A primer on Contract and Game Theory for the Applied Micro Economist

October 18, 2010

Contents

1 Introduction

These notes should provide you with a basic background of the economic theory which will be used extensively in this course. Most of you will have seen a lot of this before but it can serve as a reference book or a refresher. These notes will talk mostly about contract theory and will also touch on game theory.

Contract theory, or the theory of incentives, occupies itself with the problems that arise when delegating a task to someone with private information. If this private information is relevant to the task at hand, the informed party has to be given correct incentives to either reveal this information or to act in the interest of the person who delegated the task.

The Principal Agent framework is a tool economists use to analyze these situations. The term **principal** refers to the uninformed party who would like to delegate a task (e.g. a manager who delegates a project to his employee) and the term **agent** refers to the informed party (e.g. the worker who knows exactly how much work he has done on the project). Throughout this course you will see the term 'principal' being used interchangeably with lender, bank or landlord. The term 'agent' will refer to borrowers and tenants.

The private information of the agent is usually one of two types: either the agent's actions are not observable, which is known as *moral hazard*; or the agent has some private knowledge relevant to the contracting problem, which is referred to as the problem of *adverse selection*. These notes will discuss both in detail and you will see applications of this theory throughout the course.

Unless it is stated explicitly, you should throughout the entire course assume that both the principal and the agent are risk neutral. Risk neutrality implies that both parties maximize expected payoff, as expected payoff and expected utility are the same thing under risk neutrality.

2 Moral Hazard

Moral hazard - sometimes referred to as 'hidden action'- is the problem that arises when the principal cannot observe the actions of the agent. For example a manager who delegates work to an employee cannot constantly observe how hard the employee is working on the project. He will only observe the final outcome and unless output is a perfect signal of the worker's effort, the manager will never know how much effort the agent invested. Effort is costly to the agent, so he will exert as little as possible.. The root of the problem is that the worker and the manager have different objective functions and their incentives are not always aligned.

2.0.1 The set up

To make matters concrete, assume the principal has delegated a task to an agent. This task will succeed with probability e in which case it produces output of value y to the principal. With probability $(1-e)$ it will fail and output will be zero. The agent himself does not directly care about the outcome of the task, but he can influence the probability of success by choosing how much effort $e \in (0,1)$ to exert. Effort has to somehow be correlated with the likelihood of success and/or the payoff of the project and to keep matters simple, we simply say that the probability of success is equal to the effort level the agent exerts. This means the agent can directly choose e, the likelihood of success. However, the agent incurs a disutility of $\frac{1}{2}ce^2$ from exerting effort.

2.0.2 Observable Effort - The first best

Before we turn to the situation where effort is unobservable, we first solve the case where effort is completely and costlessly observable to the principal. In this situation there are no informational asymmetries and we can solve for the optimal effort level and the optimal transfers that achieve overall efficiency. This will serve as a benchmark against which we can compare the situation when there is moral hazard.

With observable effort the principal can contractually stipulate how much effort the agent has to exert, but in return he has to compensate the agent for his effort. This means the principal maximises his own payoff subject to the constraint that the agent will participate in the contract. If the agent does not accept the contract, his outside option gives him a payoff of u. Formally, the principal solves

$$\max_{e,w} ey - w$$
$$\text{s.t. } w - \frac{1}{2}ce^2 > u \tag{PC}$$

With probability e the principal has payoff y from the project and in any state of the world he has to pay the agent a wage w. This is maximised subject to the constraint that the agents utility net of effort cost exceeds his outside option. This constraint is usually referred to as the *participation constraint* or often also as the *individual rationality constraint*.

We can see right away that the participation constraint is going to be binding: if it were not binding, the principal could increase his payoff by lowering w a little bit until the constraint starts to bind. We can therefore solve this problem simply by maximising

$$\max_{e} ey - \frac{1}{2}ce^2 - u \tag{2.1}$$

After taking the first order condition with respect to e, equating it with zero and rearranging we get that $e = \max\{\frac{y}{c}, 1\}$. This is the first best level of effort that maximises overall efficiency. Throughout these notes (and throughout the course) we are going to assume $c \geq y$, which means we will be getting an interior solution for effort. Hence we see that the first best expected surplus is going to be $\frac{y^2}{2c} - u$. The payoff to the agent is exactly equal to his outside option u and the principal's payoff is the remaining $\frac{y^2}{2c} - u$.

2.0.3 Unobservable Effort

Now let's consider the situation where effort is not observable to the principal. He can, however, observe wether or not the project succeeded, which means he has the ability to make the agent's payoff contingent on project outcome. Let's call h the agent's payoff in the state of the world when the project is successful and l the agent's payoff in the state of the world when the project fails. Note that the payoff to the agent need not necessarily be positive: the principal can also punish the agent for not successfully completing the project.

The principal now solves:

$$\max_{l,h} e(y-h) + (1-e)(0-l)$$

such that

$$eh + (1-e)l - \frac{1}{2}ce^2 \geq u \tag{PC}$$

$$e = \arg\max_{e} eh + (1-e)l - \frac{1}{2}ce^2 \tag{ICC}$$

As in the previous section, the principal maximises his own expected payoff subject to the constraint that the agent is willing to participate in the contract. However there is now an additional constraint. The second constraint is knowns as the *incentive compatibility constraint* and it states that the effort the agent exerts has to maximise his expected payoff. Though the principal cannot directly control effort, he can indirectly choose how much effort the agent will exert by choosing the agent's payoffs in both states of the world.

We know from the previous section that in the absence of the incentive constraint, the principal would like to set the agent's effort level to $e = \frac{y}{c}$. Taking the first order condition of the incentive constraint gives $\frac{h-l}{c} = e$. Putting those two conditions together we see that any combination of h and l for which $h - l = y$ will result in the first best effort level.

For example, the principal could set $h = y$ and $l = 0$. The principal would then make zero profit and the delegation problem essentially disappears. If, on the other hand, the principal sets $h = 0$ and $l = -y$, this would violate the agent's participation constraint and he would never accept such a contract.

