Motivating Informed Decisions

Andres Zambrano*
Universidad de los Andes

January 31, 2013

Abstract

This paper studies a principal-agent model where a risk-neutral principal delegates to a risk-neutral agent the decision of whether to pursue a risky project or a safe one. The return from the risky project is unknown and the agent can acquire costly unobservable information about it before taking the decision. The problem has features of moral hazard and hidden information since the acquisition of information and its content are unobservable to the principal. The optimal contract suggests that the principal should only reward the agent for outcomes that are significantly better than the safe return. It is also optimal to distort the project choice in favor of the risky one as a mechanism to induce the direct revelation of the uncertain state. In a managerial context, the findings explain why options and profit sharing compensation induce better decision making from CEOs, as well as why excessive risk taking might be optimal.

Keywords: Information Acquisition, Private Information, Contract, CEO compensation

JEL codes: D82, D83, D86, M12

---

*E-mail: azambrano@ucla.edu. Please download the most recent version at https://sites.google.com/site/andreszam/. This paper constitutes the first chapter of my Ph.D dissertation at UCLA. The author is indebted to Andrew Atkeson and Hugo Hopenhayn for their helpful advice during all these years. I also gratefully acknowledge the comments from Simon Board, Jernej Copic, Fernando Leibovici, Moritz Meyer-ter-Vehn, Ichiro Obara, John Riley, Venky Venkateswaran, Pierre-Olivier Weill, and Bill Zame, as well as participants of the student seminars at UCLA, and presentations at JEI 2012, Universidad del Rosario, Universidad de los Andes, Universidad Carlos III, Universitat Pompeu Fabra, the Midwest Theory Conference and the Midwest Economics Association Conference. Funding from the Colombian Central Bank and NBER is also acknowledged. Remaining errors are mine.

1 Introduction

It has been argued that the recent financial crisis has been caused by excessive risk taking from CEOs and that such risk taking is misaligned with shareholders interests. Moreover, it has been suggested that option-type contracts are the cause of this misalignment (Dong et al., 2010). This paper uses a model of delegated expertise to explain why stock options and profit sharing are optimal forms of compensation when a CEO has to be motivated to take informed decisions. It also provides an explanation of why excessive risk taking is optimal from a shareholder perspective, where excessive risk taking is understood here as pursuing a risky project even though it is ex-ante inefficient.

In the proposed model a principal can hire an agent to decide between a risky and a safe project. Before taking the decision, the agent can acquire information about the risky project by exerting costly effort. However, both the effort and the acquired information are unobservable to the principal. Thus this framework shares features of moral hazard and hidden information, and incentives must be used to motivate both information acquisition and the (partial) revelation of the obtained information.

Incentives in this scenario are potentially different than the ones in standard moral hazard problems since effort does not generate greater expected returns directly. In contrast, the unobservable action taken by the agent generates a privately observed signal that improves the decision-making. The studied setup is very general and the only restriction imposed is that signals can be ordered in the likelihood ratio order (Milgrom, 1981). Under limited liability for both individuals, the optimal wages suggest that the agent should be rewarded with the return of the risky project only if it is significantly better than the safe return. Moreover, it is optimal to distort the project choice in favor of the risky project as a strategy to reveal the uncertainty directly.

The intuition why contracts reward experimenters only for extreme good outcomes has two elements whose main instrument is the probability of adopting the risky project given the observed return. The first element is a likelihood ratio that unravels the moral hazard concerns. It suggests that we should reward the agent when the probability of choosing the risky project was greater when effort was exerted rather than when no information was acquired. Since such probability is increasing in the risky return when effort was exerted and constant when that was not the case, the moral hazard incentives must be monotone increasing.

There is a second component associated with the adverse selection problem and is summarized by a hazard rate. The principal must also provide incentives for the agent to choose the risky project whenever he observes a sufficiently high signal. Since signals are ordered and wages are monotone, there will exist a unique cutoff signal where an agent is indifferent between the safe and the risky project. Higher signals will induce the choice of the risky project. This implies that, in the limit, the principal wants to penalize agents that chose the risky project when the cutoff signal was observed. Given the ordering of the signals, the probability of being at the cutoff signal given that the risky project was chosen (the hazard rate) is decreasing in the risky return. Therefore, this effect also suggests that compensation should be increasing in risky return.

The intuition for the project choice distortion relies on the idea that the principal has two mechanisms to induce the revelation of the unknown return. The first one is through risky wages as discussed before. The second one is through the safe wage, which is equivalent to picking the cutoff signal. Instead of paying the agent more to reveal the uncertainty imperfectly (the signal), the principal can also decrease the safe wage to induce the choice of the risky project and have the uncertainty revealed directly.

The model applies to a variety of situations. The one principal one agent problem is tightly connected to the optimal compensation of CEOs who must be motivated by the shareholders to undertake risky projects that could potentially lead to higher returns. In this scenario, the effort exerted by the CEO in learning about the portfolio of projects and the learned information is usually never observed by the shareholders, only the project chosen and the realized returns are observed. In this environment, the optimal contract can be implemented using restricted stocks conditional on a performance threshold. When the optimal contract is constrained to be monotone as in Innes (1990), the optimal monotone contract is an option with strike price greater than the return of the safe project.

The conclusions derived from the model explain why options and restricted stock are so widely used in this context. It also suggests that it is the best form of compensation to align the interests of shareholders and CEOs. Moreover, the optimal distortion on the project choice suggests that it is on the shareholders interest to have a CEO taking excessive risks. Letting a CEO pursue risky projects that are not ex-ante profitable would help the shareholders ameliorate the information problems.

