out and produced the situation $\prod^{min}_0 > \prod(e\sim,\theta)$ where $e\sim < e_0$, the fraud could be prevented by threatening with a monetary punishment. Indeed, a small punishment would suffice because of our assumption of infinite risk aversion.[3] The case of infinitely risk aversion is extreme. Most people, if not all, have a more moderate risk aversion and their utility functions consider both the

[3] Tirole [1988, page 53, footnote 104] makes an error when he claims that the punishment must be very large. He must have had in mind another model with a more moderate degree of risk aversion.

spread/variance and the mean value of any stochastic income they may have. We will now turn to this general case.

Case 4: Asymmetric information, and a "normal" risk averse manager.

This model is *the model* in the corporate principal-agency literature. The model framework was presented in section 2 and it is not duplicated here. However, the essential assumptions are restated and a few new are added.

Modified Assumptions

1) The manager is "normal" risk averse. That is, $U'_w > 0$, and $U''_{ww} < 0$.
2) The shareholders cannot observe effort {e}. This is, asymmetric information.

Modified Notation and Further Assumptions

- $U(w,e)= u(w(\prod(e,\theta))) - \Phi(e)$ is the separable utility function of the manager.
- $F(\prod;e)$ is the cumulative distribution function on $\prod\in[\prod^{min}, \prod^{max}]=P$, with density function $f(\prod;e)$. Both are differentiable in effort {e}. The firm may be perceived as an asset, yielding a return {∏} conditional on the effort level {e}. The more effort, the higher the returns on the firm's asset. This relation is assumed to follow first order stochastic dominance. This is, $e_1>e_2$ => $F(\prod;e_1)\leq F(\prod;e_2)\ \forall\ \theta\in\Theta$, and $F(\prod;e_1) < F(\prod;e_2)$ for some $\theta\in\Theta$[4]. Click here to get a graphic explanation of first-order stochastic dominance. An implication of this is that $F_e(\prod;e)< 0$.

Consider the shareholders' problem in section 2 reproduced here for convenience. The shareholders must design a wage contract w*(.) and pick a effort level e* that maximizes their expected payoff:

$$\{w^*(.), e^*\} \equiv \underset{\{w(.)\in W,\, e\in E\}}{\text{argmax}}\ \pi^{Sha} = \underset{\{w(.)\in W,\, e\in E\}}{\text{argmax}} \int [\prod(e,\theta) - w(\prod(e,\theta))]f(\prod;e)d\prod \qquad (1)$$

subject to,

$$\underset{\{e\in E\}}{\text{argmax}}\ \pi^{Man} = \underset{\{e\in E\}}{\text{argmax}} \int u(w(\prod(e,\theta)))f(\prod;e)d\prod - \Phi(e) \geq U_0, \text{ and} \qquad (IR,2)$$

$$\int_{\{\prod\in sP\}} u(w(\prod(e^*,\theta)))f(\prod;e^*)d\prod - \Phi(e^*) \geq \int_{\{\prod\in P\}} u(w(\prod(e,\theta)))f(\prod;e)d\prod - \Phi(e),\ \forall\ e\in[e^{min}, e^{max}]\ (IC,3)$$

[4] Tirole's [1988] definition of first order stochastic dominance is incomplete. He only considers $F(\prod;e_1) \leq F(\prod;e_2)$. For a good reference to the concept of stochastic dominance, see Copeland and Weston [1988, page 92].