So which combination will the principal choose? The principal will choose the combination of h and l that just satisfies the participation constraint of the agent. The agent therefore exerts the first best level of effort and has expected payoff equal to u as before. The principal's expected payoff is again $\frac{y^2}{2c} - u$ and overall surplus created is still $\frac{y^2}{2c} - u$. Although effort is not observable, overall efficiency is still achieved.

It turns out that unobservable effort by itself does not lead to any market imperfections. However, as we will see in the following sections, limited liability and risk aversion combined with unobservable effort will severely affect the range of contracts that can be implemented. With limited liability or risk aversion the first best effort level will no longer be achievable.

Result 1 *Despite effort being unobservable the first best will be achieved with unlimited liability and risk neutral agents.*

2.1 Limited Liability

In the section above the principal had the option to both reward the agent in the good state and to punish him in the bad state. Often, however, this is not possible. For example, an agent with little wealth cannot be asked to pay anything when the project fails. This is known as *limited liability*. The principal can reward the agent in the good state but can only exert limited punishment in the bad state.

For simplicity, we are going to focus on the case where the agent has zero wealth. That means nothing can be deducted in case the project fails. In general, however, we could have assumed that he has a limited amount of wealth $\omega < h$ (see the section 2.2 for example). We can incorporate this into the problem above by adding a limited liability constraint (LLC):

$$\max_{l,h} e(y-h) + (1-e)(0-l)$$

subject to

$$eh + (1-e)l - \frac{1}{2}ce^2 \geq u \tag{PC}$$

$$e = \arg\max_{e} eh + (1-e)l - \frac{1}{2}ce^2 \tag{ICC}$$

$$l \geq 0, \quad h \geq 0 \tag{LLC}$$

The first step towards simplifying this is realizing that the limited liability constraint will never be binding for h. In the good state the agent wants to reward the agent and therefore will more than zero. What about l? As we worked out in the previous section, this constraint will bind.

With this in mind we can start plugging in $l = 0$ into the incentive and participation constraints. The incentive constraint therefore becomes $e = \frac{h}{c}$.

We can incorporate the fact that $l = 0$ and $e = \frac{h}{c}$ into the problem and rewrite it in a simpler form:

$$\max_{h} \frac{h}{c}(y-h)$$

subject to

$$\frac{h^2}{2c} \geq u \tag{PC}$$

Through substitution we have eliminated all the constraints except for the participation constraint. Will the participation constraint be binding? The answer is that it depends on the outside option u and that there will be two separate cases, one where the participation constraint is binding and one where it is not.

2.1.1 Case 1: Participation Constraint not binding

In this case we can simply ignore the participation constraint and solve for the first order condition right away. This gives $h = \frac{y}{2}$. In this case effort will be equal to $e = \frac{y}{2c}$, half of what effort is in the first best! The agent's expected payoff equals $\frac{y^2}{8c}$, the principals payoff equals $\frac{y^2}{4c}$ and overall surplus is equal to $\frac{3y^2}{8c}$. Note that since the agent's participation constraint is not binding, the agent gets a higher utility here than he would have if effort were observable. Intuitively, this is because the agent cannot be punished enough in the bad state of the world, so the principal tries to reward the agent for the good state. This is sometimes referred to as *limited liability rent.* Although we assumed that principal holds all the bargaining power, the agent's payoff is not driven down to his reservation utility because it is in the principal's interest to reward the agent for the good state.

Result 2 *For a low outside option such that the participation constraint is not binding, the agent's payoff is higher under unobservable effort than under observable effort.*

2.1.2 Case 2: Participation Constraint binding

What happens if $\frac{y^2}{8c} \leq u$? If the agent has a high outside option then the principal has to offer him a h at least high enough to ensure the agent's expected payoff is as good as the high outside option. Rearranging the above participation constraint for h, we get that $h = \sqrt{2uc}$. Note that the higher the agent's outside option, the more the principal has to pay the agent in the good state to incentivise effort. Overall expected surplus is therefore $\sqrt{\frac{2u}{c}}y - u$, the agent's payoff is by definition equal to his outside option u and the principal's payoff is the remainder of the surplus created $\sqrt{\frac{2u}{c}}y - 2u$. Note that both the overall surplus as well as the principal's payoff are decreasing in u. Intuitively, the worse the agent's outside option, the easier it is for the principal to incentivise effort.

Whether or not the participation constraint is binding, we find that:

Result 3 *With unobservable effort and a binding limited liability constraint, effort and overall surplus will be lower than in the first best.*

2.2 Using the Lagrangean Method

In all the above sections we solved for the optimal contract analytically. That is we thought carefully about which constraints are binding and used a combination of substitution and taking first order conditions to arrive at the optimal contract.

Another, more mechanical, way of solving such problems is to use the Lagrangean method that you should have seen in the September courses. This section will solve a more general version using this Lagrangian approach to constraint maximisation. To be able to look at all possible cases, we are going to assume that the agent owns some wealth which can be seized in the bad state of the world.

We can therefore write our problem like this

$$\max_{l,h} e(y-h) + (1-e)(0-l)$$

subject to

$$eh + (1-e)l - \frac{1}{2}ce^2 \geq u \tag{PC}$$

$$\overline{e} = \arg\max_{e} e(h) + (1-e)l - \frac{1}{2}ce^2 \tag{ICC}$$

$$l \geq -\omega \quad h \geq -\omega \tag{LLC}$$

We can go ahead and simplify this by plugging in the incentive constraint $e = \frac{h-l}{c}$

$$\max_{l,h} \big(\frac{h-l}{c}\big)\big(y-h\big) - \big(1 - \frac{h-l}{c}\big)l$$

subject to

$$\frac{(h-l)^2}{2c} + l \geq u \tag{PC}$$

$$l \geq -\omega \quad h \geq -\omega \tag{LLC}$$

Notice that the optimal incentive contract (h, l) must have $h > l$ because if $h \leq l$, then from the ICC $e = 0$ and the principal gets $-l$ whereas for the same l if he sets $1 \geq h > l$ he gets $e\{1-(h-l)\} - l \geq -l$. Intuitively, the principal wants to reward the good state and punish the bade state. To achieve this, he will set a higher payment for the good state than the bad state. This also implies that one of the two limited liability constraints, $h \geq -\omega$, cannot bind.

