Contracts, Organizations, and Markets

Instructor: Kaniṣka Dam
Semester: Fall, 2023
Centro de Investigación y Docencia Económicas

CIDE

Preface

Many years ago, the original title of this course has been *Economics of Information and Uncertainty*. In recent years the title has changed to *Contract Theory and the Theory of Organizations*. I have decided to consciously maintain a balance between the two titles. This course analyzes optimal contracts offered in diverse organizations, which has been part of a larger research agenda of Mechanism design and implementation theory. I thus preferred to call the course *Contracts, Organizations, and Markets*. Apart from contract theory, we shall try to learn how to implement market outcomes via revelation mechanisms, which probably justifies the older title of the present course. Parts I and II are heavily borrowed from *Contract Theory* (Bolton and Dewatripont, 2005). In many parts of the notes, I have literally paraphrased texts from the aforementioned book. I made these notes principally for my own use to teach the course. They must be full of errors and typos. Also, interpretations of some results are my own, which may be misconstrued. So, as we go on, we surely will discover a huge room for improvement. Please let me know about all errors or any other comments you have at kaniska.dam@cide.edu.

Mexico City
Fall, 2023

Contents

Part I

Hidden action

Chapter 1

Moral hazard with single agent

1.1 Principal-agent problems: symmetric and asymmetric information

Throughout this and the subsequent chapters we will built on the following scenario. There are two classes of individuals in the economy: *principal* who may be an entrepreneur, a landlord, an investor, etc. and *agent* who may be a worker, a manager, a tenant, an entrepreneur, etc. When a principal and an agent enter into a contractual relationship, the agent has private information over the actions he may take (moral hazard) or over his type such as productivity, efficiency, etc. (adverse selection, see Chapter 3) which influence the (random) performance of the relationship. For instance, an entrepreneur hires a manager to work in her firm. The firm's performance (say, profit) depends on the effort made by the manager. When the effort levels can be observed by the principal, we are in a situation of *symmetric information*. On the other hand, when the effort choice of the manager cannot be verified by the principal, we are in a world characterized by *asymmetric information*.

Our objective is to determine the optimal managerial compensation under both situations. In general, under symmetric information, the optimal compensation is determined by solving a Pareto program where the principal maximizes her net expected payoff subject to the constraint that the compensation package guarantees a minimum expected payoff to the agent. Such constraints are called the *individual rationality* or *participation* constraint of the agent. This situation is equivalent to maximizing the aggregate surplus of the relationship, and is called the *first best* situation. When the manager's effort choice cannot be verified, a compensation contract cannot be based on the managerial effort. In addition to the participation constraint, one then needs to impose the *incentive compatibility* constraint which implies that although the effort of the agent cannot be contracted upon, the principal can anticipate that the agent will choose the effort level following some optimizing behavior. The optimal contracts obtained in this situation are called the *second best* contracts.

1.2 Moral hazard

Consider a contracting problem between a principal and an agent where the principal hires the agent to accomplish a task. The agent chooses an action e, e.g. effort, investment, production technique, etc. which influences the performance q of the task. The principal only cares about the performance. Effort is costly for the agent and he requires to be compensated. When effort is non-verifiable, the best the principal can do is to relate compensation to performance. Such compensation scheme entails a loss of efficiency since performance is only a noisy signal of effort. Such situations give rise to what is called a *moral hazard* problem. This is a very simplified scenario – the real world situation may be more complex. In general, a principal-agent relationship faces a moral hazard problem if the parties are unable to contract upon some decision variable(s). Following are some examples of moral hazard problem.

Example 1.1: Landlord-tenant

A landlord (principal) hires a tenant (agent) to cultivate a plot of land. The crop yield depends on the effort e exerted by the tenant and a random variable ε:

$$q = e + \varepsilon,$$

where $\varepsilon \sim N(0, \sigma^2)$. The tenant's effort e is not verifiable. Effort is costly to the tenant. Let $\psi(e)$ is the cost function of effort which is strictly increasing and convex in e. The tenant faces the following trade-off. An increase in the effort level increases q, but it also raises his disutility.

Example 1.2: Firm-manager

A firm's (principal) constant marginal cost of production θ may take a high value θ_H or a low value θ_L. Initially, the firm is inefficient, i.e., $\theta = \theta_H$, and hires a manager (agent) to carry out R&D activities to make the production process more efficient, i.e., reduce θ to the lower level θ_L. Let the manager chooses an R&D effort e which lowers the firm's marginal cost with a probability $p(e)$ such that $p''(e) < 0 < p'(e)$. The effort cost function is given by $\psi(e) = e$. Effort is not observable by the firm.

Example 1.3: Investor-entrepreneur

A penny less startup firm (agent) has a setup cost of \$1 which must be raised from an external investor (principal). After the investor agrees to invest, the firm may choose one of the two projects P^1 and P^2. Choice of project is not verifiable by the investor. Each project P^i has a verifiable income y_i and a non-verifiable private benefit b_i. Assume that $y_1 > y_2$ and $b_1 < b_2$, i.e., from the viewpoint of the investor P^1 is a good project and P^2 is a bad project. By incurring a non-verifiable monitoring cost $\phi(m)$ with $\phi', \phi'' > 0$ the investor can make sure that P^1 is undertaken with probability m. There are moral hazard problems both in the choices of project and monitoring.

1.3 Timing of events

The principal-agent relationship lasts for five dates, $t = 0, 1, 2, 3, 4$. At date 0, the principal offers a contract. At $t = 1$, the agent accepts or rejects the contract. At $t = 2$, the agent exerts effort. At $t = 3$, the output is realized. Finally, at date 4 the contract is executed.

1.4 Optimal contracts with two performance outcomes

Consider a situation where performance may take only two values: $q \in \{q_H, q_L\}$ with $q_H > q_L \geq 0$. When $q = q_H$ the agent's performance is a "success", and it is a "failure" otherwise. Let the probability of success be given by Prob.$[q = q_H \mid e] = p(e)$ with $p''(e) \leq 0 < p'(e)$ where $e \in [0, 1]$ is the effort chosen by the agent. Assume that $p(0) = 0$ and $p'(0) > 1$. The principal's utility function is given by $V(q - w)$ with $V'' \leq 0 < V'$, and the agent has utility function $U(w, e) = u(w) - \psi(e)$ with $u'' \leq 0 < u'$, $\psi' > 0$, and $\psi'' \geq 0$. Assume that $\psi(e) = e$.

1.4.1 First-best contracts

Suppose that the agent's effort choice is verifiable, and hence the compensation scheme can be made contingent on this choice. Let the compensation scheme be $w_\theta = w(q_\theta)$ for $\theta = H, L$. Then the principal solves the following maximization problem:

$$\max_{\{e,w_H,w_L\}} \; p(e)V(q_H - w_H) + [1-p(e)]V(q_L - w_L) \tag{M$_1$}$$

$$\text{subject to} \quad p(e)u(w_H) + [1-p(e)]u(w_L) - e \geq \bar{u}, \tag{IR}$$

where $\bar{u}$ is the agent's outside option. Assume without loss of generality that $\bar{u} = 0$. Let λ be the Lagrange multiplier associated with (IR). The first order conditions with respect to w_H and w_L yield the following optimal coinsurance, or so-called *Broch rule*:

$$\frac{V'(q_H - w_H)}{u'(w_H)} = \lambda = \frac{V'(q_L - w_L)}{u'(w_L)}\,. \tag{1.1}$$

Given that $V' > 0$ and $u' > 0$, we have $\lambda > 0$, i.e., the (IR) constraint binds at the optimum. The first order condition with respect to effort is given by:

$$p'(e)[V(q_H - w_H) - V(q_L - w_L)] + \lambda p'(e)[u(w_H) - u(w_L)] = \lambda, \tag{1.2}$$

which, together with the Borch rule, determines the optimal effort e. Why Borch rule implies Pareto optimality? Let $x_\theta := q_\theta - w_\theta$, the principal's income at state $\theta = H, L$. Notice that the slopes of the indifference curves of the principal and the agent are respectively given by:

$$\frac{dx_L}{dx_H} = -\frac{p(e)}{1-p(e)}\,\frac{V'(x_H)}{V'(x_L)}\,,$$
$$\frac{dw_L}{dw_H} = -\frac{p(e)}{1-p(e)}\,\frac{u'(w_H)}{u'(w_L)}\,.$$

Hence, (1.1) implies tangency of the indifference curves, i.e., Pareto optimality. Notice a few important properties of the set of Pareto optimal solutions.

(a) At the optimum, it is always the case that (i) $x_H \geq x_L$ and (ii) $w_H \geq w_L$. To show (i), suppose on the contrary that $x_H < x_L$. This inequality implies that $w_H - w_L > q_H - q_L > 0$. On the other hand, since $V''(\cdot) \leq 0$, we have $V'(x_H) \geq V'(x_L)$ for $x_H < x_L$. Then it follows from the Borch rule that $u'(w_H) \geq u'(w_L)$ which implies $w_H < w_L$ which is a contradiction. In a similar fashion one can prove (ii). Also, conditions (i) and (ii) together imply $0 \leq x_H - x_L \leq q_H - q_L$ and $0 \leq w_H - w_L \leq q_H - q_L$. Full insurance occurs when either (a) $w_H - w_L = 0$ and $x_H - x_L = q_H - q_L$, or (b) $x_H - x_L = 0$ and $w_H - w_L = q_H - q_L$. What do conditions (i) and (ii) mean. Consider an Edgeworth box for this principal-agent economy in which the horizontal axes represent incomes at state H and the vertical axes stand for those in state L. This property of the optimum says that the Pareto set lies between the two 45^0 lines, the full insurance lines for the principal and the agent respectively. Clearly, due to optimal risk sharing motives, both parties cannot be fully insured at the same time as long as $q_H > q_L$.

(b) Principal is risk neutral, i.e., $V(x) = x$. Then (1.1) implies that $w_H = w_L = w^*$, i.e., the optimum entails full insurance for the agent. And the optimal effort level e^* is given by:

$$u(w^*) = e^*.$$

Notice that, at the optimum, $p'(e^*)(q_H - q_L) = 1/u'(w^*)$, i.e., the marginal productivity of effort is equal to the marginal cost the the principal incurs to compensate the agent for his effort cost. In this case we have $x_H - x_L = q_H - q_L$, i.e., the risk neutral principal assumes the entire risk of output measured by $q_H - q_L$.

(c) Agent is risk neutral, i.e., $u(x) = x$. Then the optimal entails full insurance for the principal with $x_H = x_L = x^*$ and $w_H - w_L = q_H - q_L$. All risks due to output fluctuations are absorbed by the risk neutral agent.

1.4.2 Second-best contracts

When the agent's effort choice cannot be verified by the principal, the compensation contracts cannot be made contingent upon the agent's choice of effort, and the first-best contracts cannot be implemented. Then the agent's performance-contingent compensation induces him to choose an effort level e that maximizes his expected payoff:

$$p(e)u(w_H) + [1 - p(e)]u(w_L) - e \geq p(\hat{e})u(w_H) + [1 - p(\hat{e})]u(w_L) - \hat{e} \quad \text{for all } \hat{e} \in [0,\, 1]. \tag{IC}$$

The above constraint is called the *incentive compatibility constraint*. The first order condition of the maximization problem is given by:

$$p'(e)[u(w_H) - u(w_L)] = 1. \tag{IC$'$}$$

Since $p'(\cdot) > 0$, (IC$'$) implies that $u(w_H) > u(w_L)$, which in turn implies that $w_H > w_L$ since $u'(\cdot) > 0$. Let $\Delta \equiv u(w_H) - u(w_L)$. Then, from (IC$'$) it follows that

$$\frac{de}{d\Delta} = -\frac{p'(e)}{p''(e)\Delta} > 0. \tag{1.3}$$

Now, consider an increase in $w_H - w_L$, which can be achieved either (i) by increasing w_H and keeping w_L constant, or (ii) by decreasing w_L and keeping w_H constant, or (iii) by increasing w_H and decreasing w_L. Since $u'(\cdot) > 0$, in all the three cases Δ must increase. Therefore,

Lemma 1.1

At any incentive compatible contract, i.e., any $(w_L,\, w_H)$ that satisfies (IC$'$), we must have

(a) $w_H > w_L$;

(b) effort e is increasing in $w_H - w_L$.

The optimal second-best contracts $(w_H,\, w_L,\, e)$ are determined by maximizing (M_1) subject to (IR) and (IC$'$). Let μ be the Lagrange multiplier associated with the incentive compatibility constraint. The first order conditions with respect to w_H and w_L are respectively given by:

$$\frac{V'(q_H - w_H)}{u'(w_H)} = \lambda + \mu \left[\frac{p'(e)}{p(e)}\right], \tag{1.4}$$

$$\frac{V'(q_L - w_L)}{u'(w_L)} = \lambda - \mu \left[\frac{p'(e)}{1 - p(e)}\right]. \tag{1.5}$$

When $\mu = 0$, we obtain the Borch rule, and $(w_L,\, w_H) = (w_L^*,\, w_H^*)$ where $*$ denotes the first-best contracts. Since $\mu \geq 0$ and $p'(e) > 0$, (1.4) and (1.5) imply that

$$\frac{V'(q_H - w_H)}{u'(w_H)} \geq \frac{V'(q_H - w_H^*)}{u'(w_H^*)}, \quad \text{and} \quad \frac{V'(q_L - w_L)}{u'(w_L)} \leq \frac{V'(q_L - w_L^*)}{u'(w_L^*)}.$$

Notice that, given $V'' \leq 0$ and $u'' \leq 0$, the following function

$$H(w) := \frac{V'(q - w)}{u'(w)}$$

is increasing in w, and hence the above two inequalities imply that

Lemma 1.2

The second best contract has the property that $w_H \geq w_H^*$ and $w_L \leq w_L^*$, and hence, $w_H - w_L \geq w_H^* - w_L^*$.

In Section 1.5, we will prove that, at the optimum, $\mu > 0$ under very general conditions, and hence, the contract is distorted from the optimal risk sharing contract implied by the Borch rule. Moreover, from Lemma 1.1 we have $w_H > w_L$. Notice that, even if the principal is risk neutral, insuring fully the agent is not optimal. This occurs because, as opposed to the first-best situation, the principal faces a trade-off between risk sharing and incentive provision. It easily follows from the incentive constraint that the effort exerted by the agent is increasing in $w_H - w_L$ since $u' > 0$. Hence, in order to induce the agent to exert the highest possible incentive compatible effort, the principal would like to make the difference $w_H - w_L$ as large as possible. Clearly, full insurance would undermine such incentives. We will analyze further properties of the optimal second-best contracts in more general setting in the next section.

1.5 A general model of incentive contracts

Following Hölmstrom (1979), we now consider a general nonlinear incentive scheme. Let the firm's output q depends on some random variable describing possible contingencies and the effort of the agent, i.e., $q = q(\theta, e)$ where $\theta \in \Theta$ is the random variable and $e \in A \subseteq \mathbb{R}$. We assume that both Θ and A are compact sets. The principal may be risk averse and has a utility function $V(q - w)$ with $V'(\cdot) > 0$ and $V''(\cdot) \leq 0$. The agent is risk averse and has a separable utility function $u(w) - \psi(e)$ with $u'(\cdot) > 0$, $u''(\cdot) \leq 0$, $\psi'(\cdot) > 0$ and $\psi''(\cdot) \geq 0$. Let the conditional distribution and density of output are respectively given by $F(q \mid e)$ and $f(q \mid e)$ over the support $[q_{min}, q_{max}]$. Then the principal's problem can be written as

$$\max_{\{e, w(q)\}} \int_{q_{min}}^{q_{max}} V(q - w(q)) f(q \mid e) dq \tag{M$_2$}$$

$$\text{subject to} \quad \int_{q_{min}}^{q_{max}} u(w(q)) f(q \mid e) dq - \psi(e) \geq \bar{u}, \tag{IR}$$

$$e \in \text{argmax}_{\hat{e}} \left\{ \int_{q_{min}}^{q_{max}} u(w(q)) f(q \mid \hat{e}) dq - \psi(\hat{e}) \right\}. \tag{IC}$$

The first and second order conditions of the agent's maximization problem are given by:

$$\int_{q_{min}}^{q_{max}} u(w(q)) f_e(q \mid e) dq - \psi'(e) = 0, \tag{IC$_a$}$$

$$\int_{q_{min}}^{q_{max}} u(w(q)) f_{ee}(q \mid e) dq - \psi''(e) < 0. \tag{IC$_b$}$$

As we have been doing in the previous sections, we will replace (IC$_\theta$) by (IC$_a$) leaving aside for the time being the constraint (IC$_b$). The Lagrangean is therefore given by:

$$\mathcal{L} = \int_{q_{min}}^{q_{max}} \left[V(q - w(q)) f(q \mid e) + \lambda \left\{ u(w(q)) f(q \mid e) - \psi(e) - \bar{u} \right\} + \mu \left\{ u(w(q)) f_e(q \mid e) - \psi'(e) \right\} \right] dq$$

The first order necessary condition with respect to $w(q)$ is given by:

$$\frac{V'(q - w(q))}{u'(w(q))} = \lambda + \mu \frac{f_e(q \mid e)}{f(q \mid e)} \qquad \text{for all } q. \tag{1.6}$$

1.5.1 First best contracts

The first best contract is implemented when the agent's effort can be verified by the principal. In this case any effort level can be implemented, and the incentive compatibility constraint (IC_θ) can be ignored. This is equivalent to a situation where $\mu = 0$. Then the first order condition (1.6) reduces to the following Borch rule:

$$\frac{V'(q - w(q))}{u'(w(q))} = \lambda > 0 \qquad \text{for all } \ q. \tag{1.7}$$

Now we study the shape of the compensation schedule. Differentiating (1.7) with respect to q we get

$$-V''(1 - w'(q)) + \lambda u''\, w'(q) = -V''(1 - w'(q)) + \frac{V'}{u'}\, u''\, w'(q) = 0. \tag{1.8}$$

Define by $\eta_P = -V''/V'$ and $\eta_A = -u''/u'$ the coefficients of absolute risk aversion of the principal and the agent, respectively. Then Condition (1.8) becomes

$$w'_{FB}(q) = \frac{\eta_P}{\eta_P + \eta_A}\,. \tag{1.9}$$

Notice that the coefficients of risk aversion are not constant in general, and hence the equilibrium compensation function $w(q)$ is nonlinear.

Risk neutral principal

Assume that $V'(\cdot) =$ constant, i.e., the principal is risk neutral. Then Condition (1.7) requires that $u'(w(q)) =$ constant for all q, i.e., $u'(w(q)) = u'(w(q'))$ for $q \neq q'$. This occurs only if $w(q) = w(q') = w_{FB}$. Therefore, at the optimum the agent receives a fixed salary independent of the output of the firm. In other words, the principal assumes all the output risk and fully insures the agent against the variability of the output. Since $\lambda > 0$, the individual rationality constraint binds at the optimum, and hence

$$w_{FB} = u^{-1}\left(\bar{u} + \psi(e_{FB})\right), \tag{1.10}$$

where e_{FB} is the effort implemented at the compensation scheme w_{FB}.

Risk neutral agent

Now suppose that the agent is risk neutral instead. Then Condition (1.7) requires that $V'(q - w(q)) = V'(q' - w(q'))$ for $q \neq q'$. This is true only if $q - w(q) = q' - w(q')$ for any $q \neq q'$. In this case, the principal is fully insured as she sells the firm to the agent at a fixed price $q - w(q)$ and the agent becomes the residual claimant of the firm output. This type of contracts are called *franchise* contracts.

