Unit 8.4: Moral Hazard and the Principal-Agent Problem

Michael Malcolm

June 18, 2011

1 Adverse Selection versus Moral Hazard

Recall that both adverse selection and moral hazard deal with asymmetric information. While adverse selection deals with an unobserved characteristic before the transaction takes place (e.g. a buyer who doesn't know whether a used car is high or low quality), moral hazard deals with an unobserved action that typically manifests itself *after* a contract is written.

2 Examples of Moral Hazard

Moral hazard is a problem when two parties to a transaction have incentives that are not aligned, and when one party can take some unobserved action that is detrimental to the other party. Here are a few examples.

- People with good car insurance take more driving risks than they might have otherwise. This is detrimental to the insurance company.
- A worker who has a long-term contract with a firm stops working as hard as he might have otherwise. This is detrimental to the firm.
- People with good health insurance visit the doctor more often and for smaller reasons that they might have otherwise. This is detrimental to the insurance company.
- A mechanic performs needless repairs on a car to increase his profit. This is detrimental to the customer.

In all these cases, the problem is that the interests of the two parties to the contract are not aligned. A driver wants to get to his destination quickly, while the insurance company wants him to drive safely. The worker prefers not to put such high effort into his job, while the employer wants him to work hard.

If the action in question were perfectly observable, then there would be no problem since the contract could be written contingent on this – the insurance company will cover only safe drivers, the worker will only be paid if he is working hard, health insurance will not pay for unnecessary doctor visits and a mechanic's customers will pay only for needed repairs.

The problem with moral hazard is precisely that the action in question is unobserved. The insurance company can't perfectly determine how safely you're driving and the owner can't perfectly determine how hard the employee is working. Often, though, there is a signal about this action, even if it is not perfect. For example, a driver who has many accidents is not *necessarily* a dangerous driver (maybe he's just unlucky), but it is at least an indicator. Similarly, the manager of a restaurant that produces low profit is not *necessarily* being lazy (maybe the economy is just bad), but again it is at least an indicator.

The solution is to write a contract that is contingent upon these indicators. The contract will not be perfect, in the sense that it might sometimes reward agents who take the "wrong" action or punish agents who take the "right" action. For example, raising insurance rates for drivers who have bad accidents will sometimes punish safe but unlucky drivers, and will sometimes reward lucky drivers who drive dangerously. Cutting the pay of a manager whose restaurant earns low profit might sometimes punish hard-working

managers whose restaurants had a bad year and might sometimes reward lazy managers whose restaurants are in a good location. However, often the indicator will be good enough to incentivize the correct action.

3 Principal-Agent Contracting

Here is a simple numerical example of a moral hazard problem.

An owner (the *principal*) employs a manager (the *agent*) to run a restaurant. The manager can exert high effort or low effort in running the restaurant. Putting in high effort costs 2 (i.e. it deducts 2 from his utility), while putting in low effort costs nothing. Effort is unobservable to the owner. The manager could get utility of 10 from his best outside option.

The restaurant can be profitable or unprofitable, and this *is* observable to the owner:

- If the manager puts in high effort, then the restaurant is profitable with probability 0.8 and unprofitable with probability 0.2.
- If the manager puts in low effort, then the restaurant is profitable with probability 0.4 and unprofitable with probability 0.6.

This is the usual structure of moral hazard problems. The restaurant's profit is tied to how hard the manager works, but not perfectly. High effort managers sometimes have unprofitable restaurants and low effort managers sometimes have profitable restaurants. If effort itself were observable, then the contract would simply be contingent on effort and there would be no problem.

Even though effort is unobservable, restaurant profitability is a signal and can be used to design a contract that spurs the manager to put in high effort. Consider a contract that pays the manager w_p when the restaurant is profitable and w_u when the restaurant is unprofitable.

Suppose that the owner is interested in writing a contract such that the manager will put in high effort. There are two constraints. The *participation constraint* states that a high-effort manager needs to obtain an expected utility greater than 10 by working hard, since this is the utility from his outside option. Now, by putting in high effort he gets w_p with probability 0.8 when the restaurant is profitable, but w_u with probability 0.2. Also, we need to deduct 2 for the cost of high effort.

$$\begin{aligned} EU_{high} &\geq 10 \\ 0.8w_p + 0.2w_u - 2 &\geq 10 \end{aligned}$$

The second constraint is the *incentive constraint* which guarantees that the contract is written such that the agent prefers high effort over low effort. Putting in low effort changes the probability with which the agent receives w_p versus w_u, but it also costs 0 to put in low effort.

$$\begin{aligned} EU_{high} &\geq EU_{low} \\ 0.8w_p + 0.2w_u - 2 &\geq 0.4w_p + 0.6w_u - 0 \end{aligned}$$

To maximize profit, the owner will set both of these constraints to bind with equality. In other words, he wants to pay the lowest possible expected wage bill, while still encouraging high effort on the part of the manager. From the participation constraint:

$$\begin{aligned} 0.8w_p + 0.2w_u - 2 &= 10 \\ w_u &= 60 - 4w_p \end{aligned}$$

The incentive constraint will also be set to bind with equality:

$$0.8w_p + 0.2w_u - 2 = 0.4w_p + 0.6w_u - 0$$
$$0.4w_u = 0.4w_p - 2$$
$$w_u = w_p - 5$$

Equating the two expressions for w_u gives:

$$60 - 4w_p = w_p - 5$$
$$w_p = 13$$

Substituting back gives $w_u = 8$. So the optimal contract is that the manager should be paid $w_p = 13$ when the restaurant is profitable and $w_u = 8$ when the restaurant is unprofitable. This scheme is designed so that the manager will choose to put in high effort (the incentive constraint), and so that the manager prefers this to his best outside option (the participation constraint).

Again, notice that this contracting scheme is an imperfect way of monitoring effort. Even a manager putting in high effort will sometimes get paid $w_u = 8$ when the restaurant is unprofitable. However, the imperfect signal is used to write a contract such that incentives are aligned properly.

This is the structure of most contracts that involve moral hazard. I remember once I was doing private tutoring and an economics-minded student offered to pay me based on the grade he got in the course – not a perfect signal of my effort since he might get a low grade even when I put in a lot of effort in tutoring him. But it was enough of an incentive that I worked harder than I would have otherwise.