Since the objective function is concave and the constraints are convex for $h > l \geq 0$, the Kuhn-Tucker conditions give the global maximum to this programming problem. Let λ and μ be the Lagrangian multipliers associated with the LLC and the PC respectively.

We are now ready to write out the Lagrangean:

$$L = (\frac{h-l}{c})\big(y-h\big) - \big(1 - \frac{h-l}{c}\big)l + \mu\big(\frac{(h-l)^2}{2c} + l - u\big) + \lambda\big(l-\omega\big) \tag{2.2}$$

Notice that $\lambda \geq 0$ and $\mu \geq 0$ as increasing w and reducing m relaxes the constraints and increases the principals's profits. The first-order conditions with respect to h and l are:

$$\frac{y}{c} - \frac{2}{c}(h-l) + \mu\frac{1}{c}(h-l) = 0 \tag{FOC1}$$

$$-\frac{y}{c} + \frac{2}{c}(h-l) - 1 + \lambda - \mu\frac{1}{c}(h-l) + \mu = 0 \tag{FOC2}$$

Adding (FOC1) and (FOC2) we get

$$\lambda + \mu = 1 \tag{2.3}$$

This shows under an optimal contract (h, l) at least one of the two constraints, the LLC and the PC, must be binding. You should have seen during the September course that a constraint is binding if and only if its associated Lagrange multiplier is non-zero. Given that μ and λ add up to 1, it has to be the case that at least one and possibly both are greater than zero.

2.2.1 Case 1

What does the contract look like if only the participation constraint is binding, but the limited liability constraint is not? If the LLC is not binding then $\lambda = 0$ and $\mu = 1$ from (2.3) and $l - h = y$ from (FOC1). Substituting in the $ICC, e = \frac{y}{c}$ and from the PC $l = u - \frac{y^2}{2c}$ and $h = 1 + u - \frac{y^2}{2c}$.

The agent 'pays' $\frac{y^2}{2c} - u$ in either state of the world to the principal and in return the agent gets to keep the entire returns from the project. You can think of this as a fixed rent contract for example. The agent is the full residual claimant and hence effort is at the first-best level. As we saw earlier, the surplus in the case of the first best is $\frac{y^2}{2c}$.

So long as $\omega \geq \frac{y^2}{2c} - u$, the tenant is able to pay a fixed fee equal to the landlord's share of the first-best expected surplus in all states of the world and this efficient contract is feasible. Since the PC is binding under this contract, the tenant gets a share u of the first-best surplus and the landlord gets the remaining, $\frac{y^2}{2c} - u$.

2.2.2 Case 2

Now let's turn to the case where only the limited liability constraint is binding and the participation constraint is not. In this case $\lambda = 1$ and $\mu = 0$. We get $l = -\omega$ from the LLC and $h = \frac{1}{2} - \omega$ from (FOC1) and $e = \frac{y}{2c}$ from the ICC. As you can see, effort and hence social surplus are clearly lower than in Case 1.

2.2.3 Case 3

If both the PC and the LLC are binding under the optimal contract then solving them out we get $l = -\omega$ from the LLC and $h = \sqrt{2c(u + \omega)} - \omega$ from the PC. The ICC gives $e = \sqrt{\frac{2(u+\omega)}{c}}$.

Under this contract total expected social surplus is $y\sqrt{\frac{2(u+\omega)}{c}} - (u + \omega)$.

Notice that for $u + \omega = \frac{y^2}{2c}$ the solution in this case coincides with the one in Case 1 and a first best outcome can be achieved. Both effort and social surplus are strictly increasing and concave functions of $(u+\omega)$ for $0 < u+\omega < \frac{y^2}{2c}$. Hence for $u+\omega < \frac{y^2}{2c}$, effort and surplus are lower than in the first-best, and so is the success-wage. Under this contract, the PC is binding and so the tenant's expected payoff is u and the landlord's expected payoff is the remaining amount of (expected) surplus, $y\sqrt{\frac{2(u+\omega)}{c}} - (u+\omega) - u$.

Finally, note that for $u+\omega = \frac{y^2}{8c}$ Case 3 coincides with Case 2. For all levels of $u+\omega < \frac{y^2}{8c}$ effort and hence social surplus is higher in Case 2 and for all $u+\omega > \frac{y^2}{8c}$ social surplus is higher in Case 3. This follows because again social surplus is increasing in $u+\omega$.

To summarize, for $u+\omega \geq \frac{y^2}{2c}$, a fixed-rental contract is feasible and we have the first-best effort level. Hence in this case, changes in ω have no effect on the contract or the payoffs.

2.3 Many principals competing for few agents

The previous sections all assumed that the principal holds all the bargaining power. He designed the contract that maximized his utility subject to several constraints. The underlying assumption was that there are many people willing to work and only few managers, landowners or moneylenders who are able to hire a worker. However, we could as well have assumed that there are many principals and few agents.

The analysis is very similar, but instead of designing a contract that maximizes the principal's utility, we now solve for the contract that maximizes the agent's utility subject to (i) the principal's zero profit condition (ZPC) (ii) the incentive constraint of the agent (ICC) and (iii) potentially also a limited liability constraint (LLC).

$$\max_{l,h} eh + (1-e)(l) - \frac{1}{2}ce^2$$

subject to

$$e(y-h) + (1-e)(0-l) = \rho \qquad \text{(ZPC)}$$

$$e = \arg\max_{e} eh + (1-e)l - \frac{1}{2}ce^2 \qquad \text{(ICC)}$$

$$l \geq 0 \quad h \geq 0 \qquad \text{(LLC)}$$

The first constraint here is the *zero profit condition* which states that under perfect competition the principal will make zero economic profit. This can also be thought of as a participation constraint for the principal. The principal will only participate in this contractual arrangement if his payoff is at least as high as his outside option would be. Here we refer to this outside option as ρ. For example the principal could work on the project himself in which case he would not have to hire an agent, or the principal could sell the project.

This set up is useful when it comes to studying situations in which we know that agents are rare but principals are common. In this case the agent holds all the bargaining power due to market forces driving the principal all the way to zero profit. That is, we know the zero profit condition will be binding by assumption.