Risk averse principal and agent

When both the principal and the agent are risk averse, the optimal first best contract is characterized by Condition (1.6), and hence the optimal marginal compensation is given by:

$$w'_{FB}(q) = \frac{\eta_P}{\eta_P + \eta_A} \in (0,\, 1).$$

The above equation implies that following an improvement in the firm's performance, the agent receives only a part of the increased output as increased compensation. Thus, both the principal and the agent are partially insured against risk. The more risk averse is the agent, i.e., the greater is η_A, the less the firm's performance influences his compensation.

1.5.2 Second best contracts

It is intuitive that the principal would always like the agent to exert higher effort, and possibly, reward him for not being lazy. But the problem is that since effort is not verifiable, it cannot be contracted upon and the compensation of the agent can only be based on the firm's performance q. The important question is whether high effort is at all desirable as rewarding a very high effort level may be highly costly for the principal. We will assume that $F_e(q \mid e) \leq 0$ with strict inequality for some q. This is to say that given two effort levels $e < e'$, the distribution $F(q \mid e')$ *first order stochastically dominates* the distribution $F(q \mid e)$. Intuitively, in the two-performance case, i.e., $q \in \{q_H, q_L\}$ with $q_H > q_L$, Prob.$[q = q_H \mid e'] >$ Prob.$[q = q_H \mid e]$ if $e' > e$, i.e., higher effort makes 'success' more likely. In other words, higher effort shifts the density $f(q \mid e)$ to the right. The following is an important result related to first order stochastic dominance.

Lemma 1.3

Let $h : [q_{min}, q_{max}] \to \mathbb{R}$ is an increasing function, and $F_e(q \mid e) \leq 0$ for all e. Then for any two effort levels e and e' with $e > e'$, we have

$$E[h(q) \mid e] \geq E[h(q) \mid e'].$$

In particular, $E[q \mid e] \geq E[q \mid e']$.

Proof. Notice that

$$\begin{aligned}
E[h(q) \mid e] - E[h(q) \mid e'] =& \int_{q_{min}}^{q_{max}} h(q)[f(q \mid e) - f(q \mid e')]dq \\
=& h(q_{max}) - \int_{q_{min}}^{q_{max}} h'(q)F(q \mid e)dq - h(q_{max}) + \int_{q_{min}}^{q_{max}} h'(q)F(q \mid e')dq \\
=& - \int_{q_{min}}^{q_{max}} h'(q)[F(q \mid e) - F(q \mid e')]dq.
\end{aligned}$$

The above expression is positive since $h'(q) \geq 0$ and $F(q \mid e) \leq F(q \mid e')$.[1] $\square$

First order stochastic dominance is a stochastic order. The other popular stochastic orders are second order stochastic dominance, hazard rate dominance, likelihood ratio dominance (see Krishna, 2002, Appendix B for a nice discussion).

The optimal second best contracts are characterized by Condition (1.6) along with the binding individual rationality constraint (since $\lambda > 0$). When $\mu = 0$ we are back to the first best and (1.6) represents the pure risk sharing role of the contract, given by the Borch rule, which is meant to coinsure both the principal and the agent. When $\mu > 0$, the optimal contract, in addition to its risk sharing motive, is an instrument used for incentive provision. Thus, $\mu > 0$ implies the so-called tradeoff between risk sharing and incentive. We will first prove that $\mu > 0$ under very general conditions.

Lemma 1.4: Hölmstrom, 1979

Suppose that the principal has a strictly increasing utility function, i.e., $V' > 0$, and $F_e(q \mid e) \leq 0$. Then $\mu > 0$.

[1] For any effort level e, it follows from the integration by parts that

$$\int_{q_{min}}^{q_{max}} h(q)f(q \mid e) = \underbrace{[h(q)F(q \mid e)]_{q_{min}}^{q_{max}}}_{h(q_{max})\times 1 - h(q_{min})\times 0} - \int_{q_{min}}^{q_{max}} F(q \mid e)h'(q)dq.$$

Proof. Suppose on the contrary that $\mu \leq 0$. The first order condition of the principal's maximization problem with respect to e is given by:

$$\int_{q_{min}}^{q_{max}} \left[V(q - w(q)) f_e(q \mid e) + \lambda \left\{ u(w(q)) f_e(q \mid e) - \psi'(e) \right\} + \mu \left\{ u(w(q)) f_{ee}(q \mid e) - \psi''(e) \right\} \right] dq = 0.$$

Using the optimality conditions of the agent's maximization problem (IC$_a$) and (IC$_b$), note that $\mu \leq 0$ is equivalent to

$$\int_{q_{min}}^{q_{max}} V(q - w(q)) f_e(q \mid e) dq \leq 0. \tag{1.11}$$

call the first best compensation schedules of the principal and the agent $x_{FB}(q)$ and $w_{FB}(q)$, respectively, which are, by Borch rule, are non-decreasing in q with slopes everywhere less than 1. Now consider two subsets $Q_+ := \{q \mid f_e(q \mid e) > 0\}$ and $Q_- := \{q \mid f_e(q \mid e) < 0\}$ of $[q_{min}, q_{max}]$. Notice that, when $\mu \leq 0$, $w(q) \leq w_{FB}(q)$ for all $q \in Q_+$, and $w(q) \geq w_{FB}(q)$ for all $q \in Q_-$. Therefore,

$$\int_{q_{min}}^{q_{max}} V(q - w(q)) f_e(q \mid e) dq \geq \int_{q_{min}}^{q_{max}} V(q - w_{FB}(q)) f_e(q \mid e) dq. \tag{1.12}$$

However, since $F_e(q_{min} \mid e) = F_e(q_{max} \mid e) = 0$ for all e, integration by parts implies

$$\int_{q_{min}}^{q_{max}} V(q - w_{FB}(q)) f_e(q \mid e) dq = - \int_{q_{min}}^{q_{max}} V'(q - w_{FB}(q))[1 - w'_{FB}(q)] F_e(q \mid e) dq \tag{1.13}$$

Since $V' > 0$, $w'_{FB}(q) < 1$ and $F_e(q \mid e) \leq 0$ for all q, the above expression is strictly positive, and hence (1.12) implies that

$$\int_{q_{min}}^{q_{max}} V(q - w(q)) f_e(q \mid e) dq > 0,$$

which contradicts condition (1.11). Therefore, μ must be strictly positive at the second best optimum. □

For simplicity assume that the principal is risk neutral, i.e., $V' = 1$, Condition (1.6) can be written as

$$\frac{1}{u'(w(q))} = \mu + \lambda \frac{f_e(q \mid e)}{f(q \mid e)} \qquad \text{for all } \ q. \tag{1.6$'$}$$

Ideally, we would like the principal to pay the agent more for better performance, i.e., the compensation function is monotone. It is easy to show from the above that $w'(q) > 0$ if the following *monotone likelihood ratio property* holds for all q:

$$\frac{d}{dq} \left[\frac{f_e(q \mid e)}{f(q \mid e)} \right] > 0, \tag{MLRP}$$

The MLR property of the density function means that a good performance is a signal that, with high probability, high effort was exerted. Note that (MLRP) implies first order stochastic dominance, i.e., $F_e(q \mid e) < 0$.

Lemma 1.5

The monotone likelihood ratio property implies first order stochastic dominance.

Proof. Notice that (MLRP) implies for any $q_2 > q_1$ and $e_2 > e_1$,

$$\begin{aligned} & \frac{f(q_2 \mid e_2)}{f(q_2 \mid e_1)} > \frac{f(q_1 \mid e_2)}{f(q_1 \mid e_1)} \\ \iff & f(q_2 \mid e_2) f(q_1 \mid e_1) > f(q_1 \mid e_2) f(q_2 \mid e_1) \end{aligned} \tag{$\star$}$$

The above inequality implies that

(a)

$$\int_{q_{min}}^{q_2} f(q_2 \mid e_2) f(q_1 \mid e_1) dq_1 > \int_{q_{min}}^{q_2} f(q_1 \mid e_2) f(q_2 \mid e_1) dq_1$$
$$\iff f(q_2 \mid e_2) F(q_2 \mid e_1) > f(q_2 \mid e_1) F(q_2 \mid e_2).$$

Letting $q_2 = q$ we get from the above that

$$\frac{f(q \mid e_2)}{f(q \mid e_1)} > \frac{F(q \mid e_2)}{F(q \mid e_1)} \tag{1.14}$$

(b) Similarly, integrating ($\star$) with respect to q_2 over $[q_1, q_{max}]$ and letting $q_1 = q$, we get

$$\frac{f(q \mid e_2)}{f(q \mid e_1)} < \frac{1 - F(q \mid e_2)}{1 - F(q \mid e_1)} \tag{1.15}$$

Therefore, (1.14) and (1.15) together imply

$$\frac{1 - F(q \mid e_2)}{F(q \mid e_2)} > \frac{1 - F(q \mid e_1)}{F(q \mid e_1)}$$
$$\iff F(q \mid e_2) < F(q \mid e_1) \quad \text{for any} \quad e_2 > e_1. \tag{1.16}$$

Let $e_2 = e_1 + h$ for $h > 0$. The the above inequality implies that

$$\frac{F(q \mid e_1 + h) - F(q \mid e_1)}{h} < 0.$$

Since $F(q \mid \cdot)$ is continuous, taking the limit of the above as $h \to 0$ we get $F_e(q \mid e) < 0$ which is equivalent to first order stochastic dominance of $F(\cdot \mid e)$. $\square$

FOSD simply means that a higher effort level shifts probability mass to better performance. Notice that the condition (MLRP) is a stronger condition that the FOSD. Following example shows that $F_e(q \mid e) < 0$, but the corresponding density function does not satisfy (MLRP), and hence the optimal wage schedule is non-monotone.

Example 1.4: Non-monotone compensation function

Suppose there are only three possible performance realizations: $q_L < q_M < q_H$, and that the agent has two possible effort levels: $e' < e''$. The conditional densities are given in the following table:

	$f(q_L \mid e)$	$f(q_M \mid e)$	$f(q_H \mid e)$
e'	0.5	0.5	0.0
e''	0.4	0.1	0.5

Here the second best compensation function is such that $w(q_H) > w(q_L) > w(q_M)$. The point is that when $q = q_L$ it is almost likely that the agent chose e' as e''. Therefore, the principal does not want to punish the agent too much for low performance realizations. Notice in this example that the corresponding distribution function satisfies FOSD, but this property is not enough to guarantee a monotone compensation function.

In order to solve the principal's utility maximization problem we have ignored the constraint (IC$_b$). This can be done only if (IC$_a$) is necessary and sufficient for the agent's maximization problem (IC$_\theta$). We know that this would be the case if the agent's objective function is strictly concave in e which is not true in general under any density function $f(\cdot \mid e)$. We therefore require to give additional restriction on the distribution of firm output. We omit such technicalities which can be found in (Bolton and Dewatripont, 2005, Chapter 4).

1.5.3 Informative signals

Now suppose that $s \in [s_{min}, s_{max}]$ be a signal, which in addition to q, is observed by both parties and hence can potentially be used in constructing the compensation scheme. Under what conditions the principal must condition the incentive scheme both on q and s? If the contract includes s in addition to q, i.e., $w(q, s)$ is an optimal contract, and if both the principal and the agent are strictly better off by using such rules instead of $w(q)$, then the signal s is said to be *valuable*. First consider the following definition.

Definition 1.1: Informative signals

Let $f(q, s \mid e)$ be the joint density of q and s. Performance q is a "sufficient statistic" for $\{q, s\}$ with respect to e, or s is "non-informative" about e given q if and only if f is multiplicatively separable in s and e, i.e.,

$$f(q, s \mid e) = g(q, s)h(q \mid e).$$

The signal s is "informative" about e whenever q is not a sufficient statistic for $\{q, s\}$ with respect to e.

In the context of moral hazard, Hölmstrom (1979) proves the following important result, which can be extended to more general settings.

Proposition 1.1

Let $w(q)$ be an optimal compensation scheme for which the agent's choice of effort is unique and interior in $[0, 1]$. Then there exists a compensation scheme $w(q, s)$ which strictly Pareto dominates $w(q)$ if and only if s is informative about e, i.e., a signal is valuable if and only if it is informative.

Proof. We only provide a sketch of the proof. For details see Hölmstrom (1979). The first order condition for the determination of $w(q, s)$ is given by:

$$\frac{V'(q - w(q, s))}{u'(w(q, s))} = \lambda + \mu \frac{f_e(q, s \mid e)}{f(q, s \mid e)} \qquad \text{for all } (q, s). \tag{1.17}$$

Notice that the likelihood ratio f_e/f is independent of s if and only if s is non-informative about e, and hence the optimal compensation scheme should be independent of the signal s. $\square$

The above result asserts that when q is a sufficient statistic for $\{q, s\}$ with respect to e, i.e., q contains all relevant information about the agent's effort, then there is no additional gain for the principal if she uses the signal s. Conversely, if there is an additional signal that does contain additional statistical information about the agent's effort, then such informative signal must be included in the optimal compensation contract.

1.6 Optimality of linear contracts

The simplest form of contracts under moral hazard are linear compensation schemes which are prevalent in many principal-agent relationships. For example, a tenant generally pays a fixed rent to his landlord and gets a fixed share of the crop. Franchise contracts are often linear contracts. In what follows we show that under mild restrictions on the utility function of the agent and normality of the distribution of the performance, the optimal contracts are linear. Let the performance q is given by:

$$q = e + \varepsilon, \tag{1.18}$$

where ε is normally distributed with zero mean and variance σ^2. The agent has constant absolute risk averse (CARA) risk preferences represented by the following utility function:

$$u(w,\, e) = -\exp\{-\eta[w - \psi(e)]\} \tag{1.19}$$

where w is the amount of monetary compensation and $\eta > 0$ is the agent's coefficient of absolute risk aversion. For simplicity assume that $\psi(e) = (1/2)ce^2$ with $c > 0$. Consider the following linear contract.

$$w = f + bq \tag{1.20}$$

where f is the fixed salary and b is the bonus or the "piece rate" related to performance. Notice that the expectation and variance of the agent's income are given by:

$$\begin{aligned} E[w] &= f + be, \\ Var(w) &= b^2\sigma^2. \end{aligned}$$

A risk neutral principal thus solves the following maximization problem:

$$\max_{e,\, f,\, b}\ E[q - w] \tag{M$_3$}$$

$$\text{subject to} \quad E[-\exp\{-\eta[w - \psi(e)]\}] \geq u(\bar{w}), \tag{IR}$$

$$e \in \text{argmax}_{e'}\ E[-\exp\{-\eta[w - \psi(e')]\}]. \tag{IC}$$

The first constraint (IRA) is the *individual rationality* or *participation* constraint which guarantees a minimum expected utility, called the *reservation utility*, equal to $u(\bar{w})$ where $\bar{w}$ denotes the minimum acceptable certain monetary equivalent of the agent's compensation contract. The second is the *incentive compatibility* constraint which asserts that the agent chooses the effort level that maximizes his expected utility. Notice that

$$\begin{aligned} &E[-\exp\{-\eta[w - \psi(e)]\}] \\ =&E[-\exp\{-\eta[f + be + b\varepsilon - (1/2)ce^2]\}] \\ =&-\ \exp\{-\eta[f + be - (1/2)ce^2]\}\, E[\exp\{-\eta b\varepsilon\}]. \end{aligned} \tag{1.21}$$

Since $\varepsilon \sim N(0,\, \sigma^2)$, its moment generating function is given by $M_\varepsilon(t) \equiv E[\exp\{t\varepsilon\}] = \exp\{(1/2)\sigma^2 t^2\}$. Therefore, Equation (1.21) can be written as

$$E[-\exp\{-\eta[w - \psi(e)]\}] = -\exp\{-\eta\hat{w}(e)\}, \tag{1.22}$$

where

$$\hat{w}(e) = f + be - \frac{\eta}{2}b^2\sigma^2 - \frac{1}{2}ce^2 = E[w] - \frac{\eta}{2}Var(w) - \psi(e)$$

is the *certainty equivalent compensation* of the agent, which is equal to his expected compensation net of his effort cost and of a risk premium. Because of the exponential form of the utility function the agent's maximization problem reduces to:

$$e = \text{argmax}_{\hat{e}} \left\{ f + b\hat{e} - \frac{\eta}{2}b^2\sigma^2 - \frac{1}{2}c\hat{e}^2 \right\} = \frac{b}{c}. \tag{IC$'$}$$

Substituting the above into the principal's objective function and the individual rationality constraint of the agent, the maximization problem of the principal reduces to:

$$\max_{f,b}\ \frac{b}{c} - \left[f + \frac{b^2}{c} \right] \tag{M$_3'$}$$

$$\text{subject to} \quad f + \frac{b^2}{2c} - \frac{\eta}{2}\, b^2\sigma^2 = \bar{w}. \tag{IRA$'$}$$

Solving the above, and using $e = b/c$ we get

$$b^* = \frac{1}{1 + \eta c\sigma^2}\,.$$

Thus, optimal piece rate decrease when c (marginal cost of effort), η (risk aversion) and σ^2 (randomness of performance) go up. The effect of an increase in σ^2 on b^* implies a trade-off between risk sharing and incentive.

1.7 Contracts under risk neutrality and limited liability

1.7.1 Managerial compensation contracts

Consider a contracting problem between a firm (principal) and a manager (agent). The constant marginal cost of production θ of the firm may take two values: 'high' and 'low', i.e., $\theta \in \{\theta_L, \theta_H\}$ with $\theta_H > \theta_L \geq 0$. Initially, the firm is 'inefficient', i.e., it has a marginal cost equal to θ_H. The principal hires a manager whose principal task is to exert R&D effort e in order to reduce the marginal cost to the lower level. Let Prob.$[\theta = \theta_L \mid e] = p(e)$. Assume that $p(0) = 0$ and $p'(e) > 0 \geq p''(e)$ for all $e > 0$. The incentive scheme is given by (f, b) where f is the fixed salary and b is the bonus for cost reduction from θ_H to θ_L. Both the principal and the agent are risk neutral. The principal's maximization problem is given by:

$$
\begin{aligned}
&\max_{e,\, f,\, b}\ p(e)\pi(\theta_L) + [1 - p(e)]\pi(\theta_H) - [f + p(e)b] && (\mathrm{M}_4)\\
\text{subject to}\quad & [f + p(e)b] - \psi(e) \geq \bar{u}, && (\mathrm{IR})\\
& e = \operatorname{argmax}_{\hat{e}}\ [f + p(\hat{e})b] - \psi(\hat{e}), && (\mathrm{IC})\\
& f + b \geq l, \qquad \text{and} \qquad f \geq l, && (\mathrm{LL})
\end{aligned}
$$

where $l \geq 0$ is the agent's liability limit. Constraints (LL) is the agent's *limited liability* constraints that guarantee a minimum final income l at each state of the nature. Using the first order condition of the agent's maximization problem, (IC) can be written as

$$
b = \frac{\psi'(e)}{p'(e)}. \tag{IC$'$}
$$

Notice that $f \geq l$, and $p'(e) > 0$ and $\psi'(e) > 0$ imply $f + b \geq l$, and hence the first limited liability constraint can be ignored. Also define by $\pi := \pi(\theta_H) - \pi(\theta_L)$ the marginal benefit of cost reduction for the firm. Thus, substituting for b in the principal's objective function and in the constraints, the maximization problem reduces to:

$$
\begin{aligned}
&\max_{e,\, f}\ p(e)\pi + \pi(\theta_H) - f - \frac{p(e)\psi'(e)}{p'(e)} && (\mathrm{M}_4')\\
\text{subject to}\quad & f + \frac{p(e)\psi'(e)}{p'(e)} - \psi(e) \geq \bar{u} && (\mathrm{IR}')\\
& f \geq l. && (\mathrm{LL}')
\end{aligned}
$$

The Lagrange function is given by:

$$
\mathcal{L} = p(e)\pi + \pi(\theta_H) - f - B(e) + \lambda\,[f + B(e) - \psi(e) - \bar{u}] + \mu[f - l],
$$

where λ and μ are the associated Lagrange multipliers, and

$$
B(e) := \frac{p(e)\psi'(e)}{p'(e)}\,.
$$

Notice that

$$
B'(e) = \frac{p(e)}{p'(e)}\left[\psi''(e) - \frac{p''(e)\psi'(e)}{p'(e)}\right] + \psi'(e).
$$

The KKT conditions are given by:

$$
p'(e)\pi - B'(e) + \lambda[B'(e) - \psi'(e)] = 0, \tag{1.23}
$$

$$
\lambda + \mu = 1, \tag{1.24}
$$

$$
\lambda\,[f + B(e) - \psi(e) - \bar{u}] = 0, \tag{1.25}
$$

$$
\mu[f - l] = 0, \tag{1.26}
$$

$$
\lambda,\, \mu \geq 0. \tag{1.27}
$$

First notice that both individual rationality and limited liability constraints cannot be non-binding simultaneously, otherwise it would contradict the KKT condition (1.24) [i.e., 0=1]. Hence, we consider the following three cases.