The setup is also similar to the problem faced by managers who must encourage innovation among her employees to increase the profits of the firm. Workers have to divide their time between undertaking known tasks or exploring new ideas. However, innovations are risky ventures with a high probability of failure, thus agents prefer to put more effort on known tasks which returns are well known. Again issuing options and restricted stock to workers are shown to be useful to encourage more risk-taking.

The problem also resembles the technology adoption decision faced by farmers in developing countries. Foster and Rosenzweig (1995) studied the adoption of HYVs in India and found that imperfect knowledge about the management of the new seeds was a significant barrier to adoption. Hence, the rate of adoption was much slower than the desired one. In this context the amount of trials performed by farmers as well as the information gathered by them are usually unobservable to the social planner. Pricing policies resembling the optimal contract would encourage the adoption of new technologies.

As a last example, the proposed framework also fits the situation of industries with high levels of innovation such as pharmaceuticals. Pharmaceuticals must invest in potential drugs with unknown effectiveness. In the absence of property rights, free-riding reduces innovation and the potential discovery of new drugs. This environment is close to the one studied in this paper since the investment chosen by firms in the earlier stage of research is usually not disclosed, nor are the results from such investments. Given these constraints, the optimal contract suggests the use of patents to motivate innovation. However, such patents must be issued only for breakthroughs and not marginal innovations.

1.1 Literature Review

Models of delegated expertise were first proposed by Lambert (1986) and Demski and Sappington (1987), who used a simplified environment with two or three possible outcomes. Similar models where later developed as in Feess and Walzl (2004) and Gromb and Martimort (2007). In contrast with these papers, I allow for a continuum of outcomes, which permits a more complete characterization of optimal contracts. The closest structure to my model is the one studied by Malcomson (2009) who focuses in optimal distortions of the final decision as a mechanism to encourage more information acquisition. This paper, on the other hand, characterizes the optimal contract as the main incentive mechanism.[1]

The information structure used in this paper is similar to the one used in Szalay (2009) and Persico (2000). However, the first model is used in a procurement environment where the acquired information is induced to be completely revealed, this is not the case in this paper. In the second one the agents acquire information to learn about their value for an object, not the value for a principal as in our model. A similar information structure is also used in Moscarini and Smith (2001). Nevertheless, their model focus in the optimal actions of a single decision maker when information can be acquired over time, and not in the strategic interaction.

The acquisition of information is also related to bandit problems where an agent can learn about the return of a project by undertaking it as in Manso (2011) and Ederer (2008); however the information structure is more general in our case and contracts do not depend on the realized signal. Bonatti and Horner (2011) study moral hazard in teams over time where the return of a project is unknown and effort

[1]Chade and Kovrijnykh (2011) also study the motivation of information acquisition, however in their setup the acquired information is observable.

determines the rate of arrival of the return. Our setup is different in that individuals invest one time on a signal before deciding to undertake the risky project.

The next section introduces a principal agent setup with a single agent acquiring private information about a single unknown risky return to give intuition about the optimal contracts on a simplified framework. The third section solves for the first best, while the fourth section solves for the optimal contract when moral hazard and hidden knowledge exist. The next section discusses how to implement the contract in different real-world applications, and in the last section I conclude.

2 Basic environment

Consider the case of a risk neutral principal (she) who has to decide between a safe project with known returns and a risky project with unknown returns. The principal can hire a risk neutral agent (he) to acquire information and recommend him which project to pursue.[2] The information gathered by the agent (if any) is unobservable to the principal. Hence, this is a problem that involves hidden actions since the acquisition of information is not observable to the principal, but also there is a hidden information problem since the realized signal when the agent decides to acquire information is also not observed by the principal. The only observable variable is the final return. Therefore, the optimal contract designed by the principal must induce effort and the (partial) revelation of the information. It is also assumed that both individuals have limited liability.

There are two available projects that cannot be pursued simultaneously. There is a safe project with known net return $y_s \in (0, 1)$. There is also a risky one whose return $y_r \in [0, 1]$ is unknown. Let both individuals have the same nondegenerate prior belief $g(y_r)$ over the unknown return with finite mean μ_0.

Before taking the decision, the agent can exert effort and generate information about the risky project by acquiring a signal $x \in \mathbb{R}$ at a cost c. This cost can be

[2] The individual can be in fact risk averse or risk lover, just let the returns perceived by the agent be measured in utils and let the agent maximize a Von Neuman-Morgenstern utility function

associated with the cost of running trials or the disutility of effort. In a context of bandit problems, where signals are the same returns of the risky project, the fixed cost can also be thought as the ex ante expected return of the risky project, μ_0, and reflects the fact that individuals are initially pessimistic about it. The agent can also shirk, in which case no signal is generated and there will be no cost.

Let the conditional pdf and cdf of the signal x be denoted by $f(x|y_r)$ and $F(x|y_r)$, respectively. Assume these are differentiable with respect to x. Similarly, the unconditional pdf and cdf will be denoted by $f(x)$ and $F(x)$, respectively. Let the signals be ordered in the likelihood ratio sense: a signal x is more favorable than signal x' if the posterior distribution $g(y_r|x)$ first order stochastically dominates the posterior distribution $g(y_r|x')$.[3] These type of problems are known as monotone ones and were first studied by Karlin and Rubin (1956).