The other two constraints are unchanged from the previous section, so we know the limited liability constraint still binds and so does the incentive compatibility constraint. We can go ahead and plug $l = 0$ into the ICC simplified and the ZPC.

$$\max_{l,h} eh - \frac{1}{2}ce^2$$

subject to

$$e(y-h) + = \rho \qquad \text{(ZPC)}$$

$$e = \frac{h}{c} \qquad \text{(ICC)}$$

Combining the ICC with the ZPC we get a quadratic equation in e: $e^2c - ey + \rho = 0$ which has solutions

$$\frac{y \pm \sqrt{y^2 - 4\rho c}}{2c} \tag{2.4}$$

These are the only two values of e for which the principal earns zero profits. If effort were any lower the principal would make a loss and if effort were higher the bank principal make a profit. The principal is therefore indifferent between the two solutions, the agent on the other hand as a strict preference for the bigger root.

The agents expected payoff is given by

$$\begin{aligned} eh - \frac{1}{2}ce^2 \\ = \frac{1}{2}ce^2 \end{aligned}$$

which is increasing in e, so the agent prefers the higher of the two solutions.

The optimal contract therefore is one that sets $h = ce = c\left(\frac{y \pm \sqrt{y^2-4\rho c}}{2c}\right)$, $l = 0$ and $e = \frac{y \pm \sqrt{y^2-4\rho c}}{2c}$. How does this compare to the first best? Comparing the effort level to effort under first best ($e = \frac{y}{c}$) it is clear that once again effort is lower than under the first best.

The difference to the section on many agents and few principals is only who extracts the rents created. By assumption, whenever the zero profit condition binds agents extract all the rents created and when the participation constraint binds, principals extract all the rent created. Only in the case where the participation constraint does not bind and there are many agents and few principals, do both get a share of the profits created.

2.4 Risk Averse Agents

The previous section showed that effort provision and overall surplus is less than first best when effort is not observable and when there is limited liability. In this section we look at risk averse agents. As before effort is not observable but now we drop the assumption of limited liability and replace it with risk aversion in it's place.

Before we derive the maths, it is worth to think intuitively about why risk aversion might cause a departure from the first best. We showed earlier that under the first best, the agent carries the complete risk of the project. In particular, under risk neutrality a necessary condition for effort to be at the first best level was to set $h - l = y$, that is the agent is payed significantly more in the good state of the world than he is in the bad state of the world. A risk averse agent would derive significant disutility from such a volatile income stream.

2.4.1 The Set Up

The set up necessarily is slightly different from the previous sections. Here we follow the set up used by Stiglitz (1974) which you will also see in the lecture. Let output q be determined by effort e and a random shock ϵ which has zero mean and variance σ^2.

The principal is assumed to be risk neutral and and the agent is risk-averse with the following utility function where y is his income:

$$U(y) = E(y) - \frac{r_T}{2} Var(y) \tag{2.5}$$

Where $E(\cdot)$ denotes the expectation operator and $Var(\cdot)$ is the variance. As previously, the disutility of effort is given by $\frac{1}{2}ce^2$. We focus on linear contracts which are of the form

$$y = sq - R \tag{2.6}$$

That is the agent is paid a fixed proportion s of the project's return minus a fixed fee the agent has to pay to the agent in order to work for him. A good example of this is sharecropping where the agent rents land from the landlord and in return for his effort is allowed to keep some of the returns.

Given this set up the principal solves the following maximization problem

$$\max_{R,s} e(1-s)q + R$$

subject to

$$u(y) - \frac{1}{2}ce^2 \geq u \tag{PC}$$

$$e = \arg\max_{e} u(y) - \frac{1}{2}ce^2 \tag{ICC}$$

Where $u(y)$ is the agents expected utility given his expected income y which in turn depends on the effort level chosen by the agent. r_T measures how risk averse the agent is with $r_T = 0$ corresponding to risk neutrality

Knowing that the contract is linear of the form $y = sq - R$ and knowing that $u(y) = E(y) - \frac{r_T}{2}var(y)$ we can start simplifying the above expressions. First we note that $E(y) = E[sq - R] = se - R$ and $var(y) = var(sq - R) = s^2 var(q) = s^2\sigma^2$. Putting these together we see that

$$u(y) = se - R - \frac{r_T}{2}s^2\sigma^2 \tag{2.7}$$

and the principal maximises

$$\max_{R,s} e(1-s)q + R$$

subject to

$$se - R - \frac{r_T}{2}s^2\sigma^2 - \frac{1}{2}ce^2 \geq u \tag{PC}$$

$$e = \arg\max_{e} se - R - \frac{r_T}{2}s^2\sigma^2 - \frac{1}{2}ce^2 \tag{ICC}$$

Taking the first order condition for the incentive constraint we get $e = \frac{s}{c}$. Plugging this into the participation constraint and the objective the problem further simplifies to:

$$\max_{R,s} \frac{s(1-s)}{c} + R$$

subject to

$$\frac{s^2}{2c} - R - \frac{r_T}{2}s^2\sigma^2 \geq u \tag{PC}$$

Will the participation constraint be binding? To see this intuitively ask yourself what would happen if it *didn't* bind. Just by looking at the constraint above it's clear that in that case the principal could simply raise R a little bit just until the constraint binds. There is no reason to leave any extra rent to the agent, so we know the participation constraint will bind.

Aside: Lagrangean Method Revisited

Another, more mechanical way of seeing that this constraint has to bind is to go back to the section on lagrange multipliers and to re-write the problem as:

$$L = \frac{s(1-s)}{c} + R - \lambda\big(u - \frac{s^2}{2c} + R + \frac{r_T}{2}s^2\sigma^2\big) \tag{2.8}$$

Taking first order condition with respect to s and R:

$$\frac{(1-2s)}{c} - \lambda\big(-\frac{s}{c} + r_T s\sigma^2\big) \quad = 0 \tag{2.9}$$

$$1 - \lambda = 0 \tag{2.10}$$

We see immediately that $\lambda = 1$ and hence the participation constraint must bind. It's the only constraint subject to which the principal maximizes his payoff so it will bind.
What would happen if we added a limited liability constraint to the problem? Would we still be sure that the participation constraint binds? The answer is no, we would no longer know which constraint binds and we would have to look at two separate cases again.

$$L = \frac{s(1-s)}{c} + R - \lambda\big(u - \frac{s^2}{2c} + R + \frac{r_T}{2}s^2\sigma^2\big) - \mu(w - R) \tag{2.11}$$

with first order conditions:

$$\frac{(1-2s)}{c} - \lambda\big(-\frac{s}{c} + r_T s\sigma^2\big) \quad = 0 \tag{2.12}$$

$$1 - \lambda - \mu = 0 \tag{2.13}$$

And it becomes obvious that either both or one of the constraints must bind now.