One way to solve it is to form the Lagrangian. The IC restriction is in itself an optimization problem that must be solved before the Lagrangian can be formed.5 The first order condition to this is:

F.O.C. with respect to e: $\int_{\{\prod \in P\}} u(w(\prod(e,\theta)))f'_e(\prod;e)d\prod - \Phi'_e(e) = 0,$

and the second order condition is:

S.O.C. with respect to e: $\int_{\{\prod \in P\}} u(w(\prod(e,\theta)))f''_{ee}(\prod;e)d\prod - \Phi''_{ee}(e) = 0,$

Elsewhere it was explained that the IR constraint is binding. The Lagrangian may now be formed.

$$L = \int_{\{\prod \in P\}} \{[\prod(e,\theta) - w(\prod(e,\theta))]f(\prod(e,\theta);e) + \lambda[u(w(\prod(e,\theta))) - \Phi(e) - U_0]f(\prod(e,\theta);e) + \eta[u(w(\prod(e,\theta)))f'_e(\prod(e,\theta);e) - \Phi'(e)f(\prod(e,\theta);e)]\}d\prod$$

and,

$$\partial L/\partial w = - f(\prod(e,\theta);e) + \lambda u'_w(w(\prod(e,\theta)))f(\prod(e,\theta);e) + \eta u'_w(w(\prod(e,\theta)))f'_e(\prod(e,\theta);e) = 0$$

$$\Leftrightarrow$$

$$f(\prod(e,\theta);e) = u'_w(w(\prod(e,\theta)))[\lambda\, f(\prod(e,\theta);e) + \eta f'_e(\prod(e,\theta);e)]$$

$$\Leftrightarrow$$

$$1 = u'_w(w(\prod(e,\theta)))[\lambda + \eta f'_e(\prod(e,\theta);e) / f(\prod(e,\theta);e)]$$

$$\Leftrightarrow$$

$$1/u'_w(w(\prod(e,\theta))) = \lambda + \eta f'_e(\prod(e,\theta);e) / f(\prod(e,\theta);e) \qquad (12)$$

The interpretation of this equation is easier when one considers the simplified case with only two effort levels: High effort $\{e_H\}$ and low effort $\{e_L\}$. The assumption that the owners get the most payoff from high managerial effort has to be added. If not the owners would implement low effort by paying the manager a fixed income because that would yield the optimal risk sharing. This is not interesting since it does not show how the trade-off between insurance and incentive is working. The Lagrangian in the two effort level case becomes:

$$L = \int_{\{\prod \in P\}} \{[\prod(e,\theta) - w(\prod(e,\theta))]f_H(\prod(e,\theta)) + \lambda[u(w(\prod(e,\theta))) - \Phi(e) - U_0]f_H(\prod(e,\theta)) + \eta[u(w(\prod(e,\theta)))f_H(\prod(e,\theta)) - \Phi(e_H) - [u(w(\prod(e,\theta)))f_L(\prod(e,\theta)) - \Phi(e_L)]]\}d\prod$$

and,

$$\partial L/\partial w = - f_H(\prod(e,\theta)) + \lambda u'_w(w(\prod(e,\theta)))f_H(\prod(e,\theta)) + \eta u'_w(w(\prod(e,\theta)))f_H(\prod(e,\theta)) - \eta u'_w(w(\prod(e,\theta)))f_L(\prod(e,\theta)) = 0$$

$$\Leftrightarrow$$

5 This way of making the IC constraint more tractable for Lagrangian optimization is referred to as the *first order approach* in the literature.

$$f_H(\prod(e,\theta)) = u'_w(w(\prod(e,\theta)))f_H(\prod(e,\theta))[\lambda + \eta - \eta f_L(\prod(e,\theta))/f_H(\prod(e,\theta))]$$

<=>

$$1 = u'_w(w(\prod(e,\theta)))[\lambda + \eta[1 + f_L(\prod(e,\theta)) / f_H(\prod(e,\theta))]]$$

<=>

$$1/u'_w(w(\prod(e,\theta))) = \lambda + \eta[1 + f_L(\prod(e,\theta)) / f_H(\prod(e,\theta))] \qquad (13)$$

The term $f_L(\prod(e,\theta)) / f_H(\prod(e,\theta))$ is equal to the likelihood ratio used to construct statistical tests.[6] When the principals observe the profit $\prod$, the ratio tell them the likelihood that the true distribution producing the observed $\prod$ comes from the $f_L(\prod(e,\theta))$ distribution rather than the $f_H(\prod(e,\theta))$ distribution. If the ratio is high then the manager will probably work with low effort, and if it is high he will probably work with high effort. If the ratio is equal to one then it is equally likely that the manager work high or low.