Since signals are ordered, the posterior mean will be a monotone increasing transformation of the signal. Thus the distribution of the posterior mean will be a transformation of the distribution of the signal. Without loss of generalization, let $x = \mathbb{E}_{y_r}[y_r|x]$ be the posterior mean of the risky project. Hence, the support of x is $[0,1]$.[4]

The acquisition of the signal and its content are privately known by the agent. The only observable variable for the principal is the final return of the chosen project y_s or y_r. Thus the principal must provide incentives to the agent to exert effort and reveal the hidden information. The optimal contract $w(\hat{x}, y)$ is then a function of the observed returns y and the reported signal $\hat{x}$. It will also be assumed that individuals have a limited liability constraint. The wage for the agent cannot be lower than 0 and the principal cannot pay more than the return he receives. Formally, optimal wages must satisfy $0 \leq w(\hat{x}, y) \leq y$ for any $\hat{x}, y \in [0,1]$. Given that a project j is chosen and a signal has been acquired, the payoff for the principal is given by $y_j - w(\hat{x}, y_j)$ and the payoff for the agent is $w(\hat{x}, y_j) - c$.

[3] Equivalently, $f(x|y_r)$ is log supermodular

[4] To have an interesting problem we need that $F(y_s) > 0$, otherwise it is optimal to always choose the risky project.

The game consists of two stages. In the first stage the principal designs a payment schedule and makes a take it or leave it offer to the agent. The agent accepts or rejects the contract. If she accepts the contract, she decides whether to acquire or not a signal, which is privately observed by her. In the second stage, the agent will report a signal $\hat{x}$, the principal will update her beliefs and choose which project to pursue. Finally a return y is realized and the principal pays to the agent the contracted wage $w\left(\hat{x}, y\right)$.

3 First Best

As a benchmark we will first focus on the case where there are no information assymetries. We will assume that the effort exerted by the agent to acquire a signal is observed, as well as the content of the acquired information. Because of the risk neutrality of both individuals, there will be multiple wage schedules that implement the first best. For simplicity, we also assume there are no limited liability constraints.

Since the acquired information by the agent is observed by the principal we have that $w\left(\hat{x}, y\right) = w\left(x, y\right)$. If the principal decides to hire the agent, she will face the following problem:

$$\max_{w(x,y_r),w(x,y_s)} \mathbb{E}_x \left[\max_{j(x)\in\{s,r\}} \mathbb{E}_{y_{j(x)}} \left[y_{j(x)} - w\left(x, y_{j(x)}\right) |x \right] \right]$$

$$\text{s.t } \mathbb{E}_x \left[\mathbb{E}_{y_{j(x)}} \left[w\left(x, y_{j(x)}\right) |x \right] \right] - c \geq \underline{u}$$

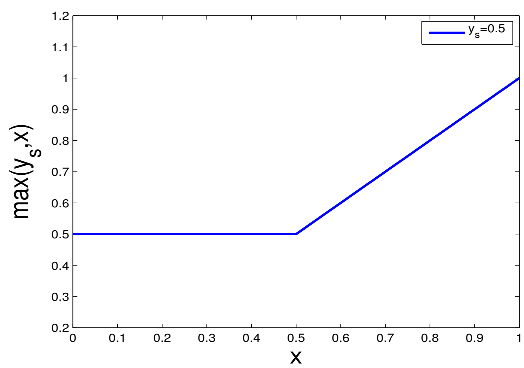
where $j\left(x\right)$ is the project chosen when x is realized. The constraint assures the agent will accept the contract by making sure his expected utility is greater than or equal to an outside option $\underline{u}$. The first best can be obtained by either a constant payment from the principal to the agent equal to the cost, or by selling the agent the returns from the project. Either of the alternatives lead us to the following indirect utility when the agent is hired:

$$\mathbb{E}_x\left[\max_{j(x)\in\{s,r\}}\mathbb{E}_{y_{j(x)}}\left[y_{j(x)}|x\right]\right]-c-\underline{u}$$

Since there are two stages, we proceed to solve the individual's problem using backward induction. That is, I will first determine which project is going to be chosen given the information acquired. Then, I will characterize when the principal decides to hire an agent as a function of y_s and c.

The individual will choose the risky project if $x > y_s$, thus the payoff of the second period is given by $\max\{x; y_s\}$. Note this is a convex function of x.

Figure 1: **Utility in second period**



The value of experimentation is defined as the ex ante expectation of the utility in the second period, that is

$$U(y_s) = \mathbb{E}_x[\max\{x; y_s\}]$$

From the previous properties we can prove the following lemma:

Lemma 1. *The value of experimentation $U(y_s)$ is strictly increasing and convex in y_s, and strictly greater than* $\max\{\mu_0; y_s\}$

The result states that information is always valuable in this setup. However, since information is costly, the principal may not want to hire the agent to acquire information. At the beginning of the first period a principal will choose to hire an agent if and only if $c \leq \hat{c}$, where $\hat{c}$ is defined by

$$U(\mu_0) - \hat{c} - \underline{u} = \mu_0 \tag{1}$$

Let the maximized objective function of the principal be denoted by $V(y_s, c) = \max\{U(y_s) - c - \underline{u}; y_s; \mu_0\}$. This function is also strictly increasing and convex in y_s. The next proposition characterizes when the principal decides to hire an agent as a function of y_s.

Lemma 2. *The principal decides to hire the agent when* $y_s \in (a_c, b_c) \subseteq (0, 1)$*, where* $\mu_0 \in (a_c, b_c)$*. Moreover, such interval is decreasing in* c*, that is* $(a_c, b_c) \subset (a_{c'}, b_{c'})$ *for any* $c < c' < \hat{c}$*, with* $(a_0, b_0) = (0, 1)$ *and* $a_{\hat{c}} = \mu_0 = b_{\hat{c}}$*.*

Even if beliefs are relatively pessimistic the individual decides to acquire information because of the potential gain represented by the value of experimentation. The lower is the fixed cost c, the greater is the interval over which the principal decides to hire the agent. Furthermore, if there is no fixed cost, the principal will always decide to hire an agent to collect information.

