1 Introduction

This chapter investigates contract design when at least one of the players is not fully informed about some payoff relevant features of the contracting environment until the agreement has been selected, and signed or rejected. Uncertainty over the outcome of a contracting game can arise from two sources. First, the players might be differentially informed about the future outcome of the contract at the time it is written. Uncertainty that arises from this source is called hidden information. If their information is not verifiable in a settling up phase of the contract, there might be opportunities for informed players to misrepresent the situation to increase their payoffs. For example an repair firm might claim some damage is worse than it really is. Second, the contract might specify certain contractual obligations that are unobserved and cannot be directly verified by the other players. This creates a situation of moral hazard, where players might choose to neglect their contractual obligations with impunity. For example a person renting a car increases its wear and tear through aggressive driving.

2 Contracts for Goods and Services

Studying a model where there are only three choices and three outcomes provides a useful introduction to the three concepts of rent extraction, the participation constraint, and the incentive compatibility constraint. But the variety of contract types can only be appreciated by looking at richer models.

4 Internalizing Incentives

The costs of hidden information and moral hazard arise because resources are used to induce truth telling and efficient behavior while simultaneaously compensating all the parties to the contract sufficiently well so that they sign on. The revelation principal shows that in a world of rational players negotiations do not reduce these costs, because for every bargaining game there is a direct revelation game with the same solution outcomes.

These losses can, however, be eliminated if everyone foresees the dual problems of hidden infomration and moral hazard before they arise and contracting is conducted while all the players are symetrically informed.and can be enforced throughout

4.1 Coordinating activities

Profit centers. Divisions within large corporations are sometimes encouraged to pursue some clearly defined objectives that are mainly but not completely determined by their own actions. If the firm is well managed, selling the divisions would cause the total value of the firm to fall. Headquarters, some promotion actitives, aspects of human resources, and supply network are not solely divisional activities but have synergies and complementarity across several divisions of the firm. In this setting managers face two related questions. In implementing policies that have ramifications across several divisions how can central management accurately elicit preferences about the value of different policies from its divisions? Furthermore, when divisions themselves take actions that affect the goals of the others, how can central management align the incentives so that the division's goals are aligned with the firms?

Consider a firm with N profit centers, where the value of center $n \in \{1,\ldots,N\}$ depends on its own actions, x_n, as well as the actions of the others, denoted by $x_{-n} \equiv (x_1,\ldots,x_{n-1},x_{n+1},\ldots,x_N)$. We assume the value of the corporation, $v(x)$, is

$$v(x) = E\left[\sum_{n=0}^{N} \pi_n(x) + \varepsilon_n\right] = \sum_{n=0}^{N} \pi_n(x)$$

where $x \equiv (x_1,\ldots,x_N)$ is the vector of the actions by all the profit centers, $\pi_n(x)$ is the goal of the n^{th} profit center for $n \in \{1,\ldots,N\}$, and $\pi_0(x)$ is a compl factor that reflects the complementaries (or between the different centers that add value to the firm for any given vector of divisional choices x. We interpret $\pi_n(x) + \varepsilon_n$ as the returns from the n^{th} profit center if it operated as an independent corporate entity whose value is affected through the choices of profit centers stand alone.

The value of coordinating profit centers becomes apparent when we compare the value of the firm under an idealized centralize management versus a decentralized decision making made at the divisional level. Let $x^o \equiv (x_1^o,\ldots,x_N^o)$ denote the divisional choice vector that maximizes the firm's value. That is $v(x^o) \geq v(x)$ for all $x \in X$, the set of feasible choice vectors for the firm. We first note that if decison making in the firm was fully decentralized, then the n^{th} division would pick x_n to maximize the value of its profit center $\pi_n(x)$ that is given the choices of the other centers. We model this decentralized approach as a simultaneous move game, or as the strategic form of any noncooperative game. Solutions to the game are found using the techniques we developed in Part III of this book. Accordingly let $x^e \equiv (x_1^e,\ldots,x_N^e)$ denote a solution to this game. It is certainly a Nash equilibrium, defined by the property for each division $n \in \{1,\ldots,N\}$ that if every other division plays their part of the solution, x_{-n}^e, then division n choosing x_n^e maximizes the value of the division. that is for each $n \in \{1,\ldots,N\}$ and $x_n \in X_n$, the strategy x^e is a Nash equilibrium if

$$\pi_n(x^e) \geq \pi_n(x_n, x_{-n}^e)$$

There is no reason to suppose that $v(x^o) = v(x^e)$ since that would require

$$\pi_n(x^o) \geq \pi_n(x_n, x_{-n}^o)$$

for all $x_n \in X_n$ and $n \in \{1,\ldots,N\}$.

For example consider a multiproduct firm of interlocking divisions in which net revenue generated by each division mainly depends on its own inputs but also on inputs for other divisions. We suppose the net revenue generated by the n^{th} division is:

$$p_n(x) = \sum_{k=1}^{N} \theta_{nk} x_k - x_k^2$$

where the fixed coefficient θ_{nn} denotes its own inputs, and the other coefficients θ_{nk} for $k \neq n$ measure the effects of the other divisions on the profitability of the n^{th}. It is reasonable to assume $\theta_{nn} > 0$ for each $n \in \{1, \ldots, N\}$, but that θ_{nk} could either postive or negative for $k \neq n$, depending on the particular application. The contribution of x_n, the inputs from n^{th} division, to total firm value is thus:

$$\pi_n(x) = \sum_{k=1}^{N} \theta_{nk} x_n - x_n^2$$

The optimal divisional input vector x^o is found by maximizing the total benefits to the firm, which are defined as:

$$\sum_{n=1}^{N} \pi_n(x) = \sum_{n=1}^{N} \left(\sum_{k=1}^{N} \theta_{nk} x_n - x_n^2 \right)$$