Knowing that the participation constraint binds, we can simply substitute $R = \frac{s^2}{2c} -$

$\frac{r_T}{2}s^2\sigma^2 - u$ into the objective function, giving

$$\frac{s(1-s)}{c} + \frac{s^2}{2c} - \frac{r_T}{2}s^2\sigma^2 \tag{2.14}$$

With first order condition with respect to s:

$$\frac{(1-2s)}{c} + \frac{s}{c} - r_T s\sigma^2 = 0 \tag{2.15}$$

and solving for s yields the optimal s^*

$$s^* = \frac{1}{1 + cr_T\sigma^2}$$

3 Adverse Selection

This chapter turns to the problem of hidden information. The previous chapter on moral hazard occupied itself with the problem of not being able to observe what the agent does. Adverse selection, on the other hand, is the problem of not being able to observe what type of agent the principal is facing.

In particular, we are going to look at adverse selection in the credit market and the various problems that can arise from it. Imagine there are two types of borrowers, some of whom are inherently risky and some of whom are inherently safe. If the principal (the lender) had perfect information then he could charge different rates to different kinds of borrowers. We will see in this chapter what happens if the lender is not able to do this.

3.0.2 Set up

All agents have access to a project which requires exactly one unit of capital to undertake. However p_r and produces R_r and other agents have access to a low risk/low yield project which succeeds with probability p_s and produces R_s. The high risk project succeeds less often ($p_r < p_s$) but if it does succeed it produces more ($R_r > R_s$). Additionally we are going to assume that *in expectation* both technologies produce just as much. That is we assume $R_r p_r = R_s p_s = \overline{R}$. A fraction θ of the population has access to the risky project while the remaining $(1 - \theta)$ have access to the safe project.

Both types of agents have zero wealth so to be able to undertake their projects, they will have to borrow capital from a lender. We will also assume that borrowers have no collateral whatsoever so that in case the project fails, borrowers will not be able to repay anything. The lender faces an opportunity cost of capital equal to ρ and we assume the credit market is completely competitive.

Before we solve this problem, let's take note of a couple of definitions that we will use a lot in the following sections.

Definition 4 *A **pooling equilibrium** is an equilibrium where all agents, regardless of type, accept the same contract*

Definition 5 *A **separating equilibrium** is an equilibrium where different types of agents accept different contracts*

3.0.3 The First Best

As a benchmark, we first work out what would happen if the lender had complete information about the which agent is risky and which agent is safe. The lender, therefore, has the option to charge different interest rates to different types of borrowers. The expected return from lending to a borrower of type i where $i \in \{r, s\}$ is given by

$$\Pi_i = p_i r_i - \rho \tag{3.1}$$

A safe borrower for example will repay r_s with probability p_s, while a risky borrower will only repay with a probability of p_r. The zero profit condition of the lender requires that $\Pi_i = p_i r_i - \rho = 0$ for risky and safe borrowers. Hence the first best, full information interest rates are given by

$$r_s = \frac{\rho}{p_s} \quad \& \quad r_r = \frac{\rho}{p_r} \tag{3.2}$$

The risky borrower therefore is charged a higher interest rate. Since the safe borrower can repay more often, he will be charged a lower interest rate, so that the lender makes zero profit from lending to either type.

To be able to tell what kind of projects will receive funding in this setting, we also need to look at the borrower's *participation constraint*. A borrower of type i will only accept a credit contract stipulating interest rate r_i if

$$p_i(R_i - r_i) \geq u \tag{PC}$$

where u is the borrower's outside option. So which projects will be funded? Plugging in the interest rate we found above, we see that all socially viable projects such that $p_i R_i > \rho + u$ will be financed and there is no credit rationing.

3.0.4 The Second Best

This section will demonstrate that even with a perfectly competitive lending market, credit rationing can occur. This is the problem known as *underinvestment* as first demonstrated by Stiglitz and Weiss (1981). We now assume that the lender knows the relative proportions of risky and safe agents, he knows their respective likelihoods of success and the expected payoffs. The only thing he does not know is which agent is risky and which agent is safe: he cannot tell them apart.

Pooling equilibrium

The lender can charge one interest rate, $\overline{r}$ for both types. The average probability of success is given by the weighted average of the probability of success of both types $\theta p_r + (1-\theta)p_s$. In order for the lender to break even we need

$$\Pi = (\theta p_r + (1-\theta)p_s)\overline{r} - \rho = 0 \tag{3.3}$$

$$\Rightarrow \overline{r} = \frac{\rho}{(\theta p_r + (1-\theta)p_s)} \tag{3.4}$$

Turning to the participation constraint of the two types of borrowers we see that no longer all economically viable projects can be financed. The safe type's project has to be sufficiently productive so that it is still worth while to the safe type to cross subsidise the risky type.

$$p_s(R_s - \overline{r}) \geq u \tag{PC-S}$$

$$p_r(R_r - \overline{r}) \geq u \tag{PC-R}$$

Since $p_r < p_s$ we know that if the participation constraint of the safe type is satisfied, the participation constraint of the risky type must be satisfied also. Solving for the minimum $p_s R_s$ such that save types will still borrow, we get

$$p_s R_s \geq u + \rho \frac{p_s}{\theta p_r + (1-\theta) p_s} \tag{3.5}$$

Comparing this directly to the first best, we see that for all safe projects in this range the credit market fails.

$$p_s R_s \in \left[u + \rho, u + \rho \frac{p_s}{\theta p_r + (1-\theta) p_s}\right] \tag{3.6}$$

This is the range of projects which are socially efficient, but which are excluded form the credit market because the pooled interest rate is too expensive to meet the borrower's participation constraint.