4 Constrained Efficiency

Now suppose the principal does not observe the effort of the agent, nor the information gathered by the individual. Also assume individuals have limited liability as described before. In this context a fixed wage will not induce any effort from the agent. Therefore the principal must provide incentives to the agent by imposing more risk in her payoff, and by distorting the choice of the best project. This distortion indicates that the first best will not be attained.

The constrained efficient problem for a principal who decides to hire an agent is the following:

$$\max_{w(y_r),w(y_s)} \mathbb{E}_x \left[\max_{j(x)\in\{s,r\}} \mathbb{E}_{y_{j(x)}} \left[y_{j(x)} - w\left(\hat{x}, y_{j(x)}\right) |x \right] \right] \tag{P}$$

subject to

$$\mathbb{E}_x \left[\mathbb{E}_{y_{j(x)}} \left[w\left(\hat{x}, y_{j(x)}\right) |x \right] \right] - c \geq \underline{u} \tag{IR}$$

$$\mathbb{E}_x \left[\mathbb{E}_y \left[w\left(\hat{x}, y_{j(x)}\right) |x \right] \right] - c \geq \max_{j\in\{s,r\}} \left\{ \mathbb{E}_y \left[w\left(\hat{x}, y_j\right) \right] \right\} \tag{IC1}$$

$$x = \arg\max_{\hat{x}\in[0,1]} \mathbb{E}_{y_{j(x)}} \left[w\left(\hat{x}, y_{j(x)}\right) |x \right] \text{ for all } x \tag{IC2}$$

$$0 \leq w\left(\hat{x}, y_j\right) \leq y_j \text{ for all } \hat{x} \text{ and } j = r, s \tag{LL}$$

Equation (IR) is the same individual rationality constraint as before. Equation (IC1) is the incentive compatibility constraint that ensures the agent will exert effort. It states that the agent's utility when he exerts effort is greater than the expected utility when he does not, in which case he picks the project that gives her the greatest ex-ante expected wage. Equation (IC2) is another incentive compatibility constraint to make sure the agent reports truthfully the information he gathered. The last equation represents the limited liability constraint.

This problem is hard to solve because the second incentive compatibility constraint involves a continuum of restrictions for each possible realization of the signal. Nevertheless, it is shown in the next lemma that optimal wages are not a function of the reported signal, they are just a function of the chosen project and the observed return. Thus the problem is simplified to one where the principal must provide incentives to the agent to choose her desired project.

Lemma 3. *The optimal wage schedule $w\left(\hat{x}, y_j\right)$ is not a function of the reported $\hat{x}$ and thus can be simplified to $w\left(y_r\right)$ and $w\left(y_s\right) = w_s$*

The intuition behind this result is different if the safe or the risky project is undertaken. If the safe project is chosen then the expected wage would just be a function of the reported signal, not on the realized one. Hence there is no way to induce the truthful revelation of the information since the agent will always choose to report the most convenient signal regardless of the realization.

On the other hand, if the risky project is chosen, having a wage as a function of the signal induces more informational rents for the agent without any benefit to the principal. Since the interest of the principal is whether to take the risky or the safe project, she is only willing to learn which project the signal suggests to take, not the exact realization of the signal. Therefore she prefers to have a contract that is only contingent on the chosen project but not on the realized signal. Therefore constraint (IC2) can be replaced by the following constraint:

$$\mathbb{E}_{y_{j(x)}}\left[w\left(y_{j(x)}\right)|x\right] \geq \mathbb{E}_{y_{-j(x)}}\left[w\left(y_{-j(x)}\right)|x\right] \text{ for all } x, \tag{2}$$

where $-j(x)$ denotes the project not chosen by the principal. In other words, the expected wage perceived by the agent must be greater when he follows the suggestion made by the principal.

Equation (2) can be further reduced to only one constraint when optimal wages for the risky project are monotone nondecreasing. Since signals are ordered, a posterior generated by a signal x first order stochastically dominates any posterior generated by any less favorable signal x'. Therefore a less favorable signal implies that the expected wage is lower.

Since distributions are continuous in x, there must exists a cutoff x_e such that the expected wage when the risky project is chosen given such signal is equal to the safe wage. Any more (less) favorable signal than x_e implies the agent will choose the risky (safe) project and that the constraint will not be binding. Formally, there exists cutoff x_e such that:

$$\int_0^1 w(y_r)\, g(y_r|x_e)\, dy_r = w_s$$

Using this fact, the ex-ante agent's utility can be rewritten as:

$$\int_{x_e}^{1}\int_{0}^{1} w\left(y_r\right) f\left(x|y_r\right) g\left(y_r\right) dy_r + w_s F\left(x_e\right) - c$$