This is a linear quadratic objective function with an interior solution derived form its first order condition

$$x_n^o = \frac{\sum_{k=1}^{N} \theta_{nk}}{2}$$

Substituting the optimal inputs into the net revenue function for the firm we obtain the total value of the firm:

$$v_0 = \frac{1}{4} \sum_{n=1}^{N} \left(\sum_{k=1}^{N} \theta_{nk} \right)^2$$

In a simultaneous move game, where each division maximizes the net benefits from its own profit center heedless of its effects on the other divisions, the n^{th} divisional manager maximizes $p_n(x)$with respect to x_n given the input choices made by the other divisions x_{-n}. In this game the dominant strategy for each division, derived from the first order condition of the division's optimization problem, is to set:

$$x_n^e = \frac{\theta_{nn}}{2}$$

The value of the firm is then

$$v_e = \frac{1}{4} \sum_{n=1}^{N} \left(\sum_{k=1}^{N} 2\theta_{nk}\theta_{nn} - \theta_{nn}^2 \right)$$

The equilibrium yields the same allocation as the optimum if and only if

$$\theta_{nn} - \sum_{k=1}^{N} \theta_{nk} = 0$$

for each $n \in \{1, \ldots, N\}$. Typically this condition is not met unless $\theta_{nk} = 0$ for all $k \neq n$, that is when the divisions have no interactions with each other. But if $\theta_{nk} = 0$ for all $k \neq n$ then the ratinale for combining the various divisions under common ownership is lost.

Suppose that if is not feasible for central management to dictate to each division the optimal action x_n^o, perhaps because it cannot be observed, or cannot be enforced effectively. Rather than command the division managers to choose x_n^o rather than x_n^e, headquarters could distort the incentives of the profit centers to align them with the goals of the firm. In particular suppose, the goal of division n is now to choose x_n to maximize:

$$\pi_n(x) + t_n(x)$$

where

$$t_n(x) = s(x_{-n}) + \sum_{k=1,k\neq n}^{N} \pi_k(x)$$

The distortion has two parts. The first expression, $s(x_{-n})$, is an arbitrary function that does not depend on what the n^{th} divisional manager chooses, included only to recenter the profit level of the division. If headquarters knows the optimal policy x^o but cannot enforce is, then a natural choice for $s(x_{-n})$ is the constant $-v(x^o)$, the negative of the maximal value of the firm under the optimal policy. The second expression is the sum of the profits of all the other divisions and is the key to realigning the divisions goals with those of the company. Substituting the formula for $t_n(x)$, under the proposed incentive scheme the n^{th} divisional manager maximizes

$$s(x_{-n}) + \sum_{k=1}^{N} \pi_k(x)$$

with respect to x_n. From the definition of an optimum choice x^o it follows for each $n \in \{1,\ldots,N\}$ that:

$$\pi_n(x_{-n}^o, x_n^o) + t_n(x_{-n}^o, x_n^o) \geq \pi_n(x_{-n}^o, x_n) + t_n(x_{-n}^o, x_n)$$

for all possible choices x_n. But this is a Nash equilibrium in a game where each divisional manager $n \in \{1,\ldots,N\}$ is given the distorted objective function for his operation and chooses x_n individually. In paricular every division will realize a loss unless they play the Nash equilibrium srategy profile of x^o when the tax for the n^{th} dvivison is set as:

$$t_n(x) = -v(x^o) + \sum_{k=1,k\neq n}^{N} \pi_k(x)$$

In the example we introduced above, this formula for $t_n(x)$ reduces to:

$$t_n(x) = \left(\sum_{k=1,k\neq n}^{N} \sum_{j=1}^{N} \theta_{kj} x_n - x_k^2\right) - \frac{1}{4}\sum_{n=1}^{N}\left(\sum_{k=1}^{N} \theta_{nk}\right)^2$$

4.2 Eliciting information

A similar situation arises when headquarters are contemplating changes that will affect all its branches (such as a new accounting system, or a marketing campaign that involves all its product lines), but have less knowledge than the branches about its effects on overall profitability. We now assume that each division has private information about how alternative policies might affect the profitability of its own branches. If central office had this information, it would choose amongst policies $z(\theta)$, where $\theta = (\theta_1,\ldots,\theta_N)$ is a vector of values representing all the information that is accessible to the divisions, where the value of θ_n is known only to division $n \in \{1,\ldots,N\}$. Analogous to above suppose the value of the firm is

$$v(z,\theta) = \sum_{n=0}^{N} \pi_n(z,\theta_n)$$

To interpret this equation, we should note that is not $v(z,\theta)$ directly observed by headquarters even expost. Otherwise

If central office ignored the information potentially available to them, they would presumably maximize the expected value of the firm taken over θ. We denote by z^u the optimal uninformed choice, satisfying the inequalities $E[v(z^u,\theta)] \geq E[v(z,\theta)]$ for all z. Let us suppose the value of θ is θ^* and let z^* maximize $v(z,\theta^*)$. By definition $v(z^*,\theta^*) \geq v(z^u,\theta^*)$, and typically this inequality is strict. There is little reason to suppose that every division within the firm prefers the informed choice over the uninformed choice. When the optimal informed choice for the firm is not unanimously supported by all the divisions, it is even possible that for one (or more) division(s) prefer the uniformed choice, that is $\pi_n(z^*,\theta_n^*) < \pi_n(z^u,\theta_n^*)$ for some $n \in \{1,\ldots,N\}$. In any event there is conflict amongst the profit centers about the optimal choice of z for the corporation whenever the optimal choice for the corporation is not the optimal choice for every division, that is whenever $\pi_n(z^*,\theta_n^*) < \pi_n(z^n,\theta_n^*)$ for some n and z^n, a common occurrence.