Separating Equilibrium

If a pooling equilibrium is not possible, which as was shown above is possible, the only option is a separating equilibrium in which the safe type is completely excluded from the credit market. In this case the lender will lend to the risky type at the break even interest rate of $r_r = \frac{\rho}{p_r}$ and the safe will not be willing to take out a loan at this rate.

3.0.5 Overinvestment

The previous section demonstrated that adverse selection in the credit market can lead to underinvestment because not all socially profitable projects have access to credit.However, that result depended in large part on our assumption that, on average, both types were equally productive ($p_r R_r = p_s R_s = \overline{R}$).

A different, equally interesting, assumption one might make is that both projects produce the same output R if they succeed but some agents succeed more often (p_s) than others ($p_r < p_s$). Under these conditions de Mezza and Webb (1987) show that this will in fact lead to *overinvestment*.

First Best

The first best, full information scenario is identical to that in the previous section. Each type is charged according to their particular risk adjusted interest rate such that the lender just breaks even

$$r_s = \frac{\rho}{p_s} \quad \& \quad r_r = \frac{\rho}{p_r} \tag{3.7}$$

Again, both risky and safe agents will only choose to borrow at these rates of interest if it is socially efficient to do so, that is if $p_i R > \rho + u$.

Second Best

If the lender cannot tell the different types of borrowers apart he again faces the pool of borrowers with average probability of sucess $\theta p_r + (1-\theta)p_s$ and lends at the average risk adjusted interest rate of

$$\overline{r} = \frac{\rho}{(\theta p_r + (1-\theta)p_s)} \tag{3.8}$$

The participation constraint of the borrower is now given by

$$p_s(R - \overline{r}) \geq u \tag{PC-S}$$

$$p_r(R - \overline{r}) \geq u \tag{PC-R}$$

Where the only difference to the previous section is the fact that they both produce the same R in case of success. However, unlike the previous section, now it is the case that if the participation constraint of the risky is satisfied then it must also be the case that the participation constraint of the safe is satisfied. The opposite was the case before. We therefore now focus on the participation constraint of the risky type. Plugging in $\overline{r}$

gives that the risky type will want to take out credit as long as

$$p_r R > \rho \frac{p_r}{\theta p_r + (1-\theta) p_s} + u \tag{3.9}$$

Given that $\theta p_r + (1-\theta)p_s) > p_r$ we know that the threshold given by the above participation constraint is lower than the socially viable threshold of $p_r R > \rho + u$. Projects that fall in the range

$$p_r R \in [\rho \frac{p_r}{\theta p_r + (1-\theta) p_s} + u, \rho + u]$$

will be financed even though they are too risky to be socially viable. This is the problem of overinvestment. The inability to tell risky borrowers apart from safe borrowers implies that safe borrowers cross-subsidize the risky ones. Therefore, risky borrowers might invest when it is not socially optimal to do so.

4 Game Theory

4.1 Introduction

Game theory is used to study the decision making of rational players, taking into account the decision of other rational players around them. The aim of this section is to equip the reader with the most basic knowledge of game theory required to understand applied theory. The term *game* is used to describe the setting in which the players interact. The *actions* available to the player are all the possible choices he can choose from. A *strategy* is a function that assigns an action to every possible history. A *history* consists of all the actions that have been played so far.

4.2 Static Games

A static game is a game in which all players choose their actions simultaneously and it's most common solution concept is the *Nash Equilibrium*.

Definition 6 *A Nash Equilibrium is an action profile such that no player can gain by changing his action given the actions of all other players*

Consider the case of two players who both have two actions available to choose from. For example two people wish to go to the cinema together to watch either 'Terminator' or 'Cinderella'. Their main concern of both players is to see the movie together, but in fact both would rather watch 'Terminator'. However, unfortunately the two have no means of communicating. Each can either go to the cinema which shows 'Terminator' or the cinema which shows 'Cinderella'. If they show up at different cinemas, they both derive zero utility. The game is depicted in normal form in figure 5.1.

		Player 2	
		Terminator	Cinderella
Player 1	Terminator	2, 2	0, 0
	Cinderella	0, 0	1, 1

Figure 4.1: **Normal form game**

Returning to the definition of Nash Equilibrium, we see that there are two Nash Equilbria in pure strategies. Both (Terminator, Terminator) and (Cinderella, Cinderella) are Nash Equilibria. Although both players would prefer to watch Terminator, nothing rules out the equilibrium that they will end up watching 'Cinderella'. For example imagine two American tourists in the UK driving past each other on the motorway. Ideally both would like to drive on the right hand side and this would yield the highest utility to both of them. However being in the UK, they stick to driving on the left hand side and while this might slightly reduce their payoff relative to both driving on the right hand side, it is nonetheless and equilibrium.

4.3 Dynamic Games

In many situations choices are, however, not made simultaneously but one after the other. For example how would the game above change if player one got to decide first and then player two could either decide to follow or to go to the other cinema? Such dynamic games are modeled in *extensive form*, also known as a *game tree* as in figure **??**.

The Nash Equilibria of this game are exactly the same as the Nash Equilibria of the normal form game. If player 2 threatens to go to watch 'Cinderella' no matter what player 1 does, then player $1'$s optimal response is to choose 'Cinderella' as well.

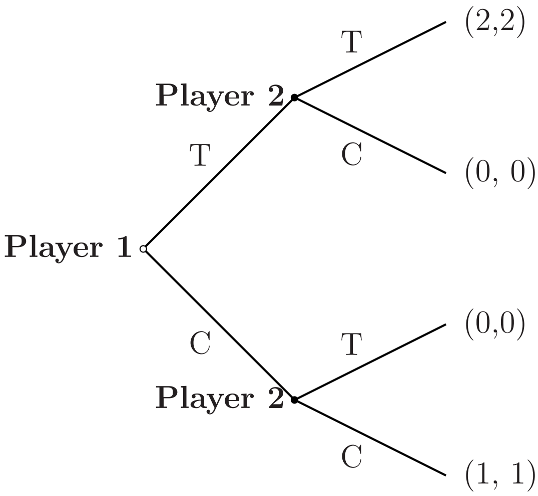

Figure 4.2: **The Game in Extensive Form**

However something about this equilibrium is not quite credible. If player 1 chose ‘Terminator’, then it is not in player 2′s interest any longer to go watch ‘Cinderella’. The threat is *not credible*. To formalize this notion of non-credible threats, we introduce the solution concept of subgame perfection which is a refinement of the Nash Equilibrium.