By integrating the possible signal values when they are higher than the cutoff we obtain the relevant object to provide incentives: $1-F\left(x_e|y_r\right)$. This is the probability that the agent chooses the risky project when he exerts effort. Let us for now assume that wages are nondecresing, and in the next proposition I will prove that is the case. The problem can be reexpressed as:

$$\max_{w(y_r),w(y_s)} \int_{\underline{y}}^{1} \left(y_r - w\left(y_r\right)\right)\left(1 - F\left(x_e|y_r\right)\right) g\left(y_r\right) dy_r + \left(y_s - w_s\right) F\left(x_e\right) \tag{P*}$$

subject to

$$\int_{0}^{1} w\left(y_r\right)\left(1 - F\left(x_e|y_r\right)\right) g\left(y_r\right) dy_r + w_s F\left(x_e\right) - c \geq \underline{u} \tag{IR*}$$

$$\int_{0}^{1} w\left(y_r\right)\left(1 - F\left(x_e|y_r\right)\right) g\left(y_r\right) dy_r + w_s F\left(x_e\right) - c \geq \max\left\{\mathbb{E}_y\left[w\left(y_r\right)\right], w_s\right\} \tag{IC1*}$$

$$\int_{0}^{1} w\left(y_r\right) \frac{f\left(x_e|y_r\right) g\left(y_r\right)}{f\left(x_e\right)} dy_r = w_s \tag{IC2*}$$

$$0 \leq w\left(y_j\right) \leq y_j, \quad \text{for } j = r, s \tag{LL}$$

Let λ, δ_r, δ_s, and ϕ be the Lagrange multipliers for the first three constraints. The distinction between δ_r and δ_s arises because the expected wage from the risky project does not need to be necessarily equal to the safe wage. Since the problem is linear on the wages, the optimal wages are determined by a bang-bang solution that is bounded by the limited liability constraint. After rearranging the derivative with respect to wages we obtain

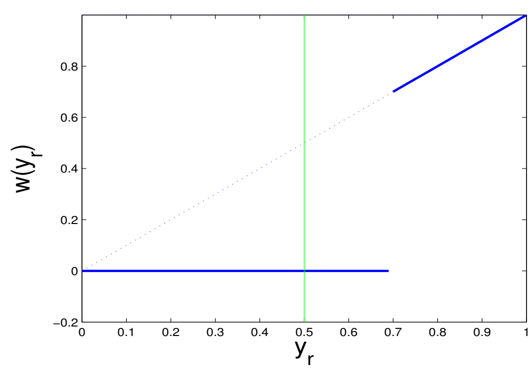
$$-1 + \lambda + \delta_r\left(1 - \frac{1}{1 - F\left(x_e|y_r\right)} + \frac{\delta_s}{\delta_r}\right) - \phi \frac{f\left(x_e|y_r\right)}{\left(1 - F\left(x_e|y_r\right)\right) f\left(x_e\right)} \tag{3}$$

Whenever this condition is positive, the wage will be set to the upper bound; if it is negative, then the optimal wage is zero. Using the structure of the signals and the latter equation we can indeed prove the that wages are nondecreasing in the following proposition. Figure 2 illustrates the contract.

Proposition 4. *If $f(x|y_r)$ satisfies the likelihood ratio order then the optimal wage schedule $w(y_r)$ is monotone and is characterized by a cutoff z such that*

$$w(y_r) = \begin{cases} y_r & \text{if } y_r \geq z, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Figure 2: **Optimal Contract**



The cutoff z is given by the value of y_r such that condition (3) is equal to zero. Wages are thus monotone if for any greater (lower) y_r the derivative is positive (negative). Thus, showing that the derivative is increasing in y_r implies that wages are monotone.

Condition (3) has four elements. The first one is the marginal cost to the principal of increasing the wage. The second one is the benefit from relaxing the IR constraint which might be 0 if such constraint is not binding, that is if $\underline{u} < w_s$. This case arises when the agent gets information rents for the private information.

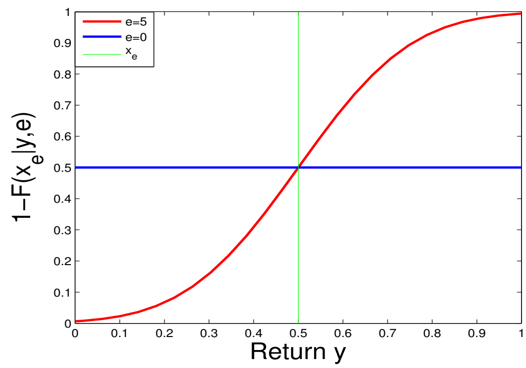
As it is common in moral hazard problems, the third term is related to a likelihood ratio. It is related to the probability of undertaking the risky project when no information has been acquired relative to that same probability when the information has been acquired. Whereas the first one is independent of y_r, the second one is increasing in y_r since the likelihood ratio order induces first order stochastic dominance.

There are three possible cases. If w_s is greater than $\mathbb{E}_y [w(y_r)]$, implying that $\delta_r = 0$, then it would become obvious that the agent exerted effort whenever the risky project is chosen. In this case the problem simplifies to one of just hidden information where incentives must be given to encourage the adoption of the risky project. On the other hand, if $w_s < \mathbb{E}_y [w(y_r)]$, then an agent choosing the safe project would suggest the acquisition of information. Thus, there will always be a punishment for choosing the risky project, but such punishment will decrease the greater is the return. When y_r approaches to 1, there will be no punishment.

When $w_s = \mathbb{E}_y [w(y_r)]$, the principal is minimizing the cost of inducing acquisition of information. To see this note the similar structure of restriction (IC1) with equation (1) and its subsequent properties. In this case the agent will be indifferent between the safe and the risky project when he does not acquire information and it will appear as if he randomizes his decision and chooses the risky project with probability $\frac{\delta_r}{\delta_r+\delta_s}$. Then the principal can reward the agent for risky returns that induced a greater probability of choosing the risky project when the information was acquired, otherwise the agent will be punished. As argued before, such likelihood ratio will be monotone increasing as it is shown in Figure 3.