To elicit truthful information about the state of the firm's divisions, the incentives of each division must be aligned with those of the whole corporation. Instead of implicitly rewarding divisional heads with $\pi_n(z^u,\theta_n^*)$, headquarters asks each division to report its information θ_n^*, receives the report δ_n, and distorts the profits of the division with a transfer $t_n(\delta)$ defined as:

$$t_n(\delta) = s_n(\delta_{-n}) + \sum_{k=1,k\neq n}^{N} \pi_k(z,\delta_k)$$

This correction closely resembles the transfer distortion we proposed in the case where the choices made by the profit centers are voluntary. The first term, $s_n(\delta_{-n})$, does not depend on the announcement of division n, while the second expression aligns the incentives facing the division with the firm. Under this incentive scheme, the n^{th} profit center receives

$$\pi_n(z(\delta),\theta_n^*) + s_n(\delta_{-n}) + \sum_{k=1,k\neq n}^{N} \pi_k(z,\delta_k)$$

When $\theta = (\delta_{-n},\theta_n^*)$ the optimal choice for the corporation is $z(\delta_{-n},\theta_n^*)$ by definition. It now follows that truthful reporting is a weakly dominant strategy for all N divisions and for all values of the vector θ because:

$$\pi_n(z(\delta_{-n},\theta_n^*),\theta_n^*) + \sum_{k=1,k\neq n}^{N} \pi_k(z(\delta_{-n},\theta_n^*),\delta_k) \geq \pi_n(z(\delta),\theta_n^*) + \sum_{k=1,k\neq n}^{N} \pi_k(z(\delta),\delta_k)$$

An unattractive feature of this scheme is that the transfers between the divisions do not balance, so this scheme is not self financing.

For example consider a company producing related product lines (such as electric appliances or outdoor gear) that wishes to determine whether a proposed expansion into a new regional market would be profitable or not. We suppose the cost of undertaking the expansion is γ and that the value of the n^{th} division of the company would increase by θ_n^*. If headquarters were fully informed, then the expansion would proceed if and only if

$$\sum_{n=1}^{N} \theta_n^* > \gamma$$

If the company was decentralized, and could not easily aggregate the divisional values with the full cooperation of the divisional managers, one approach might be for headquarters to require each divisional manager n to submit a contribution to the expansion, denoted δ_n that would be paid only if aggregate contributions exceeded the cost. The payoff to the division from playing this game of incomplete information is

$$\lambda_n = \begin{cases} \theta_n^* - \delta_n \text{ if } \sum_{k=1}^{N} \delta_k > \gamma \\ 0 \text{ if } \sum_{k=1}^{N} \delta_k \leq \gamma \end{cases}$$

Each division reports δ_n to maximize its expected value, conditional on the response that the other divisions use as a function of their types

$$\pi_n(\delta_n, \theta_n^*) = \Pr\left[\sum_{k=1}^{N} \delta_k > \gamma | \delta_n\right](\theta_n^* - \delta_n)$$

Differentiating we see that

$$\pi_n'(\delta_n, \theta_n^*) = \Pr'\left[\sum_{k=1}^{N} \delta_k > \gamma | \delta_n\right](\theta_n^* - \delta_n) - \Pr\left[\sum_{k=1}^{N} \delta_k > \gamma | \delta_n\right]$$

and the first order condition implies $\delta_n < \theta_n^*$, or that the division under-reports its true value. Since this is true for all divisions, it now follows that the expansion is not always undertaken when it is optimal to do so.

An alternative policy for headquarters is to request each division $n \in \{1, \ldots, N\}$ to submit an estimate of the value of the regional expansion to them, which we denote by δ_n, announcing that the project will be undertaken if the sum of the estimates exceed the cost, and that the n^{th} division will receive a transfer of

$$t_n = \left(\sum_{k=1}^{N} \delta_k - \gamma - \delta_n\right)$$

if the project is undertaken. In this case the divisional heads play a different game of incomplete information. Given the announcements of the other players $(\delta_1, \ldots, \delta_{n-1}, \delta_{n+1}, \ldots, \delta_N)$ and his own private information θ_n^*, the payoff to the n^{th} division head is

$$1\left\{\sum_{k=1}^{N} \delta_k > \gamma\right\}\left[\theta_n^* + \sum_{k=1}^{N} \delta_k - \gamma - \delta_n\right]$$

Regardless of the announcements of the other divisions, a weakly dominant strategy for the n^{th} division head is to truthfully report his private information. If he reports a value of less than θ_n^*, then his payments are not affected if the facility is added, but profitable opportunities for construction are missed whenever

$$\sum_{k=1}^{N} \delta_k < \gamma < \theta_n^* + \sum_{k=1}^{N} \delta_k - \delta_n$$

If he reports a higher value than θ_n^*, then his payoff is negative whenever the project is undertaken and both inequalites above are reversed. Therefore this scheme has the advantage of being efficient. Its main disadvantage is that the divisional heads are necessarily exposed to uncertainty about their payouts as a function of information beyond their control.

5 Hidden information

Contracts are often made between parties who are differentially informed. Contracting between a principal who is less informed about the value of a joint project than the agent poses a challenge to the principal, who would like to extract as much of the value for himself yet runs the risk of the agent rejecting his contract offer or misrepresenting the true value of the project. This scenario can arise within an organization. For example employees are often better situated than their managers to recognize the cost saving, efficiency improving, quality enhancing innovations to manufacturing and distributing the firm’s products. Presumably managers are keen to implement bonuses and promotion schemes that elicit proactive suggestions from their workers to capitalize on their insider knowledge, but only compensate them enough to deter them from quitting the firm.

Asymmetric information also presents a hurdle to reaching contracts between parties who belong to different organizations. For example when a venture capitalist finances the project of an entrepreneur, the venture capitalist typically knows much less about the value of the project than the entrepreneur. Recognizing his informational disadvantage, the venture capitalist, nevertheless, would like to structure payments to the entrepreneur that encourage him to develop products with great potential, spend less effort developing products with limited chances of market success, and yet extract maximal rent from the entrepreneur on all projects. Another example is when an organization hires an engineering or construction firm to undertake some repairs or building. The expertise the construction firm has about the project gives it much more precise estimates of how much work is involved than the organization. Similarly the expertise and effort a legal partnership or a consulting agency brings to a project is sometimes harder to value by the principal firm than the agency itself.