Definition 7 *A **Subgame Perfect Nash Equilibrium (SPNE)** is a strategy profile which is a Nash Equilibrium for every subgame of that game*

Importantly this definition also includes subgames that are never played in equilibrium. A SPNE must be a Nash Equilibrium for every single subgame (branch) of the overall game.

The way to solve for a SPNE is by *backwards induction*. Solving the above game by backwards induction, we begin at the final node and eliminate the strategies which are dominated for player 2. This leaves player 2 playing ‘C’ if player 1 played ‘C’ and playing ‘T’ if player 1 played ‘T’. Player 1 anticipates this and therefore can choose between the equilibria (C,C) and (T,T). Preferring (T,T) himself, he will choose to play ‘T’ in the first period.

To conclude, both (C,C) and (T,T) are Nash Equilibria in this game, but only (T,T) is a SPNE.

4.4 Repeated Games

To examine long term interaction of players, we study the model of repeated games. It captures the idea that a player will take into account the effect of his current actions on future payoffs.

To study a repeated game, we first look at the famous prisoner dilemma in a single stage game. The payoffs are in the matrix in figure 5.3 and the only Nash equilibrium is (D,D). While this is clearly pareto-dominated by (C,C), players will always benefit from choosing 'D' no matter what the other player does. The fact that players benefit from playing 'D' *regardless* of what the other player does, makes the rational behind this equilibrium particularly strong.

		Player 2: D	Player 2: C
Player 1	D	-1, -1	2, 0
	C	0, 2	1, 1

Figure 4.3: **The Prisoner's Dilemma**

However what happens if players get to interact in this setting for more than one period? While in a given period player 1 say, would benefit from playing 'D', the idea behind repeated games is that in the long run, both players would prefer to be playing (C,C).

The One Period Deviation Principle

The task of checking whether or not a particular strategy is subgame perfect in an infinitely repeated game might appear daunting. However, it turns out that it is *sufficient* to check if a single period deviation is profitable. It's obvious that if a player finds it profitable to deviate from a strategy in any given period, then the strategy is not a Nash Equilibrium. However it turns out that this is all we have to check and we do not have to worry about more complicated strategies.

To see why, imagine a strategy that satisfies the One Period Deviation Principle (OPDP) but is not subgame perfect in a finitely repeated game. Call this strategy s. Then there must be another strategy (call it s') that from some point during the game onwards (call it period t) will be better for one player. Strategy s' differs from strategy s for a finite number of periods because the entire game is finite (call the maximum period for which they differ t'). We can now construct another strategy (call it s'') that consists of s from t' onwards and s' before. From t' onwards, s'' agrees with s one period more than s'. Since s satisfies the OPDP so does s'' from t' onwards. Therefore s is as good as s' from t' on. Since s'' is otherwise the same as s' from time t onwards it is better than s' from t onwards. Similarly construct s''' and show that it is better than s''. Proceeding like this we obtain a contradiction. Thus if s satisfies the OPDP it is a SPE.

The same result holds also for infinite horizons, as long as players discount the future. The only difference is that for infinite games, we now also have to consider if infinite deviations can do better. In essence because of discounting, payoffs far in the future do not matter and we can use the same argument as above.

An Example

Returning back to our prisoner's dilemma above, we can now ask ourselves under what conditions there will be a SPNE in which players will choose 'C' instead of 'D' in an infinitely repeated game. All we have to check is whether or not a deviation in one single period is profitable. Assume that players discount the future by $\delta < 1$.

The harshest punishment available for deviating from the (C,C) equilibrium is to play (D,D) forever. This is known as the *grim trigger strategy* and it always consists of playing the one period Nash Equilibrium forever after one player played anything other than (C,C).

The payoff from cooperating is:

$$1 + \delta + \delta^2 + \delta^3 + \ldots = \frac{1}{1-\delta} \tag{4.1}$$

The payoff from deviating is:

$$2 - \delta - \delta^2 - \delta^3 + \ldots = 2 - \frac{\delta}{1-\delta} \tag{4.2}$$

Comparing these two payoffs we find that cooperation is preferable to deviating iff:

$$1 \geq 2(1-\delta) - \delta \Rightarrow \quad \delta \geq \frac{1}{3} \tag{4.3}$$

If the two players are sufficiently patient, that is if $\delta \geq 1/3$, then (C,C) is indeed a SPNE in the repeated game, although it would not be a Nash Equilibrium in the stage game.

5 Asset Ownership and the Hold Up Problem

In this section, we discuss the role of property rights in assigning ownership to an asset to maximize its productive potential. Our discussion of these issues is based on the literature on the property rights approach to the theory of the firm developed in Grossman and Hart (1986) and Hart and Moore (1990).

Consider two individuals, A and B, who undertake investments e_A and e_B that, in combination with the asset, generate returns $a\sqrt{e_A} + b\sqrt{e_B}$. The costs of these investments to A and B are, respectively, e_A and e_B. The terminology "investment" here as opposed to "effort" in the earlier section on moral hazard emphasizes the durability of the activity. We have in mind that the effort undertaken by each party creates something which is potentially of value to the future output from the asset even if the party who makes the decision is separated from the asset.

The first-best levels of these investments are:

$$e_A^* = \frac{a^2}{4} \text{ and } e_B^* = \frac{b^2}{4}. \tag{5.1}$$

The associated total surplus is:

$$S^* = \frac{1}{4}(a^2 + b^2). \tag{5.2}$$

Without any contracting problems, ownership does not have allocative implications, i.e. a contract can be written in which investment levels $\{e_A^*, e_B^*\}$ are prescribed.