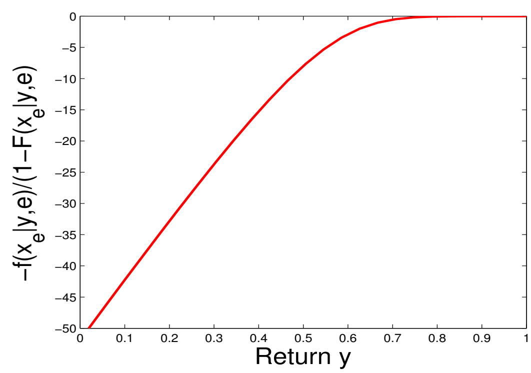
The final component is associated to the hidden information problem and is represented by the hazard rate $\frac{f(x_c|y_r,e)}{1-F(x|y_r,e)}$. The hazard rate in this context is interpreted as the probability of having observed the cutoff signal given that the risky project was chosen. Intuitively, the principal wants to make sure the right project is chosen, which implies that, in the limit, she wants to penalize agents who chose the risky project when they just observed the cutoff signal. Given the MLRP condition of the signals with respect to the return, this hazard ratio is monotone decreasing with respect to the return (see Athey (2002) and Figure 4 for an example). In other words, a lower observed return increases the chances that the agent observed the cutoff signal.

Figure 3: **Likelihood Ratio**



Therefore, this effect also suggests that compensation should be increasing in y_r.

Figure 4: **Negative Hazard Rate**



In sum, a monotone contract encourages information acquisition and the partial revelation of information. On one hand, paying more for higher outcomes induces the acquisition of information since this increases the probability of choosing the best project. On the other hand, paying more for higher outcomes induces the agent to choose the best project once the information has been revealed since a higher signal is associated to a higher return. Thus the optimal wage schedule creates the incentive for the agent to adopt the risky technology after observing favorable signals.

The safe wage is set according to constraint (IC2∗). Formally, choosing w_s is equivalent to choosing x_e. The first order condition with respect to x_e is simplified to:

$$y_s - x_e - \phi \frac{\partial \mathbb{E}\left[w\left(y_r\right)|x_e\right]}{\partial x_e} = 0$$

Raising the threshold will yield a marginal benefit of y_s since the safe project will be pursued more often. However, it also generates a cost of $x_e = \mathbb{E}\left[y_r|x_e\right]$, the return of the risky project at the threshold, and a cost of increasing the expected wage when such threshold increases. The last cost arises because the safe wage must increase to generate the appropriate incentives to the agent to choose the desired project. [5]

This condition implies that the optimal decision is always distorted in the constrained efficient solution in favor of the risky project, $x_e < y_s$. This distortion is purely generated by the hidden information problem. Note that constraints (IR$*$) and (IC1$*$) remained constant after a change on x_e precisely because of constraint (IC2$*$). Since it is costly for the principal to induce the agent to reveal the information through the risky wage, she favors the decision that reveals the uncertain state.

This distortion also suggests that the first best is never attained. To avoid the distortion the risky wage should be constant. Although the hidden information would be solved in this case, the moral hazard will persist since information will have no value to the agent. The first best is also never attained because the principal will hire the agent only for c strictly lower than $\hat{c}$, where $\hat{c}$ was defined in equation (1). To see this just note that the value of experimentation decreases for the principal since she will only appropriate risky returns lower than z, thus her payoff is no longer convex.

Finally note that z must be greater than y_s. If this is not the case then the principal will never want to hire the agent since she is better off by pursuing the

[5]Remember that wages are monotone increasing, implying that a higher threshold generates a higher expected wage.

safe project. This result suggests that the agent should be rewarded for pursuing the risky project only when its return is significantly better than the safe return, not just marginally better.

4.1 Debt Contracts

The optimal contract found in the previous section is not continuous. In particular, the payment for the principal is not monotone since any return greater than the threshold will yield him zero profit. As argued by Innes (1990), this type of contracts could be manipulated by either the principal or the agent if any of them can affect the return before the contract is paid. For example, the principal would have incentive to sabotage the risky project by burning profits in excess of the threshold. Similarly, the agent would have an incentive to inflate profits by borrowing money and "revealing" a higher apparent profit to the principal.

In order to prevent this behavior, a monotonicity constraint must be imposed, thus modifying the limited liability constraint:

$$w\left(y_r-\epsilon\right)\leq w\left(y_r\right)\leq w\left(y-\epsilon\right)+\epsilon$$

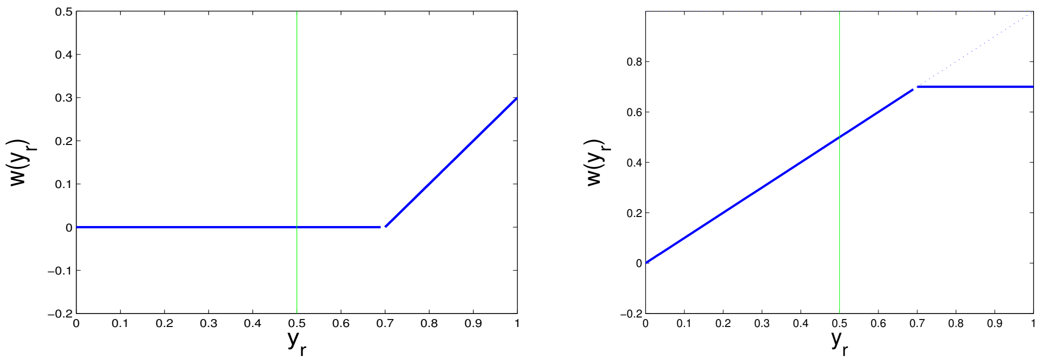
The same argument in the preceding section applies, and the optimal monotone contract will be option like, with an strike price z_0 greater than the safe return:

$$w\left(y_r\right)=\max\left\{0,y_r-z_0\right\} \tag{5}$$

This type of contract was first obtained by Innes (1990) using a principal-agent setup. In his paper the argument relies on the standard assumption that a greater effort generates better distributions of profits in the likelihood ratio sense. Therefore an optimal contract rewards the agent for higher profits and punishes him otherwise.