5.1 Research and development

We develop these ideas within the following model of a pharmaceutical company that undertakes research and development. Suppose the research division makes a discovery that can be developed into a new product line generating revenue for the company. Product development is indexed by a positive real number x (which represents the number and size of demand for the new products, potential risks from the side effects of new products exploiting the discovery, and so on). The expected summed discounted net revenues from the new product line depends on this index through the logarithmic mapping $\log(1+x)$. The cost of undertaking the development depends on whether the discovery is major or minor. If the discovery is minor, the cost function takes the linear form $\delta_1 x,$ but the cost function of product development stemming from a major discovery is $\delta_2 x,$ where $0 < \delta_2 < \delta_1.$ Drawing upon expertise in its marketing and finance divisions headquarters can verify the product development index x and the gross value to the firm of the $\log(1+x)$ but not the cost function $\delta_j x$ for $j \in \{1,2\}$. If the research division is not assigned a sufficiently large research budget to develop the new product line, they continue experimenting on existing product lines, which has a budget line of $b.$ In this game management initially proposes a policy on budgets for discoveries, and then the head of the research division announces whether there has been a discovery or not, and if so, whether it is major or minor. If the division head claims that the discovery is minor, then the division is assigned a budget of b_1 and directed to develop the product line to index x_1, and if he claims his discovery is major, then the division is assigned a budget of b_2 and ordered to attain a product development index of x_2. When a discovery occurs, the division head knows the whether the discovery is minor or major even before the contract is proposed. However management are not privy to this information, and if they informed about a discovery by the head, they assesses the probability of a minor discovery at $\theta \in (0,1)$. We assume that any agreement reached between management and the research division head regarding the budget and the product development can be enforced, but that the company cannot penalize the head for internally misrepresenting the significance of his division's discovery.

5.2 Full information revisited

As a point of departure and for future comparisons, let us briefly review what management would do if it could directly observe whether the discovery was modest or revolutionary. These are the two special cases when $\theta = 0$ (a major discovery) or $\theta = 1$ (a minor discovery). Whatever the type of discovery, the division must be compensated sufficiently well to achieve the product development goals set by management. The participation constraints require that for $j \in \{1,2\}$ the head must be induced to announce a discovery when it occurs:

$$b_j - \delta_j x \geq b$$

In the full information case, management chooses wages and hours (b_j, x_j) for each state $j \in \{1,2\}$ to maximize the present value of the firm

$$\log(x_j) - b_j$$

subject to the participation constraint. We denote the solution to the full information case by (b_1^*, x_1^*) and (b_2^*, x_2^*). Since the objective is declining in b_j, the participation constraint is optimally met with equality in both states. In other words management extracts all the rent from the research division regardless of which state occurs.

Substituting the budget equation implied by the participation constraint into the management's objective function we obtain:

$$\log(1 + x_j) - \delta_j x - b + \lambda_j x_j$$

where λ_j is a Lagrange multiplier to capture the constraint that $x_j \geq 0$. Maximizing the firm's value over the product development index x_j we obtain the first order and complementary slackness conditions for x_j^* and λ_j:

$$1 = (1 + x_j^*)(\delta_j - \lambda_j)$$
$$0 = \lambda_j x_j$$

Solving for the optimal x_j^* and λ_j we obtain

$$x_j^* = \max\{0, \delta_j^{-1} - 1\}$$

Intuitively speaking, in the event of a major discovery, the marginal benefit of product development $(1 + x_2)^{-1}$ exceeds its marginal cost δ_2 for small positive x if and only if $\delta_2 < 1$. Thus if $\delta_2 \geq 1$, the discovery is ignored and its potential is untapped. If $\delta_2 < 1$ it is profitable for the firm to develop new products, choosing an interior point over the boundary point $x_2 = 0$. In particular the marginal benefit from increasing the product development index $(1 + x_2)^{-1}$ should be equated with its marginal cost δ_2.

5.3 Incomplete information

Turning to the incomplete information case, we now consider contracts or policies that provide sufficient incentive for the division head to truthfully announce the nature of his division's discovery. The head truthfully reports the type of discovery when it is in his interest to do so. There are two parts to the truth telling constraint. In the event of making a major discovery, then the division should not be worse off if he acted as through the discovery was of minor significance. This constraint is captured by the inequality

$$b_2 - \delta_2 x_2 \geq b_1 - \delta_2 x_1$$

Also, if the discovery is indeed minor, the head should not be encouraged to act as if it was major. Thus

$$b_1 - \delta_1 x_1 \geq b_2 - \delta_1 x_2$$

In addition to the truth telling constraints defined above, the head must be compensated sufficiently well to complete the job. The participation constraints which guarantee this are identical to those in the full information case described above. Although there are four constraints in all, the participation constraint that applies when the discovery is major, is redundant. To see this, note from the first and last constraints that

$$b_2 - \delta_2 x_2 \geq b_1 - \delta_2 x_1 > b_1 - \delta_1 x_1 \geq b$$

To manage the research division budgets (b_1, b_2) and product development goals (x_1, x_2) must be chosen to maximize

$$\theta[\log(x_1) - b_1] + (1-\theta)[\log(x_2) - b_2]$$

subject to the three constraints

$$b_2 - \delta_2 x_2 \geq b_1 - \delta_2 x_1$$
$$b_1 - \delta_1 x_1 \geq b_2 - \delta_1 x_2$$
$$b_1 - \delta_1 x_1 \geq b$$

The solution $(b_1^o, b_2^o, x_1^o, x_2^o)$ can be derived sequentially. First we show that the optimal budget for a minor discovery exactly offsets the extra cost incurred by announcing it. Next we show that the optimal budget in the event of a major discovery is just sufficient to deter the research head from dishonestly announcing that the discovery is minor. This only leaves the optimal product development goals to derive. If a major discovery is made the head is ordered to achieve the same level of product development that he would have achieved if management could directly observe the importance of the discovery. Finally we derive the product development requirement for a minor discovery, and show that it is less than the amount that management would demand if they could directly observe the importance of the discovery.

5.4 Research budget

To derive the solution for b_1^o, suppose that, contrary to our claim, the optimal budget for a minor discovery required two dollars more than $b + \delta_1 x_1$, and that the three constraints are met. We now consider reducing both b_1 and b_2 by one dollar. Notice that the three constraints will continue to be met, and that the objective of the objective function rises by one dollar. A simple extension of this argument, replacing the two dollars with any amount and the one dollar by half the amount, shows that the optimal payment cannot lie strictly above $b + \delta_1 x_1$. Therefore:

$$b_1^o = b + \delta_1 x_1$$

as we claimed above. Observe that this is precisely the payment that would be made if management independently knew the division's discovery was minor and had demanded x_1 product development.