CASE 1: $\lambda = 0$ and $\mu = 1$, i.e., individual rationality does not bind but limited liability does. In this case $f = l$. From the first order condition (1.23) we get

$$p'(e)\pi = B'(e). \tag{1.28}$$

Call the level of effort that solves the above equation $\underline{e}$. Since the individual rationality constraint does not bind, we have

$$\bar{u} < l + B(\underline{e}) - \psi(\underline{e}).$$

The above inequality implies that for low values of $\bar{u}$ the solutions $f = l$ and $e = \underline{e}$ are optimal. The optimal bonus is given by $\underline{b} = \psi'(\underline{e})/p'(\underline{e})$.

CASE 2: $\lambda, \mu \in (0, 1)$, i.e., both the constraints bind at the optimum. In this case also $f = l$. Call the solution $\hat{e}(\bar{u})$ which is implicitly defined by

$$l + B(\hat{e}(\bar{u})) - \psi(\hat{e}(\bar{u})) = \bar{u}.$$

It is easy to show that the solutions $f = l$ and $e = \hat{e}(\bar{u})$ are optimal only if

$$l + B(\underline{e}) - \psi(\underline{e}) \leq \bar{u} \leq l + B(e^*) - \psi(e^*),$$

where e^* is given by the equation $p'(e^*)\pi = \psi'(e^*)$.

CASE 3: $\lambda = 1$ and $\mu = 0$, i.e., individual rationality binds but limited liability does not. From the first order condition (1.23) we get

$$p'(e)\pi = \psi'(e), \tag{1.29}$$

which gives $e = e^*$. From the incentive compatibility constraint we have $b^* = \pi$. The optimal fixed salary is determined from the individual rationality constraint which is given by:

$$f^* = \bar{u} - [B(e^*) - \psi(e^*)].$$

The non-binding limited liability constraint implies that

$$\bar{u} \geq l + B(e^*) - \psi(e^*),$$

which is the necessary condition for the solutions $e = e^*$, $b = \pi$ and $f = f^*$ to be optimal.

Exercise 1.1

Solve for the optimal contracts (e, f, b) when the firm faces a linear demand $P(q) = a - q$, and $p(e) = e$ and $\psi(e) = 1/2ce^2$. Define by $\phi(\bar{u})$ the value function of the principal's maximization problem. Show that $\phi(\bar{u})$ is weakly concave in $\bar{u}$.

1.7.2 Optimal debt contracts

Debt is a prevalent contracting schemes in financial contracting between entrepreneurs and outside investors. Innes (1990) shows that a 'risky' debt contract emerges as an optimal contract under two-sided limited liability when the repayment to investors is constrained to non-decreasing in the firm's performance. Consider a penny-less risk neutral entrepreneur of a startup firm who requires to raise I, which is the setup cost, from external sources. The profit of the firm $q \in [0, \infty)$ has a distribution function $F(q \mid e)$ with the corresponding density function $f(q \mid e)$ satisfies (MLRP). Let $r(q)$ denote the performance-based repayment scheme to the investor. A risky debt contract is defined as

$$r_D(q) = \min\{q, D\}, \tag{1.30}$$

where D is the face value of debt. We impose the following two constraints on the repayment scheme $r(q)$:

1. Two-sided limited liability: $0 \le r(q) \le q$ for all $q \in [0, \infty)$;
2. monotonicity: $r'(q) \ge 0$ for all $q \in [0, \infty)$.

The two-sided limited liability constraint implies that neither the investor nor the entrepreneur will get negative incomes at any state of the nature. To rationalize the monotonicity constraint, suppose over a subset of the range of values of q that $r'(q) < 0$. Then the entrepreneur will benefit by secretly borrowing money at par from some other source, and show a higher output than the realized one.

The maximization problem of the entrepreneur is given by:

$$\max_{\{e, r(q)\}} \int_0^\infty [q - r(q)] f(q \mid e) dq - \psi(e) \tag{M$_5$}$$

$$\text{subject to} \quad \int_0^\infty r(q) f(q \mid e) dq - I = 0, \tag{IRP}$$

$$\int_0^\infty [q - r(q)] f_e(q \mid e) dq = \psi'(e), \tag{IC}$$

$$0 \le r(q) \le q, \tag{LL}$$

$$r'(q) \ge 0. \tag{Mon}$$

Let λ and μ be the multipliers associated with (IRP) and (IC'_θ), respectively. Then the Lagrange expression is linear in $r(q)$ with the corresponding coefficient

$$\lambda - 1 + \mu \frac{f_e(q \mid e)}{f(q \mid e)}.$$

Therefore, without invoking the monotonicity constraint, the optimal solution to $r(q)$ is given by:

$$r^*(q) = \begin{cases} q & \text{if } \frac{\lambda - 1}{\mu} > \frac{f_e(q|e)}{f(q|e)}, \\ 0 & \text{if } \frac{\lambda - 1}{\mu} \le \frac{f_e(q|e)}{f(q|e)}. \end{cases}$$

Now (MLRP) implies that there exists an output level $\hat{q}$ such that

$$r^*(q) = \begin{cases} q & \text{if } q < \hat{q}, \\ 0 & \text{if } q \ge \hat{q}. \end{cases}$$

The above contract is certainly not monotone, and $r^*(q) \ne r_D(q)$. However, invoking (Mon), it is easy to see that the constrained optimal contract takes the form of a risky debt contract $r_D(q)$ where D is the lowest value that solves the (IRP) constraint:

$$\int_0^D q f(q \mid e) dq + [1 - F(D \mid e)] D = I.$$

Notice that the above proves that $r_D(q)$ is an optimal contract that solves the maximization problem (M$_5$) subject to the constraints (IRP)-(Mon). There might be other monotone non-debt contract that solves the entrepreneur's maximization problem. Innes (1990) shows that if there is an optimal debt contract that at which the investor obtains the same expected income compared with an optimal non-debt contract $r_{ND}(q)$, then the effort level induced by the debt contract will be strictly higher than that induced by the equivalent non-debt contract, and the entrepreneur will choose the debt contract. The proof is lengthy and technical, and hence is left as an exercise. Let us give the intuition behind the result.

Let $w_{ND}(q) = q - r_{ND}(q)$ be the income of the entrepreneur at a monotone non-debt contract. Notice that $r'_{ND}(q) > 0$ and $0 \leq r_{ND}(q) \leq q$ together imply that $r'_{ND}(q) < 1$, and hence $0 \leq w'_{ND}(q) < 1$. Let $w_D(q)$ be the income schedule of the entrepreneur at the debt contract $r_D(q)$. Clearly,

$$w_D(q) = \begin{cases} 0 & \text{if } \; q \leq D, \\ q - D & \text{if } \; q > 0. \end{cases}$$

Since $w'_D(q) = 1$ for all $q > D$, there exists a unique output level $q^* > D$ such that $w_{ND}(q) > (<)\, w_D(q)$ for $q < (>)\, q^*$. Since (MLRP) implies first order stochastic dominance, a marginal increase in effort shifts probability mass to the hight performance states. Therefore, the entrepreneur's compensation is also shifted to higher outcomes, and he has stronger incentives to exert effort. In particular, beyond q^*, the debt contract gives full marginal incentives to the entrepreneur, i.e., $w'_D(q) = 1$, while the non-debt contract gives less than full marginal incentives $[w'_{ND}(q) < 1]$. Thus for the same expected income for the investor, the debt contract induces the entrepreneur to spend more effort than any monotone non-debt contract.

1.8 Grossman and Hart's approach to the principal-agent problem

Grossman and Hart (1983) take an alternative approach to solve the standard principal-agent problem under moral hazard. Suppose there are only N states: $q_i \in \{q_1, \ldots, q_N\}$ with $0 \leq q_1 < \ldots < q_N$. Let $p_i(e)$ denote the probability of outcome q_i given effort choice e. Assume that the principal is risk neutral with utility function $V(q - w) = q - w$, and the agent's utility function is given by:

$$U(w, e) = \phi(e)u(w(q)) - \psi(e).$$

The above general representation contains as special cases the multiplicatively separable utility function $[\psi(e) = 0$ for all $e]$ and the additively separable utility function $[\phi(e) = 1$ for all $e]$. Moreover, assume that $u(w(q))$ is continuous, strictly increasing, and concave on $(w_{min}, +\infty)$, and that $\lim_{w \to w_{min}} u(w) = -\infty$. The principal's objective is to solve

$$\max_{e, \{w_i\}} \sum_{i=1}^{N} p_i(e)(q_i - w_i) \tag{M$_6$}$$

$$\text{subject to} \quad \sum_{i=1}^{N} p_i(e)\{\phi(e)u(w_i) - \psi(e)\} \geq \bar{u}, \tag{IR}$$

$$\sum_{i=1}^{N} p_i(e)\{\phi(e)u(w_i) - \psi(e)\} \geq \sum_{i=1}^{N} p_i(\hat{e})\{\phi(\hat{e})u(w_i) - \psi(\hat{e})\} \quad \text{for all } \; \hat{e} \in A, \tag{IC}$$

where $w_i := w(q_i)$. We concentrate on the second best contract which is determined in the following two stages.

1.8.1 Implementation

This stage involves solving the following problem:

$$\min_{\{w_i\}} \sum_{i=1}^{N} p_i(e) w_i \tag{Min}$$

$$\text{subject to} \quad \sum_{i=1}^{N} p_i(e)\{\phi(e)u(w_i) - \psi(e)\} \geq \bar{u}, \tag{IR$'$}$$

$$\sum_{i=1}^{N} p_i(e)\{\phi(e)u(w_i) - \psi(e)\} \geq \sum_{i=1}^{N} p_i(\hat{e})\{\phi(\hat{e})u(w_i) - \psi(\hat{e})\} \quad \text{for all } \hat{e} \in A. \tag{IC$'$}$$

This program solves for any effort $e \in A$. Now define $h := u^{-1}$ and $u_i := u(w_i)$ for $i = 1, \ldots, N$. Let

$$\boldsymbol{U} = \{u \mid u(w) = u \ \text{ for some } \ w \in (w_{min}, +\infty)\},$$

assume that

$$\frac{\bar{u} + \psi(e)}{\phi(e)} \in \boldsymbol{U} \quad \text{for all } \ e \in A,$$

i.e., for every effort e, there exists a compensation level that meets the individual rationality constraint of the agent for that effort. Then the transformed program is

$$\min_{\{u_i\}} \sum_{i=1}^{N} p_i(e) h(u_i) \tag{Min$'$}$$

$$\text{subject to} \quad \sum_{i=1}^{N} p_i(e)\{\phi(e)u_i - \psi(e)\} \geq \bar{u}, \tag{IR$''$}$$

$$\sum_{i=1}^{N} p_i(e)\{\phi(e)u_i - \psi(e)\} \geq \sum_{i=1}^{N} p_i(\hat{e})\{\phi(\hat{e})u_i - \psi(\hat{e})\} \quad \text{for all } \hat{e} \in A. \tag{IC$''$}$$

We now have linear constraints and a convex objective function, and hence the KKT conditions are necessary and sufficient. Let $\boldsymbol{u} = (u_1, \ldots, u_N)$ and define

$$C(e) = \inf \left\{ \sum_{i=1}^{N} p_i(e) h(u_i) \mid \boldsymbol{u} \ \text{ implements } \ e \right\}.$$

If there is no $\boldsymbol{u}$ that implements some $e \in A$, for such effort let $C(e) = +\infty$. Grossman and Hart (1983) show that when $p_i(e) > 0$ for $i = 1, \ldots, N$, there exists a solution $\boldsymbol{u}^* = (u_1^*, \ldots, u_N^*)$ to the problem (Min$'$), so that the cost function $C(e)$ is well defined. If one assumes that all relevant $h(u_i)$ are bounded and that the constraint set is compact, then $\boldsymbol{u}^*$ exists by Weierstrass' theorem.

1.8.2 Optimization

The second stage is to choose $e \in A$ to solve

$$\max_{e} \sum_{i=1}^{N} p_i(e) q_i - C(e).$$

To summarize, first find the minimum cost for the principal to implement a given effort level e. Once the cost function is determined, find the effort level that maximizes the principal's expected net profit.

1.9 Optimal contracts with multiple tasks

Hölmstrom and Milgrom (1991) analyze the optimal incentive schemes when the agent may undertake more than one tasks. The agent undertakes two tasks $i = 1, 2$ whose efforts are e_1 and e_2. The output of task i is given by $q_i = e_i + \varepsilon_i$ where ε_i is a task-specific noise which is normally distributed with mean 0 and variance σ_i^2, and is uncorrelated with ε_j where $j \neq i$. The effort-cost function is given by:

$$\psi(e_1, e_2) = \frac{1}{2}\left(c_1 e_1^2 + c_2 e_2^2\right) + \delta e_1 e_2 \quad \text{with} \quad 0 \leq \delta \leq \sqrt{c_1 c_2}.$$

When $\delta = 0$, the two efforts are technologically independent, and they are perfect substitutes if $\delta = \sqrt{c_1 c_2}$. Whenever $\delta > 0$, raising effort on one task raises the marginal cost of effort on the other task, the so-called *effort substitution* problem. The agent's risk preferences are given by a CARA utility function of the following form:

$$u(w, e_1, e_2) = -\exp\{-\eta[w - \psi(e_1, e_2)]\},$$

where η is the coefficient of absolute risk aversion, and w is a linear compensation scheme which is given by:

$$w = f + b_1 q_1 + b_2 q_2.$$

The agent's certainty equivalent compensation is given by:[2]

$$\hat{w}(e_1, e_2) = f + b_1 e_1 + b_2 e_2 - \frac{\eta}{2}\left[b_1 \sigma_1^2 + b_2 \sigma_2^2\right] - \frac{1}{2}\left(c_1 e_1^2 + c_2 e_2^2\right) - \delta e_1 e_2. \tag{1.31}$$

The incentive compatibility constraints are given by:

$$b_i = c_i e_i + \delta e_j \quad \text{for} \quad i, j = 1, 2, \quad \text{and} \quad i \neq j. \tag{IC$_i$}$$

Solving the above equations we get

$$e_i = \frac{b_i c_j - \delta b_j}{c_i c_j - \delta^2} \quad \text{for} \quad i, j = 1, 2, \quad \text{and} \quad i \neq j.$$

The principal's maximization problem is given by:

$$\max_{e_1, e_2, f, b_1, b_2} \quad e_1(1 - b_1) + e_2(1 - b_2) - f \tag{M$_7$}$$

$$\text{subject to} \quad f + b_1 e_1 + b_2 e_2 - \frac{\eta}{2}\left[b_1 \sigma_1^2 + b_2 \sigma_2^2\right] - \frac{1}{2}\left(c_1 e_1^2 + c_2 e_2^2\right) - \delta e_1 e_2 \geq \bar{w}, \tag{IR}$$

$$e_1 = \frac{b_1 c_2 - \delta b_2}{c_1 c_2 - \delta^2}, \tag{IC$_1$}$$

$$e_2 = \frac{b_2 c_1 - \delta b_1}{c_1 c_2 - \delta^2}. \tag{IC$_2$}$$

Substitute for e_1 and e_2 from (IC$_1$) and (IC$_2$) into the objective function and the constraint (IR$_\theta$). So the maximization program is expressed only in terms of f, b_1 and b_2. Since the individual rationality constraint will bind at the optimum, we get $f = f(b_1, b_2)$ which we substitute in the objective function to get an unconstrained maximization problem in terms of b_1 and b_2 alone. Solving the unconstrained maximization problem we get the optimal piece rates which are given by:

$$b_1^* = \frac{1 + (c_2 - \delta)\eta\sigma_2^2}{1 + \eta c_1 \sigma_1^2 + \eta c_2 \sigma_2^2 + \eta^2 \sigma_1^2 \sigma_2^2 (c_1 c_2 - \delta^2)},$$

$$b_2^* = \frac{1 + (c_1 - \delta)\eta\sigma_1^2}{1 + \eta c_1 \sigma_1^2 + \eta c_2 \sigma_2^2 + \eta^2 \sigma_1^2 \sigma_2^2 (c_1 c_2 - \delta^2)}.$$

[2] For derivation see Section 1.6.

- When the two tasks are technologically independent, i.e., $\delta = 0$, the optimal piece rates are given by:

$$b_i^* = \frac{1}{1 + \eta c_i \sigma_i^2} \quad \text{for} \quad i = 1,\, 2,$$

  the results we have obtained in Section 1.6.

- It is easy to check that both $\partial b_i^* / \partial \sigma_i^2$ are $\partial b_i^* / \partial \sigma_j^2$ are negative. The first one is the simple trade-off between risk and incentive. The second one is the complementarity between the piece rates in the presence of effort substitution problem.

Chapter 2

Moral hazard with multiple agents

This chapter analyzes optimal incentive schemes under moral hazard when a principal interacts with many agents. In a partnership/firm the output $Q = (q_1, \ldots, q_n)$ is jointly affected the efforts $a = (a_1, \ldots, a_n)$ of the n agents, where q_i is the individual output of agent i. Let $F(Q|a)$ be the joint conditional distribution of outputs. Throughout we will assume that the principal is risk neutral and agent i has separable preferences over salary w_i and effort a_i as follows

$$U_i(w, a) = u_i(w) - \psi_i(a), \quad \text{with} \quad u_i''(.) \leq 0.$$

We will analyze the optimal incentive schemes under two different situations: when Q, the aggregate output is observed publicly, and when the individual outputs, q_i's are observed.