The difference with our setup is that effort does not lead to such ordering on the distribution of returns. In fact, effort does not directly influence the observed returns.

Figure 5: **Optimal Monotone Contract**



(a) Payment to Agent (b) Payment to Principal

The returns for either project are given ex-ante, the problem is that individuals are uncertain about the risky return. Effort in our context lets the agent take a more informed decision and that is why the agent is rewarded for good outcomes.

5 Implementation

The studied setup closely resembles the interaction between a CEO and the shareholders of a firm. The shareholders hire a CEO to take decisions concerning the future of the company. It is usually the case that the CEO has more expertise than the shareholders in making such decisions, or at least is more efficient at gathering information related to such decisions. Decisions can range from acquisitions to the marketing of new products, whose main characteristic is the uncertainty of their return. Such uncertain returns will be reflected in the value of the firm and thus affect the shareholders payoff. The effort exerted by the CEO to acquire information is not observable. It is also common that if information has been acquired by the CEO, it is not (completely) observed by the shareholders.

The optimal contract derived here suggests that stock options and profit sharing compensation are optimal ways to motivate the acquisition of information and its revelation. Stock option programs give workers the right to buy company's shares at a fixed price for a given period of time. These will only be exercised if the market price is higher than the strike price originally agreed to. Stock options are thus used as a long- term motivator and the employee is constrained on exercising the option only after their performance has been (partially) observed. Likewise, profit sharing is also used as a long-term motivator where individuals are entitled to a percentage of the profits of a firm after a given period. To implement the contract the firm could set a threshold on the profits such that the CEO can only claim his share if profits are greater than such level.

The constrained efficient solution also explains why it is optimal to pursue risky projects even when the information suggests the safe project is better. This feature is in close connection to the debate on excessive risk taking of CEOs. Although this is inefficient ex-ante, it is a valuable strategy because it helps to reveal the uncertain state of nature. This is also related to the result obtained by Manso (2011) where early failure is not punished and long-term success is rewarded.

The environment also describes situations that involve technology adoption. The return of new technologies is usually uncertain until they are tried. For example, it has been documented the lack of adoption of high yielding variety seeds in developing countries because of the uncertainty that farmers must bear (Foster and Rosenzweig, 1995). A social planner interested on the adoption of new technologies but unable to observe the effort (or number of trials) performed by the farmers and their subsequent results, could use an option-type contract to encourage farmers.

The optimal contract might be also interpreted as a patent policy to encourage innovation. Innovation is usually thought as a process whose return is proportional to the amount of R&D expenditures. However, one could think of such expenditures as being a measure of the precision of the signals that emerge from the trials. If a great quantity of money is invested and the trials suggest that the new product would not work, a good decision is to stop pursuing such project. On the other hand, if not enough money has been invested, a bad trial does not necessarily imply is not a good project.

Following such interpretation, a social planner who wants to increase the levels of innovation in a society but cannot observe R&D expenditures and their outcomes, should use the optimal contract derived here to provide incentives for innovation. The contract can be interpreted in terms of a patent. It suggests that patents should only be given if it is shown that the new technology is significantly better than the previous one, and not for marginal improvements. This result cannot be interpreted as a restriction on the use of the new technology as often happens with patents. In other words, the optimal contract does not allocate the property rights of the new technology. On the contrary, it encourages the adoption of the new technology by all the population, while rewarding innovators with the surplus they generated, suggesting an optimal pricing policy.

6 Conclusions

This paper studies the problem faced by a risk-neutral principal who must hire a risk-neutral agent to take a decision between a safe project with known return and a risky one with unknown return. The agent can acquire costly information about the new project before taking the decision of whether to pursue it or not. The acquisition of information and its content are unobservable to the principal, thus the problem features moral hazard and hidden information frictions.

The main result is that the wage schedule when the risky project is pursued must be monotone increasing. The agent will be rewarded with the whole output only if the return is significantly better than the safe return, otherwise he will be paid nothing. Thus the risk of the decision is imposed to the agent. It has been also shown that decision making is distorted in favor of the risky project. That is, it is optimal to pursue the risky project when a signal reveals that is slightly worse than the safe one.

The optimal contract resembles restricted stock in an environment where shareholders hire a CEO to make decisions. In this case the CEO will be given firm shares when he decides to undertake the risky project and its returns are significantly greater than the returns from the previous project. When the optimal contract is constrained to be continuous it resembles an option with a strike price greater than the return of the safe project. The model also provides explanation as to why excessive risk taking is optimal, where excessive risk taking is understood as pursuing the risky project even when it seems to be worse than the safe one.

The conclusions of the model are robust when there is a finite number of possible projects. However, the same framework does not apply when the decision belongs to a continuum as in Malcomson (2009). In this case the optimal contract is analogous to one where the gathered information is completely revealed. This is a natural extension of the proposed model that we leave for further study. Another useful extension is the characterization of the optimal contract when the agent is risk-averse as well as in a dynamic setup where reputation for undertaking good projects can be built.

References

S. Athey. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187, 2002.

A. Bonatti and J. Horner. Collaborating. *The American Economic Review*, 101(2): 632–663, 2011.

H. Chade and N. Kovrijnykh. Information acquisition, moral hazard, and rewarding for bad news. *Working Paper*, 2011.

J.S. Demski and D.E.M. Sappington. Delegated expertise. *Journal of Accounting Research*, pages 68–89, 1987.