To derive b_2^o, we first note that from above, the participation constraint in this case is met with an inequality, meaning $b_2^o > b + \delta_2 x_2$. We now show that the budget allocation for a major discovery has a cost plus structure, paying the basic fee of b_1^o (the budget for a minor discovery) plus compensation to induce extra product development in the event of a major discovery:

$$b_2^o = b_1^o + \delta_2(x_2 - x_1)$$

To see this note first that b_2^o must be at least this level otherwise the incentive compatibilty constraitn is not met. But following a similar argument to the one we have just used, and b_2^o was set higher than this level, profits to the firm could be impoved by reducing it a dollar, since neither of the incentive compatitbility constraints would be affect and expected profits would increase by the fraction $(1-\theta)$ of a dollar. This establishes the optimal budget for the research division.

5.5 Product development

If the product development specifications were set exogenously, for example because the nature of the discovery did not affect which products would be developed but only affected an initial setup cost, this would complete the solution to the problem. Temporarily dropping the dependence of revenue and costs on the product development index, the discussion above shows that head of research expects to garner an expected rent of:

$$(1-\theta)(b_2^o - b_2^*) = (1-\theta)(\delta_1 - \delta_2)x_1^o$$

for his division if the nature of their discoveries are kept confidential within the division, and this rent is at the expense of management or shareholders.

When product development is endogenous (and thus depends on which state occurs), the budgets are determined as above. To solve for the optimal product development specifications, we first substitute the solutions for b_1^o and b_2^o into the firm's objective function to obtain a mapping in the remaining two variables x_1 and x_2 :

$$\theta[\log x_1 - \delta_1 x_1] + (1-\theta)[\log x_2 - \delta_2 x_2 - (\delta_1 - \delta_2)x_1] - b + \lambda_1 x_1 + \lambda_2 x_2$$

where λ_j reflects the non-negativity constraint on x_j for $j \in \{1,2\}$. The first order condition for x_2^o is exactly the same condition that pertains to a world where management is as informed as the division head. In the event of marginally increasing the index should have equal and opposite effects on revenue and costs,

In the event of a minor innovation, the linear cost structure and the log revenue assumption implies the net marginal benefits of product development are:

$$\theta(x_1+1)^{-1} - [\theta\delta_1 + (1-\theta)(\delta_1 - \delta_2)]$$

We derive a necessary and sufficient condition for ignoring a minor discovery by focusing on the values of $\delta_1,\ \delta_2$ and θ that yield negative marginal benefits at the boundary $x_1 = 0$. The condition is

$$\delta_1 + \frac{(1-\theta)}{\theta}(\delta_1 - \delta_2) \geq 1$$

Notice the condition is weaker than in the full information case (which corresponds to setting $\theta = 1$ where management knows the discovery is minor). This reinforces a point we made in the general framework: to induce truth telling about major discoveries management may block the development of minor discoveries that would have been undertaken in a full information world. If the blocking condition above is violated, minor discoveries should be developed too, and we obtain the interior solution by equating the first derivative to zero and solving for x_1^o as

$$x_1^o = \frac{1}{\delta_1 + \xi(\delta_1 - \delta_2)} - 1$$

where $\xi \equiv (1-\theta)/\theta$. Since $\delta_2 < \delta_1$ the marginal cost of expending effort is lower under a major discovery than a minor one, and thus our previous discussion implies

$x_2^o > x_1^* > x_1^o$, as is evident from the formulae for the optimal product development choices.

Having derived the optimal indices, the research budget is determined by substituting (x_1^o, x_2^o) into the budget equations derived for the general model. For example if all discoveries are developed so that $x_1^o > 0$ then

$$b_1^o = b + \frac{\delta_1 - \delta_2}{\xi^{-1}\delta_1 + \delta_1 - \delta_2}$$

and

$$b_2^o = b + \frac{\delta_1 - \delta_2}{\xi^{-1}\delta_1 + \delta_1 - \delta_2}(1 + \delta_1 - \delta_2)$$

whereas if modest discoveries are discarded, meaning $x_1^o = 0,$ then all the rent from major discoveries accrue to the firm and research scientists are invariably paid their reservation wage $b.$ The efficiency losses from private information can be calculated in a similar manner for this example.

How does x_1^o compare with x_1^*? This inequality states that at the optimal product development level, there are net marginal benefits of further product development when only a minor innovation occurs. This occurs because management must provide incentives for the divisional head to declare the truth in the event of a major discovery. Management achieve this by increasing the division's budget in that state and also by making it less attractive for him to withhold information about a major discovery, by directing that only modest product development be undertaken if he declares the discovery is minor.

5.6 Losses attributable to hidden information

When product development is determined endogenously, the gain to the division from keeping the nature of their discoveries private is $\theta[c_1(x_1^o) - c_2(x_1^o)]$ whereas the loss to shareholders is

$$\theta[c_1(x_1^o) - c_2(x_1^o)] + (1 - \theta)\{[\pi(x_1^*) - c_1(x_1^*)] - [\pi(x_1^o) - c_1(x_1^o)]\}$$

If the discovery is major, shareholder losses exactly offset the gains to the research division. If the discovery is modest, the division does not garner any rent but shareholders lose

$$[\pi(x_1^*) - c_1(x_1^*)] - [\pi(x_1^o) - c_1(x_1^o)]$$

relative to what would happen in a world of full information. This expression must be positive because h_1^* maximizes $\pi(h_1) - c_1(h_1)$ whereas h_1^o does not. The product of this expression and $(1 - \theta)$ is therefore the efficiency loss associated with private information about the nature of the discovery.

6 Moral Hazard

Moral hazard arises when the unobserved choices of one player affects the payoff he receives from a person he is contracting with. Since the player's choice is not observed by other parties to the contract, the contract cannot directly specify which choice should be taken. Linking the player's payments to the consequences of his action, can help align his incentives with those of the other players, even though the consequences are only partly attributable to or caused by the action itself. For example, managers are paid to make decisions on behalf of the shareholder interests they represent. If they were paid a flat rate, why would they pursue the objectives of shareholders? Lawyers representing clients are more likely to win if they are paid according to their record, and also whether they win the case in question or not. The extent of warranties against product defects may affect how a product is used, and how much care is taken.