2.1 Moral Hazard in Teams

This section is based on Hölmstrom (1982). In a firm only the single aggregate and deterministic output is observed publicly, which is given by the production function $Q = Q(a)$. We assume that $Q_i(a) > 0$, $Q_{ii}(a) < 0$, $Q_{ij} \geq 0$ and the Hessian matrix $\nabla Q(a)$ is negative semi-definite. Since the output is deterministic, without loss of generality we can assume that the agents are risk neutral. A partnership is defined by the following scheme of sharing rules

$$w(Q) = [w_1(Q), \ldots, w_n(Q)].$$

To start with we impose the following budget balance condition

$$\sum_{i=1}^{n} w_i(Q) = Q, \quad \text{for all} \quad Q. \tag{BB}$$

Here arises the free riding problem: someone else works hard, I gain. Let the first best efforts are given by a^* which are given by

$$a^* = \underset{a \in [0,\infty)^n}{\text{argmax}} \left\{ Q(a) - \sum_{i=1}^{n} \psi_i(a_i) \right\}. \tag{2.1}$$

The first order conditions are given by

$$Q_i(a^*) = \psi_i'(a_i^*), \quad \text{for all} \quad i = 1, \ldots, n. \tag{FB}$$

Further, let a^N be the efforts chosen by the agents in a Nash equilibrium, which are given by

$$a_i^N = \underset{a_i \in [0,\infty)}{\text{argmax}} \ \{w_i(Q(a)) - \psi_i(a_i)\} \quad \text{for all} \quad i = 1, \ldots, n. \tag{2.2}$$

The NE efforts are characterized by

$$w_i'(Q(a^N))\,Q_i(a^N) = \psi_i'(a_i^N), \quad \text{for all} \quad i = 1,\,\ldots,\,n. \tag{NE}$$

Now the question is whether $a^* = a^N$. Clearly, (FB) and (NE) are not compatible unless $w_i'(Q(a)) = 1$. This condition boils down to

$$w_i'(Q) = Q(a) + C_i, \quad \text{for all} \quad i = 1,\,\ldots,\,n. \tag{2.3}$$

For the above to be compatible with (BB) one thus needs to introduce a third party, called the budget breaker, who can write binding contracts with the n agents to receive transfers $t_i = -C_i$ from agent i. Hölmstrom (1982) claims that such a scheme exists and implements a^* in a Nash equilibrium. In order to show this we relax (BB) to a *no deficit* condition

$$\sum_{i=1}^{n} w_i(Q) \le Q, \quad \text{for all} \quad Q. \tag{ND}$$

From (2.3) and (ND) we have

$$\sum_{i=1}^{n} w_i(Q(a^*)) = nQ(a^*) - \sum_{i=1}^{n} t_i \le Q(a^*). \tag{2.4}$$

Also

$$t_i = Q(a^*) - w_i(Q(a^*)) \le Q(a^*) - \psi_i(a_i^*). \tag{2.5}$$

Thus for such t_i's to exist we need

$$\sum_{i=1}^{n} t_i \le nQ(a^*) - \sum_{i=1}^{n} \psi_i(a_i^*) \le \sum_{i=1}^{n} t_i + Q(a^*) - \sum_{i=1}^{n} \psi_i(a_i^*). \tag{2.6}$$

Notice that (4) and (5) imply that such a scheme is profitable for both the budget breaker and the agents. From the last equation, such transfers $t = (t_1,\,\ldots,\,t_n)$ exist since $Q(a^*) - \sum_{i=1}^{n} \psi_i(a_i^*)$ is strictly positive, otherwise a^* would not be efficient. Notice that if the firm performs better then it hurts the budget breaker, since $w_{BB}(Q) = \sum_i t_i + Q - nQ$ implying that $w_{BB}'(Q) = -(n-1) < 0$ at $Q(a^*)$.

Are there other ways to support first best? Mirrlees contract does. Consider the following bonus scheme

$$w_i(Q) = \begin{cases} b_i & \text{if} \;\; Q = Q(a^*), \\ -k & \text{if} \;\; Q < Q(a^*). \end{cases}$$

Choose b_i's such that $b_i - \psi_i(a_i^*) > -k$ and $\sum_{i=1}^{n} b_i = Q(a^*)$. This is possible because of the fact that $Q(a^*) - \sum_{i=1}^{n} \psi_i(a_i^*) > 0$, otherwise a^* would not be efficient. Thus a^* is a Nash equilibrium. Mirrlees' contract can be interpreted as debt financing by the firm. Firm commits to repay debts of $D = Q(a^*) - \sum_{i=1}^{n} b_i$, and b_i to each i. If it cannot, creditors collect Q and each agent pays a penalty k.

Example 2.1: Implementing the first best

Consider $n = 3$ with $a_i \in \{0,\,1\}$ and $\psi_i(1) > \psi_i(0)$ for all $i = 1,\,2,\,3$. Let $Q^1 = Q(0,\,1,\,1)$, $Q^2 = Q(1,\,0,\,1)$ and $Q^3 = Q(1,\,1,\,0)$. Thus $a^* = (1,\,1,\,1)$. If $Q^1 \neq Q^2 \neq Q^3$, then the single shirker is identified and punished, and the other two are rewarded. Thus, first best efforts are implemented. If $Q^1 = Q^2 \neq Q^3$, then also it is possible to identify and reward the non-shirker, and hence first best can be implemented.

Example 2.2: Approximate efficiency

Let $n = 2$, $A_1 = A_2 = [0, \infty)$, $Q = a_1 + a_2$ and $\psi_i(a_i) = a_i^2/2$. Then $a^* = (1, 1)$. Consider the following incentive scheme. For $Q \geq 1$, $w_1(Q) = (Q-1)^2/2$ and $w_2(Q) = Q - w_1(Q)$, and for $Q < 1$, $w_1(Q) = Q$ and $w_2(Q) = -k$. For sufficiently high k this scheme supports an NE where agent 2 chooses $a_2 = 1$ and agent 1 chooses $a_1 = 1$ with probability $1 - \varepsilon$ and $a_1 = 0$ with probability ε. To see that this is indeed an NE note the following. Given the action choice $a_2 = 1$, agent 1's best response is found by solving

$$\max_{a_1} \; w_1(a_1+1) - \frac{a_1^2}{2} \iff \max_{a_1} \; \underbrace{\frac{a_1^2}{2} - \frac{a_1^2}{2}}_{0}.$$

Hence, agent 1 is indifferent among any actions. Thus the above randomization is his best response. As for agent 2, $a_2 = 1$ guarantees $Q \geq 1$. His payoff is given by

$$(1-\varepsilon)\left(2 - \frac{1}{2}\right) + \varepsilon(1-0) - \frac{1}{2} = 1 - \frac{\varepsilon}{2}.$$

If agent 2 chooses $a_2 \in [0, 1)$, then $Q < 1$ with probability ε. Thus his payoff is

$$(1-\varepsilon)\left(1 + a_2 - \frac{a_2^2}{2}\right) - \varepsilon\, k - \frac{a_2^2}{2} \leq 1 + a_2 - a_2^2 - \varepsilon\, k,$$

which is maximized at $a_2 = \frac{1}{2}$. Thus at $a_2 = \frac{1}{2}$, agent 2's payoff is $(5/4) - \varepsilon\, k$. Therefore, $a_2 = 1$ is optimal if

$$k \geq \frac{1}{2} + \frac{1}{4\varepsilon}.$$

This proves the assertion that the first best can be implemented approximately.

The above two examples are based on Legros and Matthews (1993).

2.2 Relative Performance Evaluation

Consider the following model as in Hölmstrom (1982).

$$q_1 = a_1 + \varepsilon_1 + \alpha\varepsilon_2,$$
$$q_2 = a_2 + \varepsilon_2 + \alpha\varepsilon_1,$$

where ε_1 and ε_2 are iid Normal random variables with mean zero and variance σ^2. The agents have CARA utility functions.

$$u(w_i, a_i) = -e^{-\eta[w_i - \psi(a_i)]}, \quad \text{where} \quad \psi(a_i) = \frac{ca_i^2}{2} \quad \text{for} \quad i = 1, 2.$$

Linear incentive schemes:

$$w_1 = z_1 + v_1 q_1 + u_1 q_2,$$
$$w_2 = z_2 + v_2 q_2 + u_2 q_1,$$

The absence of relative performance evaluation implies $u_1 = u_2 = 0$. Given the symmetry of the principal's problem we need to solve only for an individual optimal scheme $(a_i,\, w_i)$. This the principal will solve

$$\max_{\{a_i,\, z_i,\, v_i,\, u_i\}} \; E(q_i - w_i) \tag{M}$$

$$\text{subject to} \;\; E\left(-e^{-\eta[w_i - \psi(a_i)]}\right) \geq u(\bar{w}), \tag{IR$_i$}$$

$$a_i = \; \text{argmax}_{\hat{a}_i} \, E\left(-e^{-\eta[w_i - \psi(\hat{a}_i)]}\right). \tag{IC$_i$}$$

The certainty equivalent wealth of agent i with respect to a, $\hat{w}_i(a)$, is defined by:

$$-e^{-\eta \hat{w}_i(a)} = E\left(-e^{-\eta[w_i - \psi(a_i)]}\right).$$

Notice that

$$Var[v_i(\varepsilon_i + \alpha\varepsilon_j) + u_i(\varepsilon_j + \alpha\varepsilon_i)] = \sigma^2\left[(v_i + \alpha u_i)^2 + (u_i + \alpha v_i)^2\right]. \tag{2.7}$$

Then agent i's incentive constraint boils down to choosing a_i to maximize the certainty equivalent wealth:

$$\max_{\hat{a}_i} \left\{ z_i + v_i\hat{a}_i + u_i a_j - \frac{c\hat{a}_i^2}{2} - \frac{\eta\sigma^2}{2}\left[(v_i + \alpha u_i)^2 + (u_i + \alpha v_i)^2\right] \right\}. \tag{IC$'_i$}$$

The above implies that

$$a_i = \frac{v_i}{c}.$$

Substituting for a_i in the principal's objective function and the agent's binding participation constraint one can reduce the principal's problem to

$$\max_{\{z_i,\, v_i,\, u_i\}} \left\{ \frac{v_i}{c} - \left(z_i + \frac{v_i^2}{c} + \frac{u_i v_j}{c}\right) \right\}.$$

$$\text{subject to} \qquad z_i + \frac{v_i^2}{2c} + \frac{u_i v_j}{c} - \frac{\eta\sigma^2}{2}\left[(v_i + \alpha u_i)^2 + (u_i + \alpha v_i)^2\right] = \bar{w}.$$

Or, substituting for z_i from the participation constraint,

$$\max_{\{v_i,\, u_i\}} \left\{ \frac{v_i}{c} - \frac{v_i^2}{2c} - \frac{\eta\sigma^2}{2}\left[(v_i + \alpha u_i)^2 + (u_i + \alpha v_i)^2\right] \right\}.$$

The principal's problem can be solved sequentially—(a) For a given v_i, determine u_i to minimize the variance in expression (2.7), and (b) The variable v_i is then set optimally to trade off risk sharing and incentives. Minimizing the variance with respect to u_i yields

$$u_i = -\left(\frac{2\alpha}{1+\alpha^2}\right) v_i.$$

The above formula implies that the optimal u_i is negative when the two agents' outputs are positively correlated, i.e., $\alpha > 0$. In other words, an agent is penalized for the better performance of the other agent. A better performance by agent j is likely to be due to a high realization of ε_j, which also positively affect the output of agent i. By setting u_i negative, the optimal incentive scheme reduces agent i's exposure to a common shock affecting both agents' output, and thus reduces the variance of agent i's compensation.

In the second step, substitute the above formula and solve for optimal v_i:

$$\max_{\{v_i\}} \left\{ \frac{v_i}{c} - \frac{v_i^2}{2c} - \frac{\eta\sigma^2\left(1-\alpha^2\right)^2 v_i^2}{2\left(1+\alpha^2\right)} \right\}.$$

The first order condition with respect to v_i yields

$$v_i = \frac{1+\alpha^2}{1+\alpha^2+\eta c\sigma^2\left(1-\alpha^2\right)^2}.$$

For $\alpha = 0$, the above formula reduces to the formula for the optimal share in the one-agent case. Notice that $v_i(\alpha)$ is convex, $v_i(-1) = v_i(1) = 1$ and the minimum is reached at $\alpha = 0$. A situation of perfect correlation is similar to first best, and hence the agent gets the full share. The reason for using relative performance evaluation is not to induce higher effort through greater competition, but to induce higher effort by lowering their risk exposure. Interested readers should refer to Mookherjee (1984) for a more general treatment of relative performance evaluation contracts under moral hazard with many agents.

2.3 Tournaments

This subsection is based on Lazear and Rosen (1981), which consider a situation two risk neutral agents 1 and 2 produce individual outputs that are independently distributed. Obviously, following the previous analysis there is no reason why tournaments would be efficient incentive schemes. However, Lazear and Rosen (1981) show that the first-best outcome can be implemented using a tournament. Let

$$q_i = a_i + \varepsilon_i \quad \text{for} \quad i = 1,\, 2, \quad \text{where} \quad \varepsilon_i \sim F(0,\, \sigma^2).$$

The first-best efforts are given by

$$\psi'(a_1^*) = \psi'(a_2^*) = 1.$$

Consider the following incentive scheme:

$$w_i = z + q_i$$

At the first-best the expected utility of agent i is given by

$$z + E(q_i) - \psi(a^*) = z + a^* - \psi(a^*) = \bar{u}.$$

Now consider a symmetric "tournament": If $q_i > q_j$, then agent i is paid $z + W$ and agent j is paid only z. Under this scheme, the expected payoff for an agent i for efforts $(a_i,\, a_j)$ is

$$z + pW - \psi(a_i),$$

where p is the probability that agent i is the winner, which is given by:

$$p = \text{Prob.}[q_i > q_j] = \text{Prob.}[a_i - a_j > \varepsilon_i - \varepsilon_j] = H(a_i - a_j),$$

where $H(\cdot)$ is the cumulative distribution of $(\varepsilon_i - \varepsilon_j)$ which has mean 0 and variance $2\sigma^2$. The best response for an agent under tournament is given by

$$W\frac{\partial p}{\partial a_i} = \psi'(a_i) \quad \Longrightarrow \quad Wh(a_i - a_j) = \psi'(a_i).$$

In a symmetric NE, effort levels by both agents are identical. Thus in order to implement a^*, the prize must be

$$W = \frac{1}{h(0)}\,,$$

The fixed wage z can be set to satisfy the following condition:

$$z + \{\text{Prob.}[0 = a_i^* - a_j^* > \varepsilon_i - \varepsilon_j]\}W - \psi(a^*) = z + \frac{H(0)}{h(0)} - \psi(a^*) = \bar{u}.$$

This is the same as the first-best with the piece rate scheme.

Part II

Hidden information

Chapter 3

Adverse Selection

3.1 Examples of adverse selection

By "adverse selection" we refer to a situation where the principal cannot observe the characteristics of the agent she is contracting with. An agent may have private information about his "type", e.g. ability, productivity, efficiency, etc. which he may use in order to extract "rent" for the principal. The principal's objective is to design a contract, often called *mechanism*, that offers adequate incentives to the agent so that he reveals such private information. Following are some typical examples of adverse selection.

Example 3.1

Suppose a seller (principal) sells wines of high and low quality. The buyer may be of two types: sophisticated who is willing to pay a high price for vintage wine, and with moderate taste who is content with low-quality wine. If the seller could observe perfectly the type of buyer she is dealing with, she could have used discriminatory pricing: (a) a wine of high quality for a high price, and (b) a wine of low quality for a low price with the objective that the sophisticated agent would choose the first contract and the other type would settle for the other contract. One main objective of this chapter is to analyze under what conditions such allocation rules are possible to implement.

Example 3.2

n a labor market, the workers differ in their abilities which are private information to them. The firms which hire them do not know to which type of workers the jobs are being offered. The firms then must screen the workers so as to lure the high-ability workers and discard the others.

Example 3.3

n a regulated market, the firms often possess more information about their technology than the regulator. Such informational advantages can be used in the firms' favor to increase their profits.

We will study particular economic situations consisting of adverse selection problems, although the methodology of solving such problems is quite general.

3.2 Timing of events

The principal-agent relationship lasts for four dates, $t = 0, 1, 2, 3$. At date 0, the agent discovers his type θ. At $t = 1$, the principal offers a contract. At $t = 2$, the agent accepts or rejects the contract. Finally, at date 3 the contract is executed.

3.3 A model of adverse selection

3.3.1 Monopolistic screening with two types

Consider an economy as in Maskin and Riley (1984) where a monopolist produces and sells a private good to a consumer. The total cost of production is cq where q is the quantity produced and $c > 0$ is the constant marginal cost of production. The consumer has utility function $U(t, q; \theta_i) = \theta_i V(q) - t$ with $V(0) = 0$, $V' > 0 > V''$, where t is the total price paid to the monopolist, and $\theta_i \in \{\theta_H, \theta_L\}$ represents his marginal valuation or "type". The consumer may have "high" or "low" valuation, i.e., $\theta_H > \theta_L$ which is his private information. The seller only knows that the consumer is of type θ_i with probability λ_i. This situation is equivalent, due to the law of large numbers, to the one in which the seller faces a continuum of consumer, a λ_H proportion of which are of high type. The profit of the monopolist is given by $\Pi = t - cq$. A contract is a type-contingent allocation rule (q_i, t_i) for $i = H, L$ where q_i is the quantity sold to, and t_i is the transfer received by the seller from type i.

Symmetric information: perfect price discrimination

Suppose first that the monopolist can observe perfectly the types of the consumer. The monopolist chooses (q_i, t_i) for $i = H, L$ to solve the following maximization problem:

$$\max_{\{q_i, t_i\}} \; t_i - cq_i, \tag{P$_1$}$$

$$\text{subject to} \quad \theta_i V(q_i) - t_i \geq 0, \quad \text{for} \quad i = H, L. \tag{IR$_i$}$$

The constraint (IR$_i$) is the individual rationality constraint of a consumer of type θ_i under the assumption that the consumer obtains zero if he does not consume anything. The monopolist will extract the entire consumer surplus in order to maximize profit, and hence will set $t_i = \theta_i V(q_i)$ for $i = H, L$. Then the maximization problem reduces to

$$\max_{q_i} \; \theta_i V(q_i) - cq_i. \tag{P$'_1$}$$

The first order conditions give $\theta_i V'(q_i^*) = c$ for $i = H, L$. Therefore,

$$V'(q_L^*) = \frac{c}{\theta_L} > \frac{c}{\theta_H} = V'(q_H^*) \implies q_L^* < q_H^*, \;\text{ since }\; V'' < 0.$$

Hence, $t_i^* = \theta_i V(q_i^*)$ for $i = H, L$. Assume that the surplus function is given by:

$$V(q) = \frac{1}{2}\left[1 - (1-q)^2\right].$$

The demand function of type θ_i will be

$$D_i(p) = 1 - \frac{p}{\theta_i}.$$

And the consumer surplus of type θ_i is given by

$$S_i(p) = \frac{(\theta_i - p)^2}{2\theta_i}.$$

The optimal contracts are given by:

$$q_i^* = 1 - \frac{c}{\theta_i} \iff p^* = c,$$
$$t_i^* = \theta_i V(q_i^*) = \frac{\theta_i^2 - c^2}{2\theta_i}.$$

Notice that

$$t_i^* = S_i(c) + cq_i^* \equiv A_i + cq_i^*.$$

Thus, optimality is achieved by a *type-specific two-part tariff* in which a consumer of type θ_i can buy any quantity by paying a per-unit price c and a type-contingent fixed fee equal to $S_i(c)$. This is perfect price discrimination. The entire consumer surplus is extracted, and the monopolist sets a price equal to marginal cost. The total surplus is maximized, and it is divided between the buyer and the seller according to the participation constraint.

Asymmetric information: non-linear pricing

Now suppose that the monopolist cannot observe the consumer's type, but only knows its probability distribution λ_H and λ_L. Notice that

$$U(t_L^*, q_L^*; \theta_H) = \theta_H V(q_L^*) - t_L^* = \theta_H V(q_L^*) - \theta_L V(q_L^*) + \theta_L V(q_L^*) - t_L^* = (\theta_H - \theta_L)V(q_L^*) > 0,$$
$$U(t_H^*, q_H^*; \theta_L) = \theta_L V(q_H^*) - t_H^* = \theta_L V(q_H^*) - \theta_H V(q_H^*) + \theta_H V(q_H^*) - t_H^* = -(\theta_H - \theta_L)V(q_H^*) < 0.$$

From the above it is clear that, if the symmetric information contracts $((q_L^*, t_L^*), (q_H^*, t_H^*))$ are offered, then the high-type consumer will choose the contract for the low-type. The low-type consumer although will continue choosing the contract (q_L^*, t_L^*) since by choosing the other contract, he looses. Hence, the first-best contracts will not be optimal under asymmetric information.