Z. Dong, C. Wang, and F. Xie. Do executive stock options induce excessive risk taking? *Journal of Banking & Finance*, 34(10):2518–2529, 2010.

F. Ederer. *Launching a Thousand Ships: Incentives for Parallel Innovation*. Working Paper, 2008.

E. Feess and M. Walzl. Delegated expertise–when are good projects bad news? *Economics Letters*, 82(1):77–82, 2004.

A.D. Foster and M.R. Rosenzweig. Learning by doing and learning from others: Human capital and technical change in agriculture. *The Journal of Political Economy*, 103(6):1176–1209, 1995.

D. Gromb and D. Martimort. Collusion and the organization of delegated expertise. *Journal of Economic Theory*, 137(1):271–299, 2007.

R.D. Innes. Limited liability and incentive contracting with ex-ante action choices* 1. *Journal of economic theory*, 52(1):45–67, 1990.

S. Karlin and H. Rubin. The theory of decision procedures for distributions with monotone likelihood ratio. *The Annals of Mathematical Statistics*, 27(2):272–299, 1956.

R.A. Lambert. Executive effort and selection of risky projects. *The Rand Journal of Economics*, pages 77–88, 1986.

E.L. Lehmann. Comparing location experiments. *The Annals of Statistics*, pages 521–533, 1988.

J.M. Malcomson. Principal and expert agent. *The BE Journal of Theoretical Economics*, 9(1):17, 2009.

G. Manso. Motivating innovation. *The Journal of Finance*, 66(5):1823–1860, 2011.

P.R. Milgrom. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391, 1981.

G. Moscarini and L. Smith. The optimal level of experimentation. *Econometrica*, 69 (6):1629–1644, 2001.

N. Persico. Information acquisition in auctions. *Econometrica*, 68(1):135–148, 2000.

D. Szalay. Contracts with endogenous information. *Games and Economic Behavior*, 65(2):586–625, 2009.

A Appendix

Proof for Lemma 1. To show that $U(y_s) > \max\{\mu_0, y_s\}$ integrate by parts the value of experimentation to obtain

$$\begin{aligned} U(y_s) &= y_s + \int_{y_s}^{1} (1 - F(x))\,dx \\ &= y_s + \int_{0}^{1} (1 - F(x))\,dx - \int_{0}^{y_s} (1 - F(x))\,dx \\ &= \mu_0 + \int_{0}^{y_s} F(x)\,dx \end{aligned} \tag{6}$$

Also note that

$$\frac{\partial U(y_s)}{\partial y_s} = F(y_s) > 0$$

where the strict inequality comes from footnote 4. Convexity is also easily obtained since the second derivative is the probability of having a signal equal to y_s.

□

Proof for Lemma 2. First note that $U(0) = \mu_0$ and $U(1) = y_s$. Therefore $U(y_s) - c$ will cross at most once each of the outside options. It could cross once the constant μ_0 from below since it is increasing in y_s. It could cross once y_s from above since its first derivative with respect to y_s is between 0 and 1. This in turn implies that $U(y_s)$ is farther from $\max\{\mu_0; y_s\}$ precisely when $\mu_0 = y_s$.

Since $U(y_s) - c$ is linear in c, there exists a $\hat{c}$ such that $U(y_s) - c = \mu_0 = y_s$. Thus, for any $c < \hat{c}$, there exists $a_c, b_c \in (0, 1)$ such that $U(y_s) - c > \mu_0$ for any $y_s > a_c$, and $U(e^*) - C(e) > y_s$ for any $y_s < b_c$. Obviously it must be the case that $\mu_0 \in (a_c, b_c)$.

Note that a_c and b_c are increasing and decreasing in c, respectively, precisely because the function crosses from below and above each of the corresponding outside options. Finally, for any $c > \hat{c}$, the interval is empty and the principal never experiments.

□

Proof for Lemma 3. The principal is interested in obtaining a truthful report from the agent in order to take a decision between the safe and the risky project. Let the set of signals that lead the principal to choose the risky project be denoted by R and the corresponding set that lead her to choose the safe project be denoted by S. Lets consider first the case when $x \in S$ and suppose the paid wage is $w(\hat{x}, y_s)$. Since x is unrelated to y_s, the expected safe wage will not depend on x, only on $\hat{x}$. Thus the agent will always prefer to report a $\hat{x}$ that maximizes $w(\hat{x}, y_s)$, regardless of the observed x. Therefore, in order to have truthful revelation, the safe wage must not depend on $\hat{x}$ and can be expressed as w_s given that y_s is constant.

Now, consider the case where $x \in R$. Lets first consider any deviation from the agent such that $\hat{x} \in R$. Since $j(x) = r = j(\hat{x})$, the principal is only interested in paying the agent the least possible as long as he reports a signal that leads her to choose the risky project. Formally, the principal solves the following problem:

$$\min_{w(x,y_r)} \mathbb{E}\left[w(x, y_r) \,|x\right]$$

$$\text{subject to } \mathbb{E}\left[w(x, y_r) \,|x\right] \geq \mathbb{E}\left[w(\hat{x}, y_r) \,|x\right] \text{ for all } \hat{x}, x \in R$$

$$\text{and } \mathbb{E}\left[w(x, y_r) \,|x\right] \geq w_s \text{ for all } x \in R$$

The solution to this program is to set $w(x, y_r) = w(y_r)$, thus the first constraint will be always binding and the principal will not have to pay unnecessary informational rents. □