Sometimes the unobserved action can be inferred exactly at some later point in time. If an air-conditioning unit is installed during winter, the guarantee should extend to the summer, so that the owner is compensated if the unit malfunctions during peak usage. Similarly car mechanics can be paid a fixed wage if there are also penalty provisions for poor workmanship that might only be revealed after the vehicle has been serviced. In these cases a potential moral hazard problem is resolved by extending the contract period to cover a sufficiently long warranty period.

Finaly we could perhaps internalize the moral hazard: that means not selling insurance, or no diversification, with managers owning their own firms . . . risk aversion and solvency issues.

Number and size of firms. There are about 14 million sole proprietorships and partnerships, and 4 million corporations in the U.S. About 1,500 corporations hold about 70 percent of assets of all U.S. non-financial corporations. Management objectives. As a first approximation, it is useful to think that:

1.Sole proprietors maximize their expected utility from the firm, taking account of their other life cycle considerations.

2.Partners bargain with each other, each partner maximizing her expected utility.

3.Shareholders collectively maximize the expected value of the corporations they own.

6.1 Unobserved actions and signals

Shareholders begin the game we now consider by making an offer to the manager. The manager can reject the offer by choosing to be employed elsewhere (or not at all). We denote the utility receives from taking the outside option by u_0. If the manager accepts the offer he chooses one of two actions, shirking or working diligently. Neither of these would be observed by the shareholders. Shirking gives the manager more personal satisfaction, whereas the board of directors prefers the manager to work diligently. This conflict between the board (or shareholders) is not meant to highlight the manager's trade-off between work and say, golf. It simply reflects the fact that their respective priorities are not automatically aligned. We denote the utility the manager recieves from receiving compensation of w and working diligently by $u_2(w)$, and the utility the manager recieves from receiving compensation of w and shirking by $u_1(w)$.

In our model both $u_1(w)$ and $u_2(w)$ are concave increasing functions with $u_1(w) > u_2(w)$ for all levels of compensation w.

To help motivate the analysis we first assume that the actions of the manager can be directly monitored by shareholders. If risk neutral shareholders directly observe how hard the manager works, then the optimal contract with him is a fixed wage w^*, since the manager is risk averse. There is no rationale for paying the manager on the basis of the firm's performance. To induce the manager to accept employment with the firm and work hard, the participation constraint must be satisfied:

$$u_0 \leq u_2(w^*)$$

Since $u_2(w)$ is increasing in w, its inverse $h(u)$ exists (where by definition $w = h(u_2(w))$), and is an increasing function. Hence the optimal contract is to set $\overline{w} = h(u_0)$.

Instead of assuming the manager's activities are monitored, we now suppose that shareholders only observe a signal which is generated by a probability distribution that depends on the manager's choice. Let the random variable x denote the signal, and suppose $f_j(x)$ is the probability density function for x when the manager chooses action a_j for $j \in \{1,2\}$. The ratio of the two probability density functions, denoted $g(x)$, plays an important role in our analysis of this problem:

$$g(x) \equiv \frac{f_1(x)}{f_2(x)}$$

We interpret a realization of $g(x)$ as the likelihood that the manager shirked rather than worked hard. When shareholders receive a signal x^* they are inclined to believe the manager had shirked if $g(x^*) > 1$, and conversely worked diligently if $g(x^*) < 1$. We remark that $g(x)$ ranges from 0 to ∞ : if $g(x^*) = 0$, shareholders conclude the manager worked diligently, whereas if $g(x^*) = \infty$ he surely shirked. One example of a signal is the firm's abnormal return. A reasonable goal for shareholders is to maximize the expected value of abnormal returns net of expected managerial compensation. In this case we assume that

$$E_2[x] \equiv \int_{-\infty}^{\infty} x f_2(x)dx > \int_{-\infty}^{\infty} x f_1(x)dx \equiv \int_{-\infty}^{\infty} g(x) f_2(x)dx \equiv E_2[xg(x)]$$

To further specialize, suppose that if the manager works diligently, abnormal returns are uniformly distributed on the closed interval between -1 and 1, which means $f_2(x) = 1/2$ for $x \in [-1,1]$, but that if he shirks the cumulative distribution function has a triangular shape on the same support, $f_1(x)$ taking the from $(1-x)/2$ for $x \in [-1,1]$. Then $g(x) = f_1(x)$, so the likelihood ratio is monotonically declining in x with $g(x) > 1$ if and only if $x < 0$. In this specialization $E_2[x] = 0$ but $E_1[x] = -1/3$.

6.2 Optimization

When $g(x^*)$ is finite shareholders cannot deduce from the realization x^* whether the manager worked diligently or not. Nevertheless there exist pairs of probability density functions $f_1(x)$ and $f_2(x)$ that allow shareholders to implement the outcome implied by the full information contract derived above even though they only observe the signal, not the manager's action. If there is a strictly positive probability that the manager will be caught shirking, meaning $g(x^*) = \infty$ for some values of x^* that cannot be reached if the manager works diligently, then threatening him with a very high penalty will deter him from shirking. A two part contract is optimal, comprising a constant wage $\overline{w}$, supplemented by a penalty that is incurred if the manger is caught shirking. This contract fully insures the manager, and achieves the same first best solution that could be attained if the action is observed and contracted upon.

For example suppose that $f_2(x)$ is defined as before, uniform on $[-1,1]$, but that $f_1(x) = 1/3$ with support $x \in [-2,1]$. Then $g(x) = 2/3$ on the interval $[-1,1]$ but is unbounded on the interval $[-2,-1)$. If the manager shirks, he will be caught one third of the time. By posting a sufficiently high penalty for sufficiently low signals $x < -1$, he is deterred from shirking even if his wage is constant on the interval $x \in [-1,1]$. This contract form ensures the manager and the shareholders are as well off as they would be if the shareholders monitored the manager's performance. Although optimal compensation depends on the realization of the signal, a random variable, in this particular example the manager receives the certainty wage w^*, even though diligence cannot be verified in every state.