When the types of the consumers are private information the monopolists will have to find contracts (q_L, t_L) and (q_H, t_H) which are incentive compatible as well. Such mechanism is a direct mechanism where the consumer announces his type, and based on his announcement he is offered the contracts $(q(\theta_i), t(\theta_i))$ for $i \in \{H, L\}$. Alternatively, the monopolist just quotes a menu (q_L, t_L) and (q_H, t_H), and the consumer selects his allocation. Let the utility of a consumer of type θ_i announces that his type is θ_j be $U(\theta_j; \theta_i) = \theta_i V(q(\theta_j)) - t(q(\theta_j))$. Incentive compatibility requires that $U(\theta_i) \equiv U(\theta_i; \theta_i) \geq U(\theta_j; \theta_i)$ for $\theta_i, \theta_j \in \{\theta_H, \theta_L\}$. Hence, the monopolist's problem is to

$$\max_{\{(q_L,t_L),(q_H,t_H)\}} \quad \lambda_L[t_L - cq_L] + \lambda_H[t_H - cq_H], \tag{P$_2$}$$
$$\text{subject to} \quad U(\theta_L) \geq U(\theta_H, \theta_L), \tag{IC$_L$}$$
$$U(\theta_H) \geq U(\theta_L, \theta_H), \tag{IC$_H$}$$
$$U(\theta_L) \geq 0, \tag{IR$_L$}$$
$$U(\theta_H) \geq 0, \tag{IR$_H$}$$

Notice that (IC$_L$) and (IC$_H$) together imply

$$(\theta_H - \theta_L)V(q_L) \leq U(\theta_H) - U(\theta_L) \leq (\theta_H - \theta_L)V(q_H) \tag{A}$$
$$\implies (\theta_H - \theta_L)(V(q_H) - V(q_L)) \geq 0$$
$$\implies V(q_H) - V(q_L) \geq 0$$
$$\implies q_H > q_L, \quad \text{since} \quad V' > 0.$$

The first and the second inequalities in (A) follow from (IC_H) and (IC_L), respectively. From (A) it also follows that $U(\theta_H) \geq U(\theta_L)$. This implies that $U(\theta_L)$ will be set as low as possible. Given (IR_L) we get $U(\theta_L) = 0$, and hence

$$\theta_L V(q_L) = t_L. \tag{B}$$

We impose (IC_H) to be binding. This gives

$$\begin{aligned} & U(\theta_H) = U(\theta_L, \theta_H) \\ \implies & \theta_H V(q_H) - t_H = \theta_H V(q_L) - t_L \\ \implies & \theta_H V(q_H) - t_H = (\theta_H - \theta_L) V(q_L) \qquad \text{[using (B)]} \\ \implies & t_H = \theta_H V(q_H) - (\theta_H - \theta_L) V(q_L). \end{aligned} \tag{C}$$

Substituting (B) and (C) into the objective function one reduces the principal's maximization problem as

$$\max_{\{q_L, q_H\}} \; \lambda_L[\theta_L V(q_L) - cq_L] + \lambda_H[\theta_H V(q_H) - cq_H - (\theta_H - \theta_L)V(q_L)] \tag{P$'_2$}$$

Let the optimal solutions be $(\hat{q}_L, \hat{q}_H)$. The first order conditions give

$$\theta_L V'(\hat{q}_L)\left[1 - \frac{\lambda_H}{\lambda_L}\frac{\theta_H - \theta_L}{\theta_L}\right] = c, \tag{3.1}$$

$$\theta_H V'(\hat{q}_H) = c. \tag{3.2}$$

For an interior solution to exist we need $\lambda_L > \tilde{\lambda}_L \equiv (\theta_H - \theta_L)/\theta_H$, which means that the proportions of the low type must be sufficiently high for price discrimination to be feasible. Denote by z the term in the square bracket in equation (3.1), which is strictly less than 1. Thus we have

$$\theta_L V'(\hat{q}_L) = \frac{c}{z}. \tag{3.1$'$}$$

The optimal transfers are given by:

$$\hat{t}_L = \theta_L V(\hat{q}_L), \tag{3.3}$$

$$\hat{t}_H = \theta_H V(\hat{q}_H) - (\theta_H - \theta_L)V(\hat{q}_L). \tag{3.4}$$

Finally, we need to check that, under the above solutions, (IR_H) and (IC_L) are satisfied. Notice that (IR_H) requires that $U(\theta_H) \geq 0$. Since $U(\theta_H) = \theta_H V(\hat{q}_H) - \hat{t}_H = (\theta_H - \theta_L)V(\hat{q}_L)$, it is strictly positive given that $\theta_H > \theta_L$ and $\hat{q}_L$. Next, notice that $\hat{q}_L < q_L^* < q_H^* = \hat{q}_H$. The constraint ($\text{IC}_L$) requires that $0 = U(\theta_L) \geq U(\theta_H, \theta_L)$. Now,

$$\begin{aligned} U(\theta_H, \theta_L) = & \theta_L V(\hat{q}_H) - \hat{t}_H \\ = & \theta_L V(\hat{q}_H) - \theta_H V(\hat{q}_H) - (\theta_H - \theta_L)V(\hat{q}_L) \\ = & -(\theta_H - \theta_L)[V(\hat{q}_H) - V(\hat{q}_L)] \leq 0, \end{aligned}$$

and hence (IC_L) is satisfied. Note a few important properties of the optimal contract under asymmetric information.

1. Since $z < 1$, we have
$$\theta_L V'(\hat{q}_L) = \frac{c}{z} > c = \theta_L V'(q_L^*).$$
The above implies $\hat{q}_L < q_L^*$ since $V'' < 0$, i.e., as compared with the symmetric information case, the low type consumer always consumes a lower quantity under asymmetric information. On the other hand, from (3.2) we have $\hat{q}_H = q_H^*$, i.e., the high type consumer receives the efficient quantity. This is called the "no distortion at the top" property. The monopolist offers the efficient quantity to the high-valuation consumer, and distorts the consumption of the low type so as to make the contract for the low type consumer less attractive for the high type.

2. Under perfect price discrimination, the principal, who has all the bargaining power, is able to extract the entire rent from both types of the consumer. Since $U(\theta_H) = (\theta_H - \theta_L)V(\hat{q}_L) > 0$ under asymmetric information, such rent extraction from the high type is not possible any more, at least when the seller wants both types to participate in the mechanism. The monopolist has to give up some rent to the high-type consumer because the consumer possesses informational advantages. The term $(\theta_H - \theta_L)V(\hat{q}_L)$ is referred to as the "informational rent" of type θ_H. Clearly, the higher the marginal valuation of this type relative to θ_L, the more is his informational rent. The low-type consumer gets zero informational rent.

3. Define $S(\theta_i) := [\theta_i V(q_i) - t_i] + [t_i - cq_i]$ which is the aggregate surplus of the trade between the monopolist and the consumer of type θ_i. Using this, the seller's objective function can be written as

$$\underbrace{\lambda_H S(\theta_H) + \lambda_L S(\theta_L)}_{\text{Expected allocative efficiency}} - \underbrace{\lambda_H U(\theta_H) + \lambda_L U(\theta_L)}_{\text{Expected informational rent}}.$$

The seller thus faces a clear trade-off between efficiency and rent extraction. Rewrite the first order condition (3.1) with respect to q_L as

$$\lambda_L[\theta_L V'(\hat{q}_L) - c] = \lambda_H(\theta_H - \theta_L)V'(\hat{q}_L). \tag{3.5}$$

The above equation represents the optimal trade-off between efficiency and rent extraction under asymmetric information. The expected marginal efficiency gain [the left-hand-side of (3.5)] and the expected marginal cost of the rent [the right-hand-side of (3.5)] brought about by an infinitesimal increase of consumption of the low type are equated.

The above properties that we have obtained in the simplest possible model of adverse selection with two types in general carry through under more general specifications, which we will see in the next section.

3.3.2 Monopolistic screening with a continuum of types

We extend the model in Section 3.3.1 to a case with a continuum of types. The agent has a type $\theta \in \Theta := [\underline{\theta}, \bar{\theta}]$ that constitutes his private information. Let $W(q, t) = t - cq$ and $U(q, t, \theta) = \theta V(q) - t$ denote the utilities of the principal and the agent of type θ, respectively. We assume that $V(\cdot)$ is strictly increasing and strictly concave in q. The additively separable form of $U(q, t, \theta)$ implies that the agent's utility is quasilinear in consumption and money. The principal entertains an a priori belief on the agent's type that is embodied in a cumulative distribution function $F(\theta)$ with density $f(\theta) > 0$ for all $\theta \in \Theta$.

Incentive compatibility

We restrict our attention to a *direct mechanism*, $g(\theta') = (q(\theta'), t(\theta'))$ where θ' is the announced type of the consumer. We use the following notations:

$$U(\theta', \theta) \equiv U(q(\theta'), t(\theta'), \theta) = \theta V(q(\theta')) - t(\theta'),$$
$$U(\theta) \equiv U(q(\theta), t(\theta), \theta) = \theta V(q(\theta)) - t(\theta).$$

The first of the above two expressions is the utility of the agent of type θ when he announces his type to be θ', while the second one is the utility of a type θ from telling the truth. A mechanism $g(\theta)$ must satisfy the following individual rationality and incentive compatibility constraints:

$$U(\theta) \geq 0 \quad \text{for all} \quad \theta \in \Theta, \tag{IR$_\theta$}$$
$$U(\theta) \geq U(\theta', \theta) \quad \text{for all} \quad (\theta, \theta') \in \Theta \times \Theta, . \tag{IC$_\theta$}$$

Note that (IC$_\theta$) can equivalently be written as

$$\theta = \underset{\theta'}{\text{argmax}} \ \left\{\theta V(q(\theta')) - t(\theta')\right\}. \tag{IC$'_\theta$}$$

We first prove a very important result.

Proposition 3.1: Incentive compatible mechanism

A direct mechanism $g(\cdot)$ is incentive compatible, i.e., satisfies (IC$_\theta$) if and only if the following two conditions hold for all $\theta \in \Theta$:

$$\theta V'(q(\theta))q'(\theta) - t'(\theta) = 0, \tag{FOC}$$
$$q'(\theta) \geq 0. \tag{MON}$$

Proof. We first prove the necessity of (FOC) and (MON). For any two θ and θ' with $\theta \neq \theta'$, (IC$_\theta$) implies that

$$\theta V(q(\theta)) - t(\theta) \geq \theta V(q(\theta')) - t(\theta'),$$
$$\theta' V(q(\theta')) - t(\theta') \geq \theta' V(q(\theta)) - t(\theta).$$

Summing the above two we obtain

$$(\theta' - \theta)[V(q(\theta')) - V(q(\theta))] \geq 0.$$

If $\theta' > \theta$, then the above inequality implies that $q(\theta') \geq q(\theta)$ since $V'(\cdot) > 0$. As $q(\cdot)$ must be non-decreasing, it is differentiable almost everywhere (a. e.). Therefore, $t(\cdot)$ is also differentiable with the same points of non-differentiablity. The local first-order necessary condition associated with (IC$'_\theta$) at $\theta' = \theta$ (truth-telling) is given by:

$$\theta V'(q(\theta))q'(\theta) - t'(\theta) = 0. \tag{FOC}$$

The local second-order necessary condition at $\theta' = \theta$ is given by:

$$\theta V''(q(\theta))[q'(\theta)]^2 + \theta V'(q(\theta))q''(\theta) - t''(\theta) \leq 0. \tag{SOC}$$

On the other hand, differentiating (FOC) with respect to θ, we obtain:

$$V'(q(\theta))q'(\theta) + \underbrace{\theta V''(q(\theta))[q'(\theta)]^2 + \theta V'(q(\theta))q''(\theta) - t''(\theta)}_{\leq 0 \text{ by (SOC)}} = 0,$$

and hence, $q'(\theta) \geq 0$ since $V'(\cdot) > 0$.

To show the sufficiency of (FOC) and (MON), suppose that these two conditions hold for all $\theta \in \Theta$, but by contradiction, for at least one type θ the buyer's incentive compatibility is violated, i.e.,

$$\theta V(q(\theta')) - t(\theta') > \theta V(q(\theta)) - t(\theta)$$

for at least one $\theta' \neq \theta$. Integrating the above we get

$$\int_{\theta}^{\theta'} \left[\theta V'(q(x))q'(x) - t'(x)\right] dx > 0. \tag{3.6}$$

Suppose that $\theta' > \theta$. For any $\theta \leq x$, we have

$$\theta V'(q(x)) < xV'(q(x))$$

because $V' > 0$. Thus, $q'(\cdot) \geq 0$ implies that

$$\theta V'(q(x))q'(x) < xV'(q(x))q'(x).$$

It follows from the above inequality that

$$\int_{\theta}^{\theta'} \left[\theta V'(q(x))q'(x) - t'(x)\right] dx < \underbrace{\int_{\theta}^{\theta'} \left[xV'(q(x))q'(x) - t'(x)\right] dx = 0}_{\text{because (FOC) holds for } x}. \tag{3.7}$$

Thus, (3.6) and (3.7) contradict each other. If $\theta' < \theta$, the same logic leads to similar contradiction. This completes the proof. □

The above proposition permits us to replace an infinity of constraints in (IC$_\theta$) by only two constraints—namely, (FOC) and (MON).

Solving the model

Now consider the individual rationality constraint (IR$_\theta$) of the buyer of type θ, which holds in particular for $\theta = \underline{\theta}$, i.e.,

$$U(\underline{\theta}) \equiv \underline{\theta}V(q(\underline{\theta})) - t(\underline{\theta}) \geq 0. \tag{IR$_{\underline{\theta}}$}$$

Given that the incentive compatibility holds for any type $\theta > \underline{\theta}$, this implies

$$U(\theta) \equiv \theta V(q(\theta)) - t(\theta) \geq \theta V(q(\underline{\theta})) - t(\underline{\theta}) > \underline{\theta}V(q(\underline{\theta})) - t(\underline{\theta}) \geq 0.$$

Therefore, the only relevant individual rationality constraint, among the infinitely many ones in (IR$_\theta$), we have to take into account is the constraint (IR$_{\underline{\theta}}$). Therefore, the monopolist's maximization problem can be written as

$$\max_{\{q(\theta),t(\theta)\}} \int_{\underline{\theta}}^{\bar{\theta}} [t(\theta) - cq(\theta)]d\theta,$$
$$\text{subject to (IR}_{\underline{\theta}}\text{), (MON) and (FOC).}$$

Note first that (IR$_{\underline{\theta}}$) must bind at the optimum, otherwise the seller can increase the transfer $t(\underline{\theta})$ a little bit by holding the quantity $q(\underline{\theta})$ constant, the type $\underline{\theta}$ buyer will still accept the contract and the seller would increase her expected profits. Therefore, we will have $U(\underline{\theta}) = 0$ or, $t(\underline{\theta}) = \underline{\theta}V(q(\underline{\theta}))$. We will ignore (MON) for the time being, and solve the relaxed problem with binding (IR$_{\underline{\theta}}$). Recall that

$$U(\theta) = \theta V(q(\theta)) - t(\theta) = \max_{\theta'} \{\theta' V(q(\theta')) - t(\theta')\}.$$

Applying the Envelope theorem to the above we get

$$U'(\theta) = V(q(\theta))$$

or, integrating

$$U(\theta) - U(\underline{\theta}) = U(\theta) = \int_{\underline{\theta}}^{\theta} V(q(x))dx \equiv R(\theta). \tag{R$_\theta$}$$

The above expression is the *informational rent* that accrues to a type $\theta > \underline{\theta}$ buyer, which is similar to that in the two-type case. From the above we have

$$t(\theta) = \theta V(q(\theta)) - R(\theta).$$

Thus, the seller's objective function reduces to:

$$W = \int_{\underline{\theta}}^{\bar{\theta}} \left[\theta V(q(\theta)) - \int_{\underline{\theta}}^{\theta} V(q(x))dx - cq(\theta)\right] f(\theta)d\theta. \tag{3.8}$$

Note that

$$\underbrace{\int_{\underline{\theta}}^{\bar{\theta}} \underbrace{\left[\int_{\underline{\theta}}^{\theta} V(q(x))dx\right]}_{R(\theta)} f(\theta)d\theta}_{\text{expected informational rent}} = \left[\left\{\int_{\underline{\theta}}^{\theta} V(q(x))dx\right\} F(\theta)\right]_{\underline{\theta}}^{\bar{\theta}} - \int_{\underline{\theta}}^{\theta} V(q(\theta))F(\theta)d\theta$$

$$= \int_{\underline{\theta}}^{\bar{\theta}} V(q(\theta))d\theta - \int_{\underline{\theta}}^{\bar{\theta}} V(q(\theta))F(\theta)d\theta$$

$$= [1 - F(\theta)] \int_{\underline{\theta}}^{\bar{\theta}} V(q(\theta))d\theta$$

$$= \int_{\underline{\theta}}^{\bar{\theta}} \left[\frac{1 - F(\theta)}{f(\theta)}\right] V(q(\theta))f(\theta)d\theta.$$

Thus, the principal's objective function is given by:

$$\int_{\underline{\theta}}^{\bar{\theta}} [t(\theta) - cq(\theta)]f(\theta)d\theta$$

$$= \underbrace{\int_{\underline{\theta}}^{\bar{\theta}} [\theta V(q(\theta)) - cq(\theta)]f(\theta)d\theta}_{\text{expected allocative efficiency}} - \underbrace{\int_{\underline{\theta}}^{\bar{\theta}} R(\theta)f(\theta)d\theta}_{\text{expected informational rent}}$$

$$= \int_{\underline{\theta}}^{\bar{\theta}} \left[z(\theta)V(q(\theta)) - cq(\theta)\right] f(\theta)d\theta, \tag{P$_3$}$$

where

$$z(\theta) := \theta - \frac{1 - F(\theta)}{f(\theta)}$$

is the virtual valuation of the agent. Notice that the principal's objective function is a function of $q(\theta)$ alone. Call $\hat{q}(\theta)$ the optimal solution to the maximization problem whose first order condition is given by:

$$z(\theta)V'(\hat{q}(\theta)) = c, \quad \text{for all} \quad \theta \in \left[\underline{\theta}, \bar{\theta}\right].$$

Notice that so far we have suppressed the monotonicity constraint, i.e., $q'(\theta) \geq 0$ for all $\theta \in \Theta$. Differentiating the above optimality condition, we get

$$\hat{q}'(\theta) = -\frac{z'(\theta)V'(\hat{q}(\theta))}{z(\theta)V''(\hat{q}(\theta))},$$

which implies that

$$\text{sign}[\hat{q}'(\theta)] = \text{sign}[z'(\theta)] = \text{sign}\left[1 - \frac{d}{d\theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right)\right].$$

Hence, $\hat{q}(\theta)$ is non-decreasing if and only if

$$\frac{d}{d\theta}\left[\frac{1 - F(\theta)}{f(\theta)}\right] \leq 1. \tag{3.9}$$

Now define by

$$h(\theta) := \frac{f(\theta)}{1 - F(\theta)}$$

the *hazard rate* associated with the distribution function $F(\theta)$. Notice that

$$h'(\theta) = \frac{d}{d\theta}\left[\frac{f(\theta)}{1 - F(\theta)}\right] > 0 \tag{MHR}$$

implies that the condition (3.9) is satisfied. The condition (MHR) is called the *monotone hazard rate property* of the distribution function $F(\theta)$, which is a sufficient condition for the optimum consumption allocation is non-decreasing in types. Notice that $\hat{q}(\theta)$ non-decreasing implies that the types are separated and truthfully revealed. This is a *separating* contract. Finally, notice that

$$V'(q^*(\theta)) = \frac{c}{\theta} < \frac{c}{z(\theta)} = V'(\hat{q}(\theta)) \implies \hat{q}(\theta) < q^*(\theta) \quad \text{for all} \quad \theta \in \left[\underline{\theta}, \, \bar{\theta}\right),$$

where $q^*(\theta)$ is the first best consumption allocation. Since $z(\bar{\theta}) = \bar{\theta}$, we have $\hat{q}(\bar{\theta}) = q^*(\bar{\theta})$, i.e., the "no-distortion-at-the-top" property is satisfied.