Note that if the observed profit $\prod$ is high then the ratio will be small, since it speaks for high effort level. This makes the right hand side of the equation high. For the equation to hold the left hand side must goes up as well. This is u'_w must fall. This implies that the optimal wage w must rise since $u''_{ww} < 0$. This is the mayor point of the general agency model. *The optimal wage structure must induce the agent to high effort by paying him partly as a function of profit.* Specifically $w'_\Pi > 0$. This makes the wage risky and the agent "sacrifice some utility" by not being fully insured. The IR must be fulfilled so the mean wage has to be higher than under full insurance. Relative to the first best models (case 1 and 2 above) the principals loose. However, in this case it pays because $\prod'_e > 0$ and by making w such that $w'_\Pi > 0$ the principals make a net gain on the agent's higher effort. Note that because e is not directly observable the principals are forced to make use of the $\prod'_e > 0$ effect.

With the above in mind, it should be easier to comprehend equation (12). The term $f'_e(\prod(e,\theta);e) / f(\prod(e,\theta);e)$ is the likelihood ratio in equation (12), and it is comparable to $f_H(\prod(e,\theta)) / f_L(\prod(e,\theta))$. The only difference is that the term in (12) is continuous in the effort level whereas the ratio in (13) is not.[7] The economic interpretation is the same: *The wage should be made an increasing function of profit in order to induce profitable effort levels.* This conclusion turns out to depend on the characteristic of the distribution of $\prod$. Two conditions must be satisfied for the first order approach to yield valid results that maximizes the shareholders' problem with respect to w and e:

1) *Monotonic-likelihood-ratio property*: f'_e / f (or f_H / f_L) increases with $\prod$, and
2) *Convexity of the distribution function*: This is, $F''_{ee} \geq 0$, or $\prod$; $\alpha e_1 + (1 - \alpha)e_2) \leq \alpha F(\prod;e_1) + (1 -$

[6] Formally the likelihood ratio is $r(\mathbf{x}) = L(\theta_0|\mathbf{x})/\max_{\theta\in\Theta}L(\theta|\mathbf{x}) = L(\theta_0|\mathbf{x}) / L(\theta\sim|\mathbf{x})$, where $\mathbf{x} = (x_1,...,x_n)$ is the observed values of a random variable X with density function $f(\mathbf{x}|\theta)$, $\theta\in\Theta$. This ratio, test the hypothesis H_0:θ=q_0, against the alternative; H_1: $\theta\neq\theta_0$. The closer $r(\mathbf{x})$ is to one, the better the observed data support H_0. For a statistical book including the topic, see Andersen, Jensen, and Kousgaard [1984].

[7] See Milgrom [1981] for a discussion of likelihood ratios on this particular subject.

$\alpha)F(\prod;e_2)$, $\forall$ e_1, e_2, and $\alpha \in [0, 1]$.

Condition 2 says that the distribution from a certain effort level $e = \alpha e_1 + (1 - \alpha)e_2$ is stochastically dominating the composite distribution randomizing between the two effort levels e_1, e_2. The conditions ensure that the manager has a concave objective function. This is needed for the Lagrangian to yield a global maximum.