A moral hazard problem arises only if the support of $f_1(x)$ is contained in the support of $f_2(x)$. When $g(x)$ is a finite valued function the optimal contract entails the manager accepting some risk. The compensation he receives might depend on the signal, and to indicate that dependence we now write $w_j(x)$ for the manager's compensation when he chooses action $j \in \{1,2\}$ and the shareholders subsequently observe the signal x. As above, there are two restrictions on contracts the manager accepts, the incentive compatibility and participation constraints. The participation constraint is now expressed as

$$\int u_j(w(x))f_j(x)dx \geq u_0$$

for $j \in \{1,2\}$. The incentive compatibility inducing diligent work is

$$\int u_2(w(x))f_2(x)dx \geq \int u_1(w(x))f_1(x)dx = \int u_1(w(x))g(x)f_2(x)dx$$

These two constraints must be satisfied for the manager to accept employment with the firm and work diligently.

The minimum cost contract for engaging the manager to work diligently is found by choosing w for each x to minimize the expected cost of managerial compensation subject to the participation and incentive compatibility constraints. Equivalently, we choose w for each distinct value of g that x induces through the mapping $g(x)$ to minimize expected costs. The second formulation is more convenient because it shows that tyhe opitmal contract only depends on the likelhood ratio rather than the signal itself. Thus supposing there are two values of the signal, say x' and x'', that induce the same likelihood ratio, meaning $g(x') = g(x'')$, then optimal compensation should also be the same, that is $w(x') = w(x'')$. If $g(x)$ only takes on a handful of values then a contract taking on an uncountable number of different values is unwarrented,even though the variable x might, and exposes the manager to needless risk, consequently raising the expected cost to the shareholders of satisfy his participation constraint. For example, suppose that when the manager works diligently, the signal comes from a uniform distribution with lower bound -1 and upper bound 1, and that two steps of equal height and length on the same support comprise the probability density function for the signal when the manager shirks. Mathematically $f_2(x) = 1/2$ on $x \in [-1,1]$, while $f_1(x)$ is $3/4$ on $x \in [-1,0]$ and $1/4$ on $x \in (0,1]$. These assumptions imply $g(x)$ is $3/2$ on $x \in [-1,0]$ and $1/2$ on $x \in (0,1]$. Because $g(x)$ only takes on two values, the optimal contract only takes two values too, a base salary plus a bonus, which is received when $0 < x \leq 1$ (with probability $1/2$).

6.3 A specialization

We now assume that the utility function of the manager takes the logarithmic form

$$u_j(w) = k_j \log(w)$$

for $j \in \{1,2,3\}$ and the that the outside option offers w_0, implying $u_0 = k_0 \log(w_0)$. We also assume that $g(x)$ only takes on two values α_1 and α_2, where the probability of α_1 occurring is p. This could arise because the probability density fucntions $f_1(x)$ and $f_2(x)$ take a similar form to the example described above, or because the the only observed outcome is whether the manager is successful in securing a goal or not. The discussion above implies that the optimal compensation for the manager is a two point distribution. He is paid w_1 if α_1 occurs and w_2 if α_2 occurs. The shareholders minimize expected compensation

$$pw_1 + (1-p)w_2$$

subject to the participation constraint

$$pk_2 \log(w_1) + (1-p)k_2 \log(w_2) \geq u_0$$

and the incentive compatibility constraint

$$p\alpha_1 k_1 \log(w_1) + (1-p)\alpha_2 k_1 \log(w_2) \leq pk_2 \log(w_1) + (1-p)k_2 \log(w_2)$$

If w_1 and w_2 are lowered at the same rate, then both sides of the incentive compatibility constraint decline the same amount and the inequlaity is preserved, but only the left side of the participation constraint is affected. Since shareholders prefer paying less to the manager providing they can ensure diligent work, it is not optimal to offer a contract that satisfies the participation constratin with strict inequality. Therefore the optimal package (w_1^*, w_2^*) must satisfy the participation constraint with equality:

$$pk_2 \log(w_1^*) + (1-p)k_2 \log(w_2^*) = u_0$$

The optimal compensation payments (w_1^*, w_2^*) also satisfies the incentive compatibility constraint with equality. To see this, first substitute the left side of the participation constraint into the right side of the incentive compatibility constraint to obtain:

$$p\alpha_1 k_1 \log(w_1^*) + (1-p)\alpha_2 k_1 \log(w_2^*) \leq u_0$$

If the inequality was strict, then we could change w_1^* by dw_1^* and w_2^* by dw_2^*, leaving expected total compensation unchanged if

$$0 = pdw_1^* + (1-p)dw_2^*$$

or

$$\frac{dw_2^*}{w_2^*} = \frac{-p}{(1-p)}\left(\frac{w_1^*}{w_2^*}\right)\frac{dw_1^*}{w_1^*}$$

The effect on the manager's expected utility is

$$pk_2\frac{dw_1^*}{w_1^*} + (1-p)k_2\frac{dw_2^*}{w_2^*} = pk_2\left(1-\frac{w_1^*}{w_2^*}\right)\frac{dw_1^*}{w_1^*}$$

If $w_1^* < w_2^*$ then set $dw_1^* > 0$; if $w_1^* > w_2^*$ then set $dw_1^* < 0$. In this way the manager's expected utility increases, so that the participation contraint is met with strict inequality too. But this contradicts our eariler result that the participation constraint is met with equality. Therefore the incentive compatibility constraint is also satisfied with strict equality at the optimal contract.