Pooling or bunching mechanism

What happens if the monotone hazard rate property does not hold? Since (MHR) is only a sufficient condition for $\hat{q}(\theta)$ to be non-decreasing on Θ, if this property is not satisfied $\hat{q}(\theta)$ may be strictly decreasing over a subset of Θ. Therefore, in solving the principal's maximization problem we cannot ignore anymore the monotonicity constraint. Call the solution to the principal's constraint maximization problem $\bar{q}(\theta)$, rewrite the maximization problem as the following:

$$\max_{q(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} \left[z(\theta)V(q(\theta)) - cq(\theta)\right] f(\theta) d\theta \tag{P$'_3$}$$

$$\text{subject to} \quad q'(\theta) = \mu(\theta), \tag{3.10}$$

$$\mu(\theta) \geq 0. \tag{3.11}$$

The above maximization is an optimal control problem, the Hamiltonian of which is given by:

$$H(\theta, \, q, \, \mu, \, \lambda) = \left[z(\theta)V(q(\theta)) - cq(\theta)\right] f(\theta) + \lambda(\theta)\mu(\theta).$$

Assume that $H(\theta, \, q, \, \mu, \, \lambda)$ is concave in q. Then by Pontryagin's maximum principle, the necessary and sufficient conditions for an optimum $(\bar{q}(\theta), \, \bar{\mu}(\theta))$ are given by:

$$H(\theta, \, \bar{q}(\theta), \, \bar{\mu}(\theta), \, \bar{\lambda}(\theta)) \geq H(\theta, \, \bar{q}(\theta), \, \mu(\theta), \, \lambda(\theta)); \tag{3.12}$$

$$\left[z(\theta)V'(\bar{q}(\theta)) - c\right] f(\theta) + \lambda'(\theta) = 0 \quad \text{for almost all} \quad \theta \in \Theta; \tag{3.13}$$

$$\text{and the transversality conditions} \quad \lambda(\underline{\theta}) = \lambda(\bar{\theta}) = 0 \quad \text{are satisfied.} \tag{3.14}$$

Integrating (3.13), we get

$$\lambda(\theta) = -\int_{\underline{\theta}}^{\theta} \left[z(\theta)V'(\bar{q}(\theta)) - c\right] f(\theta) d\theta,$$

and using the transversality conditions, we then have

$$0 = \lambda(\bar{\theta}) = -\int_{\underline{\theta}}^{\bar{\theta}} \left[z(\theta)V'(\bar{q}(\theta)) - c\right] f(\theta) d\theta.$$

Next, since (3.12) requires that $\mu(\theta)$ must maximiza H subject to $\mu(\theta) \geq 0$, wh must have $\lambda(\theta) \leq 0$, i.e.,

$$\int_{\underline{\theta}}^{\theta} \left[z(\theta)V'(\bar{q}(\theta)) - c\right] f(\theta) d\theta \geq 0.$$

Whenever $\lambda(\theta) < 0$, we get

$$\bar{\mu}(\theta) = \bar{q}'(\theta) = 0.$$

Thus we get the following complementary slackness condition:

$$\bar{q}'(\theta) \cdot \int_{\underline{\theta}}^{\theta} \left[z(\theta)V'(\bar{q}(\theta)) - c\right] f(\theta) d\theta = 0.$$

Now let $\bar{q}(\theta)$ is constant over the interval $[\theta_1, \theta_2] \subset \Theta$, and strictly increasing otherwise. Clearly over $\Theta \setminus [\theta_1, \theta_2]$, $\bar{q}(\theta)$ coincides with $\hat{q}(\theta)$ since $\bar{q}'(\theta) > 0$ implies that $\lambda(\theta) = 0$ which in turn implies that

$$\lambda'(\theta) = 0 \quad \Longrightarrow \quad z(\theta)V'(\bar{q}(\theta)) = c.$$

A constant $q(\theta)$ implies that the principal offers the same allocations $(\bar{q}(\theta), \bar{t}(\theta))$ to all types $\theta \in [\theta_1, \theta_2]$. Such contract is called a *bunching* or *pooling* contract. Notice that, by continuity of $\lambda(\theta)$ we must have $\lambda(\theta_1) = \lambda(\theta_2) = 0$, so that

$$\int_{\underline{\theta}_1}^{\theta_2} \left[z(\theta)V'(\bar{q}(\theta)) - c\right] f(\theta)d\theta = 0. \tag{3.15}$$

On the other hand, by continuity of $\bar{q}(\theta)$ we have

$$\hat{q}(\theta_1) = \hat{q}(\theta_2). \tag{3.16}$$

Therefore, solving the equations (3.15) and (3.16) we get the optimal bunching interval $[\theta_1, \theta_2]$. For further discussions on the bunching contracts, see (Bolton and Dewatripont, 2005, Chapter 2).

3.4 A general model of adverse selection: non-quasilinear utility

We assume that the buyer's utility is given by $U(q, t, \theta)$ with $U_q > 0$, $U_t < 0$ and $U_\theta \geq 0$. In order to have a proposition similar to Proposition 3.1, we require to impose more structure of the buyer's utility function—namely, the Spence-Mirrlees condition or the single-crossing condition, which is given by:

$$\frac{\partial}{\partial \theta}\left[-\frac{U_q(q, t, \theta)}{U_t(q, t, \theta)}\right] > 0 \quad \text{for all } (q, t, \theta). \tag{SM}$$

Note that the slope of the indifference curve of any type θ buyer, $U(q, t, \theta) = \bar{u}$ is given by $-(U_q/U_t) > 0$. Thus, condition (SM) implies that, given any two types θ and θ' with $\theta' > \theta$, the indifference curve of θ' is everywhere steeper than that of θ, and hence, they cross only once. This condition means that if type θ is indifferent between two consumption-transfer pairs (q, t) and (q', t') with $(q, t) < (q', t')$, then the higher type θ' is willing to pay more than t' to receive consumption q'. Thus, in an incentive compatible mechanism, the principal is able to separate the different types of the agent by offering larger allocations q to higher types and making them pay for the privilege. Therefore, the Spence-Mirrlees condition is also called the sorting condition. The following result is a generalization of Proposition 3.1 to a non-quasilinear environment.

Proposition 3.2

Suppose the buyer's utility function $U(q, t, \theta)$ satisfies condition (SM). Then, a direct mechanism $g(\cdot)$ is incentive compatible if and only if the following two conditions hold for all $\theta \in \Theta$:

$$U_q(q(\theta), t(\theta), \theta)q'(\theta) + U_t(q(\theta), t(\theta), \theta)t'(\theta) = 0, \tag{FOC}$$
$$q'(\theta) \geq 0. \tag{MON}$$

Proof. We restrict attention to differentiable mechanisms $(q(\theta'), t(\theta'))$. Incentive compatibility is given by:

$$U(q(\theta), t(\theta), \theta) = \max_{\theta'} \{U(q(\theta'), t(\theta'), \theta)\}. \tag{IC$_\theta$}$$

We first prove the necessity of (FOC) and (MON). The local first-order necessary condition associated with (IC$_\theta$) at $\theta' = \theta$ (truth-telling) is given by:

$$U_q(q(\theta), t(\theta), \theta)q'(\theta) + U_t(q(\theta), t(\theta), \theta)t'(\theta) = 0. \tag{FOC}$$

The local second-order necessary condition at $\theta' = \theta$ is given by:

$$\begin{aligned}\Omega(\theta) \equiv &[U_{qq}(q(\theta), t(\theta), \theta)q'(\theta) + U_{qt}(q(\theta), t(\theta), \theta)t'(\theta)]q'(\theta) + U_q(q(\theta), t(\theta), \theta)q''(\theta) \\ &+[U_{tq}(q(\theta), t(\theta), \theta)q'(\theta) + U_{tt}(q(\theta), t(\theta), \theta)t'(\theta)]t'(\theta) + U_t(q(\theta), t(\theta), \theta)t''(\theta) \leq 0. \end{aligned} \tag{SOC}$$

On the other hand, differentiating (FOC) with respect to θ, we obtain:

$$U_{q\theta}(q(\theta), t(\theta), \theta)q'(\theta) + U_{t\theta}(q(\theta), t(\theta), \theta)t'(\theta) + \underbrace{\Omega(\theta)}_{\leq 0 \text{ by (SOC)}} = 0,$$

which implies that

$$\begin{aligned} &U_{q\theta}(q(\theta), t(\theta), \theta)q'(\theta) + U_{t\theta}(q(\theta), t(\theta), \theta)t'(\theta) \geq 0 \\ \iff &q'(\theta)\left[U_{q\theta}(q(\theta), t(\theta), \theta) - U_{t\theta}(q(\theta), t(\theta), \theta) \cdot \frac{U_q(q(\theta), t(\theta), \theta)}{U_t(q(\theta), t(\theta), \theta)}\right] \geq 0 \\ \iff &q'(\theta) \cdot -U_t(q(\theta), t(\theta), \theta) \cdot \frac{\partial}{\partial \theta}\left[-\frac{U_q(q, t, \theta)}{U_t(q, t, \theta)}\right] \geq 0. \end{aligned}$$

Thus, under (SM), the above implies that $q'(\theta) \geq 0$.

Next, we prove the sufficiency of (FOC) and (MON) for incentive compatibility. Global incentive compatibility requires

$$\begin{aligned} &U(q(\theta), t(\theta), \theta) \geq U(q(\theta'), t(\theta'), \theta') \quad \text{for all} \quad (\theta, \theta') \in \Theta \times \Theta \\ \iff &\int_{\theta'}^{\theta} \left[U_q(q(x), t(x), \theta)q'(x) + U_t(q(x), t(x), \theta)t'(x)\right] dx \geq 0 \\ \iff &\int_{\theta'}^{\theta} q'(x) \cdot -U_t(q(x), t(x), \theta)\left[-\frac{U_q(q(x), t(x), \theta)}{U_t(q(x), t(x), \theta)} - \left\{-\frac{U_q(q(x), t(x), x)}{U_t(q(x), t(x), x)}\right\}\right] dx \geq 0. \end{aligned} \tag{3.17}$$

The last inequality arises because (FOC) holds for any $x \in [\theta', \theta]$. Because $q'(x) \geq 0$, $U_t < 0$, and using (SM) we can conclude that

$$\begin{aligned} &\int_{\theta'}^{\theta} q'(x) \cdot -U_t(q(x), t(x), \theta)\left[-\frac{U_q(q(x), t(x), \theta)}{U_t(q(x), t(x), \theta)} - \left\{-\frac{U_q(q(x), t(x), x)}{U_t(q(x), t(x), x)}\right\}\right] dx \\ \geq &\int_{\theta'}^{\theta} q'(x) \cdot -U_t(q(x), t(x), \theta)\left[-\frac{U_q(q(x), t(x), x)}{U_t(q(x), t(x), x)} - \left\{-\frac{U_q(q(x), t(x), x)}{U_t(q(x), t(x), x)}\right\}\right] dx = 0, \end{aligned}$$

which proves (3.17). This completes the proof of the Proposition. □

To solve the model explicitly we of course require specific form of the buyer's utility function $U(q, t, \theta)$. Once that is known, we can follow procedure similar to that in Section 3.3.2.

Marginal cost pricing under full information

Now suppose that the taxes are not distortionary, i.e., $\lambda = 0$. The the social welfare is given by:

$$\begin{aligned} W &= S(q) - qP(q) - t + \alpha U \\ &= S(q) - \theta q - (1-\alpha)U. \end{aligned} \tag{3.18}$$

The regulator, by choosing q maximizes the above expression subject to $U \geq 0$, which is the individual rationality constraint of the firm. This constraint will bind at the optimum, and hence the regulator's maximization problem reduces to:

$$\max_q S(q) - \theta q. \tag{3.19}$$

The first order condition of the above maximization problem implies that

$$P(q^*) = \theta. \tag{3.20}$$

When there is no shadow cost of public fund, the product is priced at its marginal cost.

3.4.1 Optimal regulation under asymmetric information

First we describe some general feature of a direct incentive compatible mechanism $g(\theta) = (q(t), t(\theta))$. As usual, define by

$$U(\theta', \theta) = t(\theta') + q(\theta')P(q(\theta')) - \theta q(\theta')$$

the profit of a type θ monopolist when he announces to be of type θ. By defining $U(\theta) := U(\theta, \theta)$, we have

$$U(\theta', \theta) = U(\theta') + (\theta' - \theta)q(\theta').$$

Then the incentive compatibility constraint of type θ is given by:

$$U(\theta) = \max_{\theta'}\{U(\theta') + (\theta' - \theta)q(\theta')\}. \tag{IC$_\theta$}$$

And the individual rationality constraint is given by:

$$U(\theta) \geq 0 \quad \text{for all} \quad \theta \in \Theta. \tag{IR$_\theta$}$$

A mechanism $g(\theta)$ is feasible if it satisfies (IR$_\theta$) and (IC$_\theta$) for all θ. Then, as Proposition 3.1, we can write the following characterization result.

Proposition 3.3

A regulatory mechanism $g(\theta) = (q(\theta), t(\theta))$ is feasible if and only if it satisfies the following conditions for all $\theta \in \Theta$:

(a) $U(\bar{\theta}) \geq 0$;

(b) $q(\theta)$ is non-increasing in θ;

(c) Local incentive compatibility constraint:

$$U'(\theta) + q(\theta) = 0. \tag{LIC$_\theta$}$$

Proof. Similar to the proof of Proposition 3.1. □

From the local incentive compatibility and binding individual rationality of type $\bar{\theta}$, i.e., $U(\bar{\theta}) = 0$, we get

$$U(\theta) = \int_{\theta}^{\bar{\theta}} q(x)dx \equiv R(\theta), \tag{3.21}$$

where $R(\theta)$ is the informational rent of type θ. From the above it also follows that the average informational rent is given by:

$$\int_{\underline{\theta}}^{\bar{\theta}} R(\theta)f(\theta)d\theta = \int_{\underline{\theta}}^{\bar{\theta}} q(\theta)L(\theta)f(\theta)d\theta,$$

where $L(\theta) := F(\theta)/f(\theta)$. Also,

$$t(\theta) = -[P(q(\theta)) - \theta]q(\theta) + U(\theta) = -[P(q(\theta)) - \theta]q(\theta) + R(\theta).$$

Ramsey pricing under asymmetric information

When the taxes are distortionary, using the expressions for $t(\theta)$ and the expected informational rent, the regulator's objective function reduces to:

$$\int_{\underline{\theta}}^{\bar{\theta}} [S(q(\theta)) + \lambda q(\theta)P(q(\theta)) - (1+\lambda)z(\theta)q(\theta)]\, f(\theta)d\theta,$$

where

$$z(\theta) := \theta + \frac{\lambda}{1+\lambda} L(\theta)$$

is the virtual marginal cost of the monopolist. The regulator chooses $q(\theta)$ in order to maximize the expected welfare such that $q'(\theta) \leq 0$. We ignore for the time being the monotonicity constraint. Let the optimal solution be $q(\theta)$ and $P(\theta) := P(q(\theta))$. Then the first order condition of the above maximization problem implies that

$$L(\theta, z) := \frac{P(\theta) - z(\theta)}{P(\theta)} = \frac{\lambda}{1+\lambda} \frac{1}{\eta(\theta)} .$$

The above is the virtual Lerner index of the regulated firm. When the regulated firm has private information regarding its marginal cost of production, the regulator perceives that its marginal cost is equal to its virtual value $z(\theta)$ which is higher than θ. Hence, as compared with the full information case, the efficient mechanism is offered to $z(\theta)$ instead of θ. The true Lerner index is given by:

$$L(\theta) := \frac{P(\theta) - \theta}{P(\theta)} = \underbrace{\frac{\lambda}{1+\lambda} \frac{1}{\eta(\theta)}}_{\text{Ramsey markup}} + \underbrace{\frac{\lambda}{1+\lambda} \frac{L(\theta)}{P(\theta)}}_{\text{Incentive correction}} .$$

Notice that, under asymmetric information, the true Lerner index of the regulated firm consists of a Ramsey markup term and an incentive correction term. The second term was absent in the full information case since no incentive correction via distortion was necessary. Note that, the no-distortion-at-the-top property holds here too since $z(\underline{\theta}) = \underline{\theta}$, i.e, the most efficient type of the monopolist receives the efficient contract. Finally, under the assumption of monotone hazard rate, at the optimum we have $q(\theta)$ a non-increasing function of the firm type.

Marginal cost pricing under asymmetric information

In the model of Baron and Myerson (1982), using the above procedure the regulator's objective function boils down to:

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[S(q(\theta)) - z(\theta) q(\theta) \right] f(\theta) d\theta,$$

where

$$z(\theta) := \theta + (1-\alpha) L(\theta)$$

is the virtual marginal cost of the monopolist. Notice that the virtual type in this case is different from that in the model of Laffont and Tirole (1986). From the first order condition of the regulator's maximization problem it follows that

$$P(\theta) = z(\theta) = \underbrace{\theta}_{\text{Marginal cost}} + \underbrace{(1-\alpha) L(\theta)}_{\text{Incentive correction}} .$$

The above pricing rule adheres to the marginal cost pricing where the price charged by the regulated firm is equated with her virtual marginal cost of production, which is the sum of the true marginal cost and an incentive correction term.

3.5 Subjective Performance Evaluation

In the standard principal-agent model under moral hazard [in Chapter 1] the performance q of the relationship, which is a noisy signal of the agent's effort, has been assumed to be publicly verifiable, and hence any contract can be made contingent on q. Let q represents the quality of a good produced by the agent. Often the agent does not possess ability to assess the quality, but the principal does. On the basis of such assessment, the principal designs a contract for the agent. Such situation is called the *subjective performance evaluation* (SPE) as opposed to objective evaluation where q is publicly observable. In what follows we discuss two models of incentive contracting under SPE.

3.5.1 A Simple Model of Optimal Money Burning Contract

Consider a simple version of MacLeod (2003) where a risk neutral principal (P) hires a risk neutral agent (A) to work on a project. The agent spends non-verifiable effort $e \in \{0, 1\}$ by incurring a total cost $\psi(e) = \varphi e$ with $\varphi > 0$. Agent's effort choice induces a probability distribution over the stochastic return (performance) $q \in \{q_H, q_L\}$ with $q_H > q_L \geq 0$ of the project which is given by:

$$\begin{aligned} \text{Pr.}[q = q_H \mid e = 0] &= \pi_0, \\ \text{Pr.}[q = q_H \mid e = 1] &= \pi_1, \end{aligned}$$

with $0 < \pi_0 < \pi_1 < 1$. Principal observes the realization of the return privately, and sends message $m \in M = \{m_H, m_L\}$ to the agent. Under SPE, a contract between the principal and the agent is contingent on the message sent by the principal. We may restrict attention to the direct revelation mechanism, i.e., $M = \{y_H, y_L\}$.