The key insight of the property rights approach is that ownership matters due to contractual incompleteness. In this example, the owner has some bargaining power as he can threaten to exclude the other party from using the asset (i.e., he can "fire" the other party and exclude him from the returns from his investment). Ownership is now different from residual claimancy of a profit stream: it is the residual control right over the asset.

If the owner of an asset rents it out to someone, the tenant has residual claimancy. However, the owner retains the right not to renew the lease. This will potentially affect the incentive of the tenant to improve the asset. It is these residual control rights that give the owner a bargaining advantage over the non-owner.

As we shall see, this improves investment incentives for the owner while worsening them for the tenant. The optimal assignment of ownership takes into account how important are the investment decisions of each party and how severe is the hold up problem from having each party not owning the asset. The term hold-up here refers to the fact that the owner can limit the value of an investor's input to the project by firing him ex post.

To illustrate these arguments more precisely, assume that e_A and e_B are observable but non-verifiable. The last of these assumptions implies that a court could not enforce stipulated effort levels as it would be impossible to verify whether they were implemented. Thus investment levels are non-contractible *ex ante*. The two parties are assumed to bargain over the *ex post* surplus once it has been created.

Suppose first that party A is owner. Then at the bargaining stage, he has the right to fire B. Let $\overline{u}_i^j$ denote the disagreement payoff or outside option of i when j is the owner. We assume that even if A fires B at the bargaining stage, he can still make some use of the results of B's investments. Specifically, a fraction λ of the investment remains to be exploited by A in B's absence. It is useful to think of λ as measuring the extent of asset specificity. In the model of the previous section where e_B would be generic "effort" then $\lambda = 1$. However, where there is something special about B's human capital which requires his continued involvement in the project to make the most of it, then $\lambda < 1$.

Putting this together, the outside options of the two parties are $\overline{u}_A^A = a\sqrt{e_A} + \lambda b\sqrt{e_B}$ and $\overline{u}_B^A = \overline{u}_B$ where $\overline{u}_B$ is the exogenously given level of the disagreement payoff of B. Using the symmetric Nash bargaining formula,[1] the *ex post* payoff of A is:

$$\frac{1}{2}\left(a\sqrt{e_A} + b\sqrt{e_B}\right) + \frac{1}{2}\left(\overline{u}_A^A - \overline{u}_B^A\right) \tag{5.3}$$

which simplifies to $a\sqrt{e_A} + \frac{1}{2}(1+\lambda)b\sqrt{e_B} - \frac{\overline{u}_B}{2}$. Similarly, the *ex post* payoff of B is

$$\frac{1}{2}\left(a\sqrt{e_A} + b\sqrt{e_B}\right) + \frac{1}{2}\left(\overline{u}_B^A - \overline{u}_A^A\right) \tag{5.4}$$

which in turn simplifies to $\frac{1}{2}(1-\lambda)b\sqrt{e_B} + \frac{\overline{u}_B}{2}$. The two parties will choose e_A and e_B at the *ex ante* stage anticipating the ex post payoffs derived above. As a result the optimal choice of these variables are

$$\hat{e}_A^A = \frac{a^2}{4} \text{ and } \hat{e}_B^B = \frac{b^2(1-\lambda)^2}{16}. \tag{5.5}$$

This yields a second-best net expected surplus of

$$\hat{S}^A = \frac{a^2}{4} + \frac{b^2}{16}(1-\lambda)(3+\lambda). \tag{5.6}$$

This is less than the first best surplus S^*.[2] Since B anticipates that, after the investments are made, he will be at the mercy of A, he invests less than the first best level. The higher is λ, the less costly it is for A to fire B and the greater is the incentive problem of B. However, if $\lambda = 1$, there is full exploitation of B's output and he does not invest at all. If we think of λ as representing the extent of specialized skills, then economies with greater skill intensity will suffer a smaller efficiency loss through this

[1] This is standard in the literature following Grossman and Hart (1986). The transfers solve:

$$\begin{aligned} t^* &= \arg\max\{\left(a\sqrt{e_A} + b\sqrt{e_B} - \bar{u}_A^A - t\right) \\ &\quad \times \left(a\sqrt{e_A} + b\sqrt{e_B} - \bar{u}_B^A + t\right)\}. \end{aligned}$$

[2] Observe that $(1-\lambda)(3+\lambda)$ is strictly decreasing in λ. Also, it takes the value 3 for $\lambda = 0$ and the value 0 for $\lambda = 1$. Hence, $\frac{b^2}{16}(1-\lambda)(3+\lambda) \leq \frac{3b^2}{16} < \frac{b^2}{4}$ implying that $\hat{S}^A$ is less than S^*.

effect than those which only have generic labor input. There are symmetric expressions if B is the owner. We now use $\mu \in [0,1]$ to be the investment specificity parameter analogous to λ. By a similar analysis we find:

$$\hat{e}_A^B = \frac{a^2}{16}\left(1-\mu^2\right) \text{ and } \hat{e}_B^B = \frac{b^2}{4}. \tag{5.7}$$

Second best surplus (also less than S^*) is:

$$\hat{S}^B = \frac{a^2}{16}(1-\mu)(3+\mu) + \frac{b^2}{4}. \tag{5.8}$$

As before, a larger value of μ induces a greater efficiency loss, all else equal.

We can now which party should own the asset to maximize economic efficiency (measured by total surplus) as a function of the key parameters: a, b, λ, and μ. Comparing (4.6) and (4.8), we find that A should owner the asset if if $a^2(1+\mu)^2 > b^2(1+\lambda)^2$, while B should own it otherwise. We state this finding as:

Result 8 *If the marginal return of A's (B's) investment is greater than that of B's (A's) or his investment is more asset-specific than Bs (A's), under the efficient assignment of property rights, A (B) should own the asset.*

This theory of the "optimal" allocation of property rights can be thought of as reflecting two dimensions of the skill of the investors. The parameters a and b reflect their relative productivities as investors with a presumption that the most productive should own the asset. But the specificity of their skills matters too. If one investor has a very specialized skill so that replacing him would lead to a major loss in output, then it is best that he owns the asset. If not, the investment process is more prone to hold-up. So if one party supplies generic effort which stays with the project whether or not he leaves, he will generally not optimally be the owner.