Having proved both constraints are satisfied with equality, the optimal contract isfound by solving these two equations in the two unknowns w_1^* and w_2^*. From the definition of $g(x)$ it follows that $E[g(x)] = 1$ or

$$p\alpha_1 + (1-p)\alpha_2 = 1$$

Solving for α_2 we obtain:

$$\alpha_2 = \frac{1-p\alpha_1}{1-p}$$

Hence the incentive compatibility constraint can be expressed as

$$\begin{aligned} u_0 &= p\alpha_1 k_1 \log(w_1^*) + (1-p)\alpha_2 k_1 \log(w_2^*) \\ &= p\alpha_1 k_1 \log(w_1^*) + (1-p\alpha_1)k_1 \log(w_2^*) \end{aligned}$$

or

$$\frac{u_0}{k_1} = \log(w_2^*) + p\alpha_1 \log\left(\frac{w_1^*}{w_2^*}\right)$$

Similary the participation constraint can be expresssed as

$$\frac{u_0}{k_2} = \log(w_2^*) + p\log\left(\frac{w_1^*}{w_2^*}\right)$$

Subtracting one equation from the other we solve for the ratio of th two compensation levels in logarithmic form

$$\log\left(\frac{w_1^*}{w_2^*}\right) = \left(\frac{1}{k_1} - \frac{1}{k_2}\right)\frac{u_0}{p(\alpha_1 - 1)}$$

Now substituting this solution into either constraint yields

$$\begin{aligned}\log(w_2^*) &= \frac{u_0}{k_2} - \left(\frac{1}{k_1} - \frac{1}{k_2}\right)\frac{u_0}{(\alpha_1 - 1)} \\ &= \frac{u_0(\alpha_1 - 1)}{k_2(\alpha_1 - 1)} - \frac{u_0}{k_1(\alpha_1 - 1)} + \frac{u_0}{k_2(\alpha_1 - 1)} \\ &= \frac{\alpha_1 u_0}{k_1(\alpha_1 - 1)} - \frac{u_0}{k_2(\alpha_1 - 1)}\end{aligned}$$

Also

$$\begin{aligned}\log(w_1^*) &= \left(\frac{1}{k_1} - \frac{1}{k_2}\right)\frac{u_0}{p(\alpha_1 - 1)} + \log(w_2^*) \\ \frac{p(\alpha_1 - 1)}{u_0}\log(w_1^*) &= \frac{1}{k_1} - \frac{1}{k_2} + \frac{p\alpha_1}{k_1} - \frac{p}{k_2} \\ &=\end{aligned}$$

Taking the exponential of both sides of both equations and rearranging we conclude

$$\begin{aligned}w_1^* &= \exp\left[\frac{u_0}{k_1 p(\alpha_1 - 1)} - \frac{\alpha_1 u_0}{k_2(\alpha_1 - 1)}\right] \\ w_2^* &= \exp\left[\frac{u_0}{(\alpha_1 - 1)}\left(\frac{\alpha_1}{k_1} - \frac{1}{k_2}\right)\right]\end{aligned}$$

6.4 Losses from moral hazard

The costs of moral hazard arise from the fact that the expected value of compensation under the optimal contract, which we denote by $E[w^*(x)]$ exceeds $\overline{w}$ the fixed amount that would be paid if monitoring the manager was possble. In our example the difference is

$$\begin{aligned}&pw_1^* + (1-p)w_2^* - \overline{w} \\ &= pw_1^* + (1-p)w_2^* - \frac{u_0}{k_2}\end{aligned}$$

7 Summary

The example at the beginning of this chapter sets its tone. In order to procure a component from a supplier, a firm must offer a contract that the supplier accepts, and a contract that ensures the supplier will ifnd that it is in its own to produce according to the specifications that the procurement firm wants. These two conditions are called the participation constraint and the incentive compatitibility constraint. The procurement firm minimizes its costs subject to these two constraints.

The application which follow build upon this theme.

When all the information sets at the terminal nodes of a contract are singletons, the contract has a full set of contingencies. The most straightforward mechanism design problems are those where there are fully contingent contracts. We analyzed three applications of this problem, contracts with unskilled laborers and itinerant workers, procurement procedures from component suppliers and service providers. In these examples the principal optimally designs a contract that maximizes the social surplus, that is the aggregate value of the game to all the players. In these examples the agent has the option of rejecting the contract offer, and seeking employment or business elsewhere. This alternative constrains the contracts that the principal will offer to avoid rejection. Thus the participation constraints restrict the principal to those contracts where the best response of agents is to accept the contract the principal offers. These two principles, maximizing social surplus and obeying participation constraints, lead us to a general characterization of the solution. When contracts are fully contingent, the principal should design a contract that maximizes the social surplus and expropriates all the gain to himself subject to the participation constraints applying to the agents.

For example when a firm contracts with a builder to undertake to undertake some construction, the quality of the work is sometimes hard to judge, because if there is structural failure the reasons might not be clear. Unfortunately errors of both types (punishing innocent parties and failing to convict guilty parties) in courts determining liability. In cases like the builder might be the only party to know whether his construction methods are sound or not. Recognizing the agent makes choices that the principal does not observe, the principal offers contracts that provide incentives for the agent to adopt construction methods that are compatible with the principal's objectives. This is an example of an incentive compatibility constraint, which are relevant whenever the agent makes choices affecting the principal's expected payoffs that the principal cannot observe or infer.

Another type of example is when agents know more about the problem, observing finer information partition at the terminal nodes than the principal. A question arises does. This chapter has analyzed games with first movers as problems in contract design. The first half of the chapter concentrated on games where there is a single principal designing a contract and one or more agents who have the opportunity to implement it. In a full information game with a single first mover, only the participation constraint limits the principal's power to extract rent from the agents. If the principal attempted to extract more than is dictated by the participation constraint, the agents would reject the contract. In first mover games with perfect information

One area where incomplete information about a valuations is in employment. The last topic of this chapter was on search games, in which an employer seeks to fill a position by sequentially discovering information about candidates and their abilities until one is hired, thus stopping the process. This topic also extends the earlier work on multilateral bargaining, because it involves selecting a recruit from a set of job candidates. We analyzed games where there are only a finite number of candidates who seek the job and look qualified enough to justify serious consideration, and showed how the reservation quality the firm demands increases, as the field of qualified candidates increases, and falls as the number of rejected candidates increases. We also discussed how recall affects the list of offers; if candidates who reject offers that the firm makes have the opportunity to reconsider better offers as they are made to subsequent candidates, then offers a firm makes do not increase as steeply as when candidates have no right of recall. Increased competition between candidates reduces the firm's wage bill. Finally we discussed the trade-off between two partners as they search for ways of achieving a better match as they simultaneously bargain over the share of their resulting partnership, and compare this with the results from an dynamic optimizing model of the type analyzed in Chapter 11. Does bargaining lead to too much investment, or too little?