Now consider standard bonus contract $\gamma = (z, b)$ under limited liability discussed in Chapter 1, where z is the fixed salary and b is the bonus if $q = q_H$. Let $V_i(\gamma, q_j)$ be P's payoff if she observes $q = q_i$ for $i = H, L$, but reports $m = q_j$ for $j = H, L$. Then,

$$\begin{aligned} V_H(\gamma, q_H) &= q_H - z - b, \\ V_H(\gamma, q_L) &= q_H - z. \end{aligned}$$

If $b > 0$, then $V_H(\gamma, q_H) < V_H(\gamma, q_L)$, i.e., P never has incentives to report truthfully. One way to induce principal to report truthfully is that P puts b on the table before A exerts effort, and this amount goes to A if P reports q_H, and is destroyed otherwise. In this case $V_H(\gamma, q_L) = q_H - z - b$ which implies that $V_H(\gamma, q_H) = V_H(\gamma, q_L)$. Thus, the organization must burn the amount b in order to incentivize the principal to report truthfully.

We now analyze the optimal money burning contract for such organization. Assume that P has a fixed budget W. A contingent money burning contract $\gamma = (w_H, w_L, b_H, b_L)$ is such that for $i = H, L$

$$\begin{aligned} &w_i = w(q_i), \qquad b_i = b(q_i), \\ &w_i + b_i = W \end{aligned}$$

Thus, a contract can be alternatively represented by $\gamma = (W, b_H, b_L)$. We assume that $q_H - q_L$ is high enough so that P always wants to implement the high effort $e = 1$.

Under a money burning contract A's expected payoffs are given by:

$$U(e = k \mid \gamma) = W - [\pi_k b_H + (1 - \pi_k) b_L] - \varphi k \quad \text{for} \quad k = 0, 1.$$

Therefore, the agent's incentive compatibility implies

$$U(e = 1 \mid \gamma) \geq U(e = 0 \mid z) \iff b_L - b_H \geq \frac{\varphi}{\pi_1 - \pi_0}. \tag{ICA}$$

The agent is protected by limited liability, i.e., his state-contingent income must be non-negative, which is given by:

$$b_L \leq W, \qquad b_H \leq W. \tag{LL}$$

Feasibility of money burning, on the other hand, requires

$$b_L \geq 0, \qquad b_H \geq 0. \tag{F}$$

Agent's outside option is normalized to 0, and hence his individual rationality is given by:

$$W - [\pi_1 b_H + (1-\pi_1) b_L] - \varphi \geq 0. \tag{IRA}$$

It is easy to verify that given constraints (ICA) and (LL), the individual rationality constraint is automatically satisfied. On the other hand, (ICA) implies that $b_L \geq b_H$, and hence the constraints $b_H \leq W$ and $b_L \geq 0$ can be ignored. The optimal money burning contract $\gamma^* = (W^*, b_H^*, b_L^*)$ solves the following minimization problem:

$$\begin{aligned} &\min_{\{W, b_H, b_L\}} \quad W \\ &\text{subject to } \ (\text{IC}_a), \ \ b_L \leq W \ \text{ and } \ b_H \geq 0. \end{aligned}$$

The following lemma characterizes the optimal money burning contract.

Lemma 3.1: Optimal money burning contract

The optimal money burning contract $z^* = (w_H^*, w_L^*, b_H^*, b_L^*)$ is given by:

$$\begin{aligned} w_H^* &= \frac{\varphi}{\pi_1 - \pi_0} > 0, \qquad w_L^* = 0, \\ b_L^* &= \frac{\varphi}{\pi_1 - \pi_0} > 0, \qquad b_H^* = 0. \end{aligned}$$

The above result is fairly intuitive. The state-contingent wages paid to the agent must be different in order to elicit effort. The wage paid in state H must be destroyed if the principal reports a low performance, i.e., $w_H^* = b_L^*$ so that she does not have incentives to misreport. Thus, the organization incurs a deadweight loss which is equal to $\pi_1\varphi/(\pi_1 - \pi_0)$, the expected money burning.

Clearly, it is necessary for an organization destroy resources in order to ensure incentive compatibility. What happens if money burning is not feasible? Then things are even worse in this simple model. Note that when $b_H = b_L = 0$, with a fixed budget the principal must commit a fixed wage schedule $w_H = w_L = \bar{w}$ to the agent. This destroys A's incentive to exert effort since

$$U(e = 1 \mid \bar{w}) = \bar{w} - \varphi < \bar{w} = U(e = 0 \mid \bar{w}).$$

Therefore,

Lemma 3.2

If money burning is not feasible under SPE, then there is no contract that implements high effort $e = 1$.

Thus, in order to implement the high effort some degree of inefficiency via money burning is necessary for an organization. Several attempts (e.g. Levin, 2003; MacLeod, 2003; Fuchs, 2007; Khalil, Lawarrée, and Scott, 2015) have been made to search for contractual instruments to minimize expected money burning in an organization.

Part III

Markets and mechanisms

Chapter 4

Mechanism design

4.1 A simple market: An exchange economy

An exchange economy or market consists of two individuals (traders), denoted by $i = 1, 2$, and two goods, denoted by $j = 1, 2$. Each trader is born with an endowment of each good. Formally, $\omega_{ij} \geq 0$ is trader i's endowment of good j. The aggregate endowment of good j is thus given by $\omega_j = \omega_{1j} + \omega_{2j}$.

Each individual has preferences over the bundles of the two commodities, which is assumed to be rational (complete and transitive). We would assume that preferences are continuous so that they can be represented by a utility function $u : X \to \mathbb{R}$ where $X \subseteq \mathbb{R}^2_+$ is the set of all consumption bundles. Consumption bundles are denoted by x, y, z, etc. where x_{ij} denotes trader i's consumption of good j.

The market with two traders and two goods are represented by the Edgeworth box in Figure 4.1. The two corners represents the origins of the two agents—the southwest corner corresponds to trader 1 and the northeast corner, to trader 2. The endowment is denoted by the point ω. The blue curves are agent 1's indifference curves given their preferences $u(x_{11}, x_{12})$. The red curves are the indifference curves given their preferences $v(x_{21}, x_{22})$. The size of the Edgeworth box is determined by the aggregate endowments of the two goods—the length represents ω_1 and the width represents ω_2.

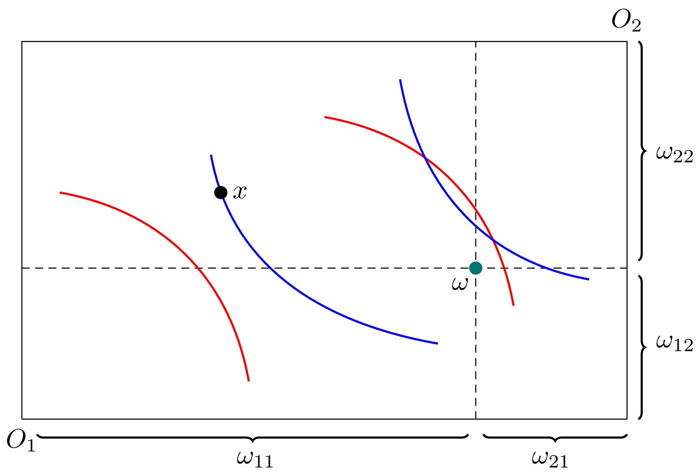


Figure 4.1: A market with two traders.

Now we turn to the concept of an allocation in the Edgeworth box. In Figure 4.1, the point x is an allocation with the property that the aggregate consumption of any good must exhaust its aggregate endowment. Formally,

Definition 4.1: Allocation

A point x in the Edgeworth box is a (consumption) allocation if $x_{1j} + x_{2j} = \omega_j$ for $j = 1, 2$. A set of allocations is given by:

$$A = \{x \mid x_{ij} \geq 0 \text{ for all } i = 1, 2, \; j = 1, 2 \text{ and } x_{1j} + x_{2j} = \omega_j \text{ for } j = 1, 2\}.$$

Definition 4.1 asserts that if a point x represents an allocation, then it cannot be outside the Edgeworth box. In other words, the way an allocation is defined incorporates the notion of feasibility, and in our context, "allocation" is simply a short-hand for "feasible allocation". Also, A, the set of allocations is the entire box. A simpler version of the Edgeworth box economy is obtained when there is only one good (a good without subscript j). Let it's aggregate endowment be $\omega = \omega_1 + \omega_2$. A (consumption) allocation, x of this economy depicted in Figure 4.2. An allocoation $x = (x_1, x_2)$ is simply a point on the line (of length ω) so that $x_1 + x_2 = \omega_1 + \omega_2 = \omega$.

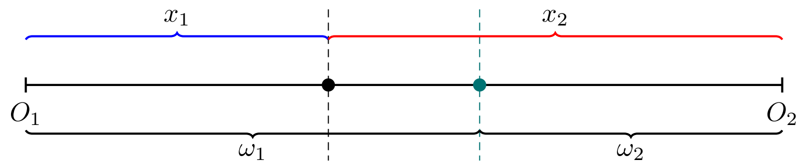


Figure 4.2: A market with two traders and one good.

Every point like x on the line represents an allocation, and, conversely, every allocation can be represented by a point, or division of the line.

In what follows, we would also assume that the traders are self-interested, i.e., the preferences of no trader depends only on their own consumption bundles, and not on that of the other trader. Formally,

$$u(x_{11}, x_{12}, x_{21}, x_{22}) = u(x_{11}, x_{12}),$$
$$v(x_{11}, x_{12}, x_{21}, x_{22}) = v(x_{21}, x_{22}).$$

In other words, we would abstract from consumption externalities.

4.1.1 Pareto efficiency, individual rationality and the core

Pareto efficient allocations. The notion of Pareto efficiency is associated with the idea of improving the situations of the traders in the exchange economy. Consider an allocation x in the Edgeworth box. Can there be a different allocation x' that makes the traders better off? If this is the case for both traders, then allocation x is not a "good" allocation. Formally,

Definition 4.2: Pareto efficient allocation

An allocation x' Pareto dominates another allocation x if $u(x'_{i1}, x'_{i2}) \geq u(x_{i1}, x_{i2})$ for all $i = 1, 2$, and $u(x'_{i1}, x'_{i2}) > u(x_{i1}, x_{i2})$ for at least one i. An allocation x is a Pareto efficient allocation if there is no allocation x' that Pareto dominates x. The set of all Pareto efficient allocations is called the contract curve.

Definition 4.2 asserts that at an alternative allocation, there is no way to make both traders better off and make at least one strictly better off. How do we compute and represent a Pareto efficient allocation in the Edgeworth box? In Figure 4.3, allocations x and y are both Pareto efficient allocations. At these allocations, at which the indifference curves of traders 1 and 2 are tangent to each other, it is impossible to make one trader better strictly off without making the other strictly worse off. The curve that joins all such points is the contract curve for the exchange economy.

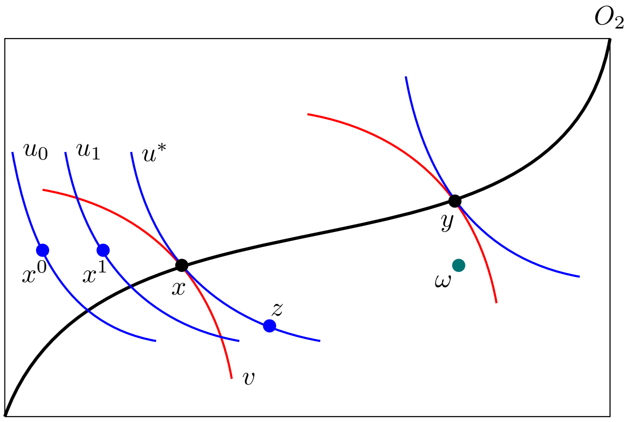


Figure 4.3: Pareto efficient allocations.

The tangency points can be derived in the following way.

- Fix a utility level of trader 2 at v, i.e., $v(x_{21}, x_{22}) = v$. Think of proposing allocations at all of which we are required to guarantee at least v to trader 2.

- Start from an allocation at which trader 1 obtains u_0. That is, start from allocation such as x^0 (on trader 1's indifference curve labeled u_0).

- Try to improve trader 1's utility. We must move to a higher indifference curve of this trader, say at level u_1, i.e., we must propose an allocation such as x^1.

- However, u_1 is not the best we can do. At an allocation like x^1, trader 2 obtains a utility level that is strictly higher than v, but there is still room for improvement for trader 1 (respecting the minimum utility constraint of trader 2).

- We can go as far as allocation x at which the two indifference curves are tangent to each other. Going further to x will now improve the utility of trader 1, but at the cost of diminishing utility of trader 2. So, x is a Pareto efficient allocation. Note also that a point like z is not Pareto efficient. It yields u^* to trader 1, but z is on the indifference curve of trader 2, which is strictly lower than that gives them v (so, the minimum utility constraint is violated).

So, to find a Pareto efficient allocation we would solve the following problem:

$$\begin{aligned} \max_{\{x_{11}, x_{12}, x_{21}, x_{22}\}} \quad & u(x_{11}, x_{12}), \\ \text{subject to } \quad & v(x_{21}, x_{22}) \geq v, \\ & x_{11} + x_{21} = \omega_1, \\ & x_{12} + x_{22} = \omega_2. \end{aligned}$$

The first constraint is the minimum utility constraint of trader 2, and the other two constraints are the feasibility constraints of goods 1 and 2. We would rather solve the following numerical example.

Example 4.1: Finding Pareto efficient allocations

Let the utility functions are given by $u(x_{11}, x_{12}) = x_{11}x_{12}$ and $v(x_{21}, x_{22}) = x_{21}x_{22}$, and the endowments of the two goods are given by $(\omega_{11}, \omega_{12}) = (2, 2)$ and $(\omega_{21}, \omega_{22}) = (2, 1)$. So, the aggregate endowments are $\omega_1 = 2 + 2 = 4$ and $\omega_2 = 2 + 1 = 3$ (i.e., the Edgeworth box has length 4 and width 3). We can write the above maximization problem as follows:

$$\begin{aligned} \max_{\{x_{11}, x_{12}\}} \; & u(x_{11}, x_{12}) = x_{11}x_{12}, \\ \text{subject to } & v(x_{21}, x_{22}) = x_{21}x_{22} \geq v, \\ & x_{11} + x_{21} = \omega_1 = 4, \\ & x_{12} + x_{22} = \omega_2 = 3. \end{aligned}$$

First, note that the first order condition of the above maximization problem is given by:

$$\underbrace{\frac{x_{12}}{x_{11}}}_{-\text{MRS}^1} = \underbrace{\frac{x_{22}}{x_{21}}}_{-\text{MRS}^2} = \alpha,$$

i.e., at any Pareto efficient allocation, the indifference curves of the two traders are tangent to each other. The above equations imply that

$$x_{i2} = \alpha x_{i1} \quad \text{for } i = 1, 2.$$

Substituting the above into the second feasibility constraint, we get

$$\alpha(\underbrace{x_{11} + x_{21}}_{=\omega_1}) = \omega_2 \quad \Longleftrightarrow \quad \alpha = \frac{\omega_2}{\omega_1} = \frac{3}{4}.$$

Therefore, the contract curve is given by:

$$x_{12} = \frac{\omega_2}{\omega_1} \cdot x_{11} = \frac{3}{4} \cdot x_{11}.$$

The contract curve is the diagonal of the Edgeworth box. It is easy to see from Figure 4.3 that the minimum utility constraint must hold with equality, i.e., $\alpha(x_{21})^2 = v$. Moreover, $x_{22} = \alpha x_{21}$. Let the Pareto efficient allocations to trader i be denoted by $(\hat{x}_{i1}, \hat{x}_{i2})$. The last two conditions alongwith $\alpha = \omega_2/\omega_1$ yield

$$(\hat{x}_{21}, \hat{x}_{22}) = \left(\sqrt{\frac{\omega_1 v}{\omega_2}}, \sqrt{\frac{\omega_2 v}{\omega_1}}\right) = \left(\sqrt{\frac{4v}{3}}, \sqrt{\frac{3v}{4}}\right).$$

Using the two feasibility constraints, we obtain

$$(\hat{x}_{11}, \hat{x}_{12}) = \left(\omega_1 - \sqrt{\frac{\omega_1 v}{\omega_2}}, \omega_2 - \sqrt{\frac{\omega_2 v}{\omega_1}}\right) = \left(4 - \sqrt{\frac{4v}{3}}, 3 - \sqrt{\frac{3v}{4}}\right).$$

Observations

- The Pareto efficient consumption x_{ij} does not depend on the an individual's endowment of each good, ω_{ij}. It only depends on the aggregate endowment of each good, ω_1 and ω_2 (i.e., the length and width of the Edgeworth box). Thus, if we change the length and width of the box, the Pareto efficient allocations (which lie on the diagonal of the box) change.
- A given Pareto efficient allocation depends on the minimum utility level v of trader 2, and hence, there are a continuum of such allocations, one for each v.

Exercise 4.1

Let the utility functions be given by $u(x_{11}, x_{12}) = \min\{x_{11}, x_{12}\}$ and $v(x_{21}, x_{22}) = \min\{x_{21}, x_{22}\}$. The endowments are $(\omega_{11}, \omega_{12}) = (2, 0)$ and $(\omega_{21}, \omega_{22}) = (0, 3)$. Find the set of Pareto efficient allocations.

Utility possibility frontier. We shall now define two concepts associated with Pareto efficiency. Consider the efficient allocations and compute the maximum value function of the maximization problem in Example 4.1:

$$
\begin{aligned}
u = \underbrace{\frac{\omega_2}{\omega_1}}_{\alpha} \cdot (x_{11})^2 &= \frac{\omega_2}{\omega_1} \cdot \left(\omega_1 - \sqrt{\frac{\omega_1 v}{\omega_2}}\right)^2 \\
\Longleftrightarrow \sqrt{\frac{\omega_1 u}{\omega_2}} + \sqrt{\frac{\omega_1 v}{\omega_2}} &= \omega_1 \\
\Longleftrightarrow \sqrt{u} + \sqrt{v} &= \sqrt{\omega_1 \cdot \omega_2} = \sqrt{12}. \qquad \text{(UPF)}
\end{aligned}
$$

Condition (UPF) describes what is called the utility possibility frontier (UPF) or the Pareto frontier or the bargaining frontier of the Edgeworth box economy. We can also write (UPF) as

$$u = \phi(v) \equiv (\sqrt{\omega_1 \cdot \omega_2} - \sqrt{v})^2. \qquad \text{(UPF}'\text{)}$$

In many contexts, the final utilities that are accrued to the traders, rather than their allocations, are more important and convenient to use. The above discussion is easily generalizable to a market with $n \geq 2$ traders and $m \geq 2$ goods.

Definition 4.3: Utility possibility set

Let $X = \{1, \ldots, m\}$ be the set of m goods that can be potentially traded in an exchange economy or market with the set of traders, $N = \{1, \ldots, n\}$. Let trader i's utility function over m goods be given by $u^i(x_i)$, where $x_i = (x_{i1}, \ldots, x_{im})$ denotes trader i's allocations of m goods, which is continuous and monotonic. The function

$$\phi(u_2, \ldots, u_n) = \max_x \left\{ u^1(x_1) \mid u^i(x_i) \geq u_i \text{ for } i \in N \setminus \{1\} \text{ and } \sum_{i=1}^{n} x_{ij} = \omega_j \text{ for all } j \in X \right\}$$

is called the utility possibility frontier of the exchange economy. The set

$$\mathcal{U} = \{(u_1, \ldots, u_n) \in \mathbb{R}^n_+ \mid u^1 \leq \phi(u_2, \ldots, u_n)\}$$

is called the utility possibility set of the exchange economy.

The Pareto frontier (UPF) and the associated utility possibility set are depicted in Figure 4.4. The utility possibility set is simply the combinations of payoffs that can be reached by the market participants by trading what they initially have (endowments) in a decentralized manner. So, we can see that there is a one-to-one correspondence between the set of (feasible) allocations and the utility possibility set. Also, there is a one-to-one correspondence between the contract curve and the Pareto frontier.

Let us derive the UPF for an exchange economy wherein there are $n \geq 2$ traders and 2 goods, 1 and 2.