The reason that condition 1 is needed, is that first order stochastic dominance is a very week form of dominance. It only considers the cumulative probability distributions. One could construct a non-unimodal density function that satisfied first order stochastic dominance. This function would have $\prod$ levels between its modes, where higher $\prod$ signal lower effort! This is avoided by condition 1. Although no empirical research has investigated how likely non-unimodal $\prod$ density functions is it seems safe to say that condition 1 is not a strong requirement. However, external factors and accounting doctrines does affect $\prod$ besides e. Therefore, it may be very likely that high $\prod$ does not signal high effort and condition 1 may seems strong.
Comparing the models may generate some insight. Moving from case 4 to case 3 where the effort level is e^{min}, the principals loose, but the agent remains on his reservation utility. When case 4 is compared to case 1 and 2, the agent gets paid more but receives the same utility $\{U_0\}$. The principals loose because E[w] is higher.

It is also interesting to analyze the incentive distortion from insurance.[8] Starting from case 3 with e^{min} and a constant wage w^{min} (obeying $u(w(\prod(e^{min},\theta))) - \Phi(e^{min}) = U_0^{min}$, complete insurance), consider the continuum of effort levels up to e* that is induced in case 2 by the other extreme of residual payment (zero insurance). With moderate risk aversion the optimal solution will be e^, where $e^{min} <$ e^ < e*, and the agent is partly insured.[9] This is true assuming decreasing marginal returns to $\prod$ by increased e, this is $\prod''_{ee} < 0$. Furthermore, assume that the principal can induce any effort level ($e = \alpha e^{min} + (1 - \alpha)e^*$) from e^{min} to e* by picking a fraction α and accordingly pay wage $w = \alpha w^{min} + (1 - \alpha)[\prod(e,\theta)f(\prod;e) - \Phi(e) - p]$. This wage scheme represents different degrees of insurance and incentive, and it includes the two extremes; case 2 and 3 when $\alpha = 0$ and 1 respectively. The wage scheme implies increased risk for increased effort levels and the marginal cost of risk is therefore increasing. It should therefore be clear why moderate risk aversion yields the optimal solution e^, where $e^{min} <$ e^ < e*. Some fraction α will induce e^. From e^ it does not pay to increase e further because then $MC^{Risk} > MR^{Incentive}$, and neither will it pays to decrease e^ since that implies $MC^{Risk} < MR^{Incentive}$.

Finally, an interesting and simple extension of the basic principal-agency model

8 This paragraph is my own idea.

9 e^ could be calculated by taking the derivative to the Lagrangian above. Both Tirole [1988], and Hart and Holmström [1987] have omitted this analysis. As can be seen from the Lagrange equation, it would be a very tedious expression probably very though to interpretate.

should be mentioned. In the real world profit $\prod$ is not the only information that shareholders observe. They may also have information on performance of agents in stochastically related firms, or general economic conditions under which the agent operate. The point is that this additional information may convey further information about the managers effort level. Suppose this information is called s for signal. It can be modeled very easy by replacing the cumulative distribution function $F(\prod;e)$ with the joint distribution of $\prod$ and s for given effort $G(\prod,s;e)$. Also $g(\prod,s;e)$ replaces $f(\prod;e)$. Everything else is as before. The optimal wage scheme equation (12) is therefore:

$$1/u'_w(w(\prod,s)) = \lambda + \eta g'_e(\prod,s;e) / g(\prod,s;e) \qquad (14)$$

This is very similar to equation (12) and indeed if

$$g'_e(\prod,s;e) / g(\prod,s;e) = f'_e(\prod;e) / f(\prod;e) = r(\prod;e) \qquad (15)$$

it would be identical. In that case it would be a waste of time to collect s. Equation (15) implies that

$$g(\prod,s;e) = A(\prod,s)B(\prod;e) \ \forall \ e\in[e^{min}, e^{max}] \qquad (16)$$

This is the called the *sufficient statistic condition.*[10] It says that $\prod$ is a sufficient statistic for $(\prod,s)$. It says that s is valuable if and only if it conveys information about effort e that adds to the information already brought about by $\prod$. This result is interesting because it opens up for the discussion of what additional information could be relevant to consider from the perspective of shareholders.

[10] Grossman and Hart [1983], Holmström [1979], and Shavell [1979] provide a more formal discussion of this theorem.