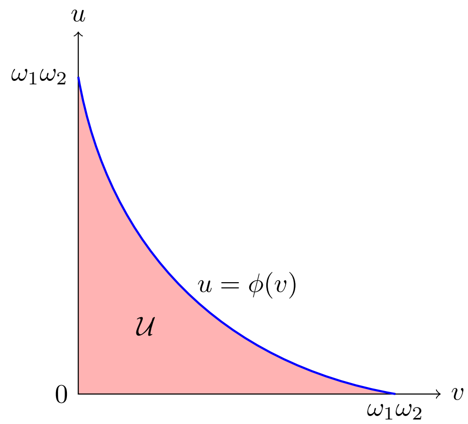


Figure 4.4: The Pareto frontier and the utility possibility set.

Example 4.2: UPF with many traders

Consider the set of traders $N = \{1, \ldots, n\}$ and two goods, 1 and 2. The utility function of trader i is given by $u^i(x_i)$ where $x_i = (x_{i1}, x_{i2})$ denotes trader i's allocations of 2 goods. Trader i's endowments are given by $(\omega_{i1}, \omega_{i2})$, and the aggregate endowment of good $j = 1, 2$ is given by $\sum_{i=1}^{n} \omega_{ij} = \omega_j$. A Pareto efficient allocation, $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_n)$ solves

$$\begin{aligned}
\max_{x} \quad & u^1(x_1) = x_{11}x_{12}, \\
\text{subject to} \quad & u^2(x_2) = x_{21}x_{22} \geq u_2, && \text{(U}_2\text{)} \\
& \vdots \\
& u^n(x_n) = x_{n1}x_{n2} \geq u_n, && \text{(U}_n\text{)} \\
& x_{11} + \ldots + x_{n1} = \omega_1, && \text{(F}_1\text{)} \\
& x_{12} + \ldots + x_{n2} = \omega_2. && \text{(F}_2\text{)}
\end{aligned}$$

At the Pareto efficient allocations, the marginal rate of substitution between the two goods (MRS) of all traders must be equal (i.e., the indifference curves of all traders are tangent to each other):

$$\text{MRS}^1 = \ldots = \text{MRS}^n \quad \Longleftrightarrow \quad \frac{x_{12}}{x_{11}} = \ldots = \frac{x_{n2}}{x_{n1}} = \alpha\,(\text{say}).$$

Therefore, we have $x_{i2} = \alpha x_{i1}$ for all $i \in N$. Substituting this into the second feasibility constraint (F$_2$), we obtain

$$\alpha(\underbrace{x_{11} + \ldots + x_{n1}}_{=\omega_1 \text{ from (F}_1\text{)}}) = \alpha \cdot \omega_1 = \omega_2 \quad \Longleftrightarrow \quad \alpha = \frac{\omega_2}{\omega_1}.$$

Constraints (U$_2$)-(U$_n$) would bind at the optimum, and hence, we have

$$\begin{aligned}
& \alpha(x_{21})^2 = \frac{\omega_2}{\omega_1} \cdot (x_{21})^2 = u_2, \ldots, \alpha(x_{n1})^2 = \frac{\omega_2}{\omega_1} \cdot (x_{n1})^2 = u_n \\
\Longleftrightarrow \quad & \hat{x}_{21} = \sqrt{\frac{\omega_1 u_2}{\omega_2}}, \ldots, \hat{x}_{n1} = \sqrt{\frac{\omega_1 u_n}{\omega_2}}.
\end{aligned}$$

Let u_1 denote the maximized utility of trader 1, which is given by:

$$u_1 = \alpha(\hat{x}_{11})^2 = \frac{\omega_2}{\omega_1} \cdot (\hat{x}_{11})^2 = \frac{\omega_2}{\omega_1} \cdot (\omega_1 - \hat{x}_{21} - \ldots - \hat{x}_{n1})^2$$

$$\iff \frac{\omega_1 u_1}{\omega_2} = \left(\omega_1 - \sqrt{\frac{\omega_1 u_2}{\omega_2}} - \ldots - \sqrt{\frac{\omega_1 u_n}{\omega_2}}\right)^2$$

$$\iff \sqrt{\frac{\omega_1 u_1}{\omega_2}} + \sqrt{\frac{\omega_1 u_2}{\omega_2}} + \ldots + \sqrt{\frac{\omega_1 u_n}{\omega_2}} = \omega_1$$

$$\iff \sqrt{u_1} + \sqrt{u_2} + \ldots + \sqrt{u_n} = \sqrt{\omega_1 \cdot \omega_2} \quad \iff \quad u_1 = \phi(u_2, \ldots, u_n) \equiv \left(\sqrt{\omega_1 \cdot \omega_2} - \sum_{i=2}^{n} \sqrt{u_i}\right)^2.$$

The above is our desired UPF.

Individually rational allocations. We have seen so far that a Pareto efficient allocation does not depend on individual's endowment of each good. The next property we would like to impose on an allocation is individual rationality. We would require that each trader must do at least as well as what they could with their endowments. Formally,

Definition 4.4: Individually rational allocation

An allocation x in the Edgeworth box is individually rational if $u(x_{11}, x_{12}) \geq u(\omega_{11}, \omega_{12})$ and $v(x_{21}, x_{22}) \geq v(\omega_{21}, \omega_{22})$.

The above definition asserts that, at an individually rational allocation, each trader obtains utility that is at least as large as their utility evaluated at their endowments of the two goods.

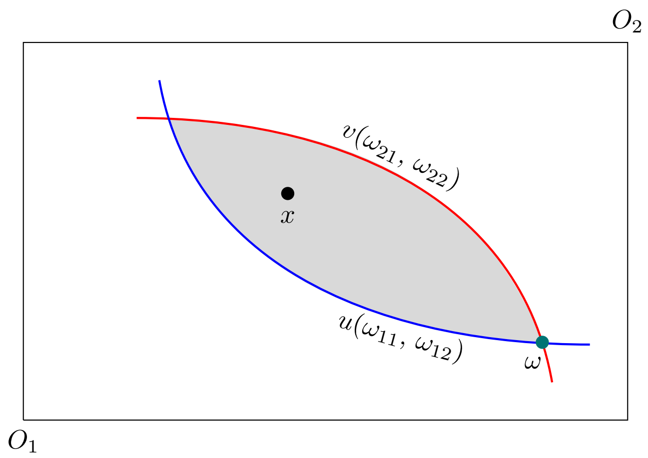


Figure 4.5: Individually rational allocations.

In Figure 4.5, both traders obtain higher utility at a point like x that that evaluated at ω because at x lies at a higher indifference curve of each trader. Therefore, the set of individually rational allocations relative to the endowment point ω is lens-shaped shaded region.

Core allocations. The next property we would like to impose on an allocation is that it is in the core of the exchange economy. This idea will be generalized in Chapter **??**. The notion of core is related to the idea that some allocations can be blocked or objected by the traders. In our 2×2 exchange economy, there are two ways to block a proposed allocation—(a) each trader on their own can object to a proposed allocation, and (b) the two traders together can block a proposed allocation. In a more general model wherein there are $n > 2$ traders, any non-empty subsets of traders, called coalitions, are allowed to block an allocation (there are $2^n - 1$ such coalitions). Formally,

Definition 4.5: Core allocations

An allocation x in the Edgeworth box is blocked by the trader(s) if there is another (feasible) allocation x' such $u(x'_{11}, x'_{12}) \geq u(x_{11}, x_{12})$ and $v(x'_{21}, x'_{22}) \geq u(x_{21}, x_{22})$ with strict inequality for at least one trader. An allocation x is a core allocation if there is no other allocation x' with which x can be blocked.

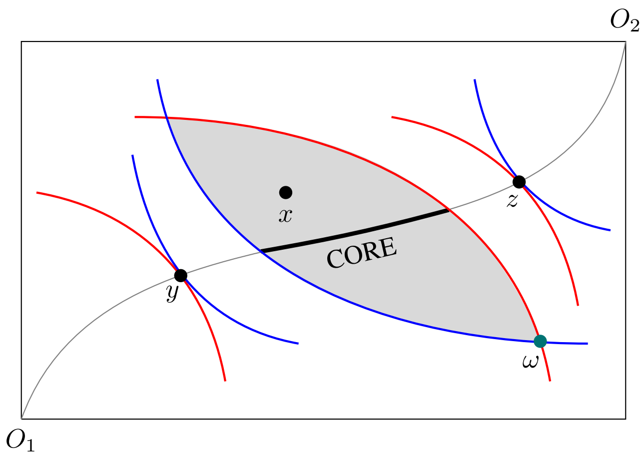


Figure 4.6: The core.

What is crucial is that the traders can object to an allocation by proposing another allocation that is feasible for them. Let us analyze the various possibilities of blocking.

- Consider the case when a single trader can block a proposed allocation. Note that the only feasible allocation that a trader has and can be used to block an allocation their endowments of the two goods, i.e., $(\omega_{i1}, \omega_{i2})$ for each $i = 1, 2$. Therefore, allocations like y and z (in Figure 4.6) cannot be core allocation either (y will blocked by trader 1 and z, by trader 2). Consequently, a core allocation must lie within the lens-shaped shaded region in Figure 4.6.

- Next, consider the possibility of blocking by the two traders together. Proposing an alternative allocation to block an initial one means that the two traders can exchange their endowments to create a "new allocation" that makes both of them better off. In Figure 4.6, an allocation like x will be blocked by the traders together. To see this, create another lens-shaped region with x and ω. Any allocation in the newly-created region yields strictly higher utility of both traders. If we keep on doing this, the only allocations that would not be blocked are those on the contract curve (when no further lens-shaped region of improvement can be created).

From the above discussion we can conclude

Theorem 4.1

An allocation in the Edgeworth box is Pareto efficient and individually rational if and only if it is a core allocation.

The proof is easy for the Edgeworth box economy. In Figure 4.6, the part of contract curve that lies inside the lens-shaped region is the core of our exchange economy. For an exchange economy with more than 2 traders, it is always true that a core allocation is Pareto efficient and individually rational, but the converse is not true in general.

Exercise 4.2: Core is a stronger concept than Pareto efficiency and individual rationality

Consider an exchange economy with three traders, 1, 2 and 3, and two goods, 1 and 2. The utility functions are given by $u^i(x_{i1}, x_{i2}) = x_{i1}x_{i2}$ for $i = 1, 2, 3$. The endowments are given by $(\omega_{11}, \omega_{12}) = (1, 9)$, $(\omega_{21}, \omega_{22}) = (5, 5)$ and $(\omega_{31}, \omega_{32}) = (9, 1)$. Show that there is at least one allocation that is Pareto efficient and individually rational, but it is not a core allocation. Find the core allocations of this market.

The reason for which Theorem 4.1 does not hold for an exchange economy with more than 2 traders is that the possibility of blocking a proposed allocation increases as the number of traders grows. If there are 5 traders, then there are $2^5 - 1 =$31 coalitions that can potentially block an allocation. However, a market with 20 traders, this number becomes $2^{20} - 1 =$1,048,575. One may also wonder what happens if we replicate our Edgeworth box economy.

Exercise 4.3: Core shrinks under replication

Consider an exchange economy with two traders, 1 and 2, and two goods, 1 and 2. The utility functions are given by $u^i(x_{i1}, x_{i2}) = x_{i1}x_{i2}$ for $i = 1, 2$. The endowments are given by $(\omega_{11}, \omega_{12}) = (9, 1)$ and $(\omega_{21}, \omega_{22}) = (1, 9)$. Show that Theorem 4.1 holds for this market.
Next, consider a replication of the Edgeworth box in the following way. Add two more traders, 3 and 4. Trader 3 is the identical twin of trader 1, and trader 4 is the identical twin of trader 2. This is to say that $u^i(x_{i1}, x_{i2}) = x_{i1}x_{i2}$ for $i = 1, 2, 3, 4$, and $(\omega_{31}, \omega_{32}) = (9, 1)$ and $(\omega_{41}, \omega_{42}) = (1, 9)$. Show that Theorem 4.1 does not hold for the replicated economy.

Edgeworth conjectured in 1881 that the core is large in a small market, whereas it is small in a large market. Aumann (1964) proves a striking result that the core of an exchange economy shrinks as the market is replicated. Moreover, he shows that if the market is replicated infinitely, the core converges to a single allocation. This motivates our analysis of a competitive or Walrasian equilibrium.

4.1.2 The Walrasian equilibrium

We now introduce a Walrasian allocation for the Edgeworth box economy. For that we require a price system which the traders would take as given while trading with each other. Let $(p_1, p_2) \in \mathbb{R}_{++}$ be the price vector. The budget set of trader $i = 1, 2$ is given by:

$$B_i(p_1, p_2) = \{(x_{i1}, x_{i2}) \in \mathbb{R}_+ \mid p_1x_{i1} + p_2x_{i2} \leq p_1\omega_{i1} + p_2\omega_{i2}\}.$$

Definition 4.6: Walrasian equilibrium

A Walrasian (or competitive) equilibrium for the Edgeworth box economy is a price vector $p^* = (p_1^*, p_2^*)$ and an allocation $x^* = ((x_{11}^*, x_{12}^*), (x_{21}^*, x_{22}^*))$ in the Edgeworth box such that for each $i = 1, 2$, (x_{i1}^*, x_{i2}^*) solves

$$\max_{\{x_{i1}, x_{i2}\}} u^i(x_{i1}, x_{i2}),$$
$$\text{subject to } p_1x_{i1} + p_2x_{i2} \leq p_1\omega_{i1} + p_2\omega_{i2}.$$

In a Walrasian equilibrium, given the commodity prices, each trader maximizes their utility subject to the budget constraint. Because x^* is an allocation, we must by definition have supply equal to demand for every good, i.e.,

$$x_{1j} + x_{2j} = \omega_{1j} + \omega_{2j} \quad \text{for each good } j = 1, 2.$$

Also, notice that a Walrasian equilibrium depends on the endowments, ω. If we change ω, we change the competitive equilibrium. The maximization problem in Definition 4.6 clearly yields

$$MRS^1_{12} = -\frac{p_1}{p_2} = MRS^2_{12}.$$

In Example 4.3, we shall compute Walrasian equilibrium allocations under specific functional forms of the utility functions. Figure 4.7 depicts a competitive equilibrium where the line joining ω and x^* represents the relative prices with slope $-p_1^*/p_2^*$. As we can clearly see that x^* is a core allocation. We state the following

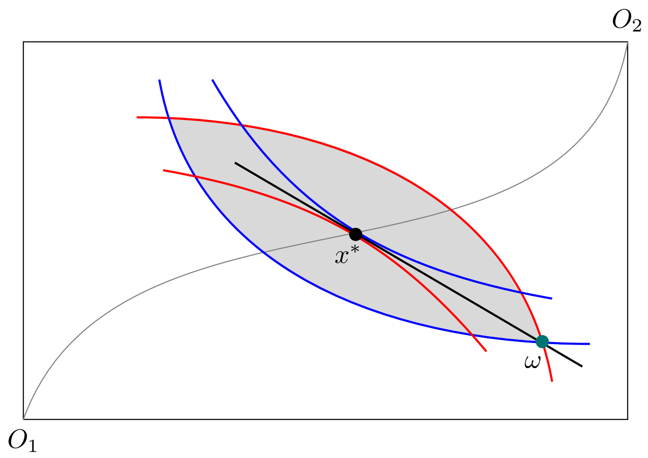


Figure 4.7: A competitive equilibrium.

result (without proof) that describes the welfare properties of a Walrasian allocation.

Theorem 4.2: First theorem of welfare economics

Assume that all traders have monotone utility functions, i.e., $u^i(x_i) > u^i(y_i)$ for $x_i > y_i$ for all $i \in N$. Let (x^*, p) be a Walrasian equilibrium. Then, x^* is a core allocation (and is, therefore, Pareto efficient as well).

Exercise 4.4: Proof of Theorem 4.2

Prove Theorem 4.2. Is the converse true? Give a graphical argument.

4.2 Mechanism design

4.2.1 Implementation of Walrasian allocations

Let us start with the following example of an Edgeworth box economy in order to understand the kind of problem we would analyze in this section.

Example 4.3: Manipulability of Walrasian allocation

Consider the following Edgeworth box economy wherein the endowments are given by $(\omega_{11}, \omega_{12}) = (3, 9)$ and $(\omega_{21}, \omega_{22}) = (9, 3)$. Trader 1's preferences are given by $u(x_{11}, x_{12}) = x_{11}x_{12}$. However, trader 2 has one of the two alternative preferences, $v(x_{21}, x_{22}) = x_{21}+x_{22}$ and $v'(x_{21}, x_{22}) = \sqrt{2}x_{21}+\frac{1}{\sqrt{2}}x_{22}$. The government, who wants to implement the Walrasian allocation (which is Pareto efficient) at prices $(p, 1)$, does not know the true preferences of trader 2. Let us first compute the Walrasian equilibrium under both circumstances.

First, consider the preference profile given by $u(x_{11}, x_{12})$ and $v(x_{21}, x_{22})$. Let (x^*, p^*) be the Walrasian equilibrium which is given by the following set of equations:

$$
\begin{aligned}
&\mathrm{MRS}^1_{12} = -\frac{x_{12}}{x_{11}} = -p,\\
&px_{11} + x_{12} = 3p + 9,\\
&\mathrm{MRS}^2_{12} = -1 = -p,\\
&px_{21} + x_{22} = 9p + 3,\\
&x_{11} + x_{21} = 12,\\
&x_{12} + x_{22} = 12.
\end{aligned}
$$

The above system yields $(x^*_{11}, x^*_{12}) = (6, 6)$, $(x^*_{21}, x^*_{22}) = (6, 6)$ and $p^* = 1$.

Next, consider the preference profile given by $u(x_{11}, x_{12})$ and $v'(x_{21}, x_{22})$. Let (x', p') be the Walrasian equilibrium which is given by the following set of equations:

$$
\begin{aligned}
&\mathrm{MRS}^1_{12} = -\frac{x'_{12}}{x'_{11}} = -p',\\
&px'_{11} + x'_{12} = 3p' + 9,\\
&\mathrm{MRS}^2_{12} = -2 = -p',\\
&px'_{21} + x'_{22} = 9p' + 3,\\
&x'_{11} + x'_{21} = 12,\\
&x'_{12} + x'_{22} = 12.
\end{aligned}
$$

The above system yields $(x'_{11}, x'_{12}) = (3.75, 7.5)$, $(x'_{21}, x'_{22}) = (8.25, 4.5)$ and $p' = 2$.

Now suppose that trader 2's true preferences is given by $v(x_{21}, x_{22})$ which is private information. Suppose further that the government asks the traders to report their preferences in an effort to implement the Walrasian allocation $(6, 6), (6, 6))$. Note that if trader 2 falsifies their preferences by reporting $v'(\cdot, \cdot)$, then the government would implement $x' = ((3.75, 7.5), (8.25, 4.5))$. In fact, this trader has incentives to do so because $v(8.25, 4.5) = 12.75 > 12 = v(6, 6)$.

The notion of mis-reporting must be understood carefully. When trader 2 misreport, they do not compare $v'(8.25, 4.5)$ with $v(6, 6)$, but compares $v(8.25, 4.5)$ with $v(6, 6)$. The idea is that a trader mis-report their preferences in an effort to obtain an allocation which is different from the one that would have been implemented if they would have reported truthfully. Then, the trader evaluates their true utility function at the allocation obtained by mis-reporting. If this utility is strictly higher than that at the allocation that would have been implemented under truthful reporting, we say that the trader has incentive to falsify their true preferences.

In Example 4.3, the reason behind the failure to induce trader 1 to report truthfully their preferences is described in the left panel of Figure 4.8. The bundles in the red-shaded region in the left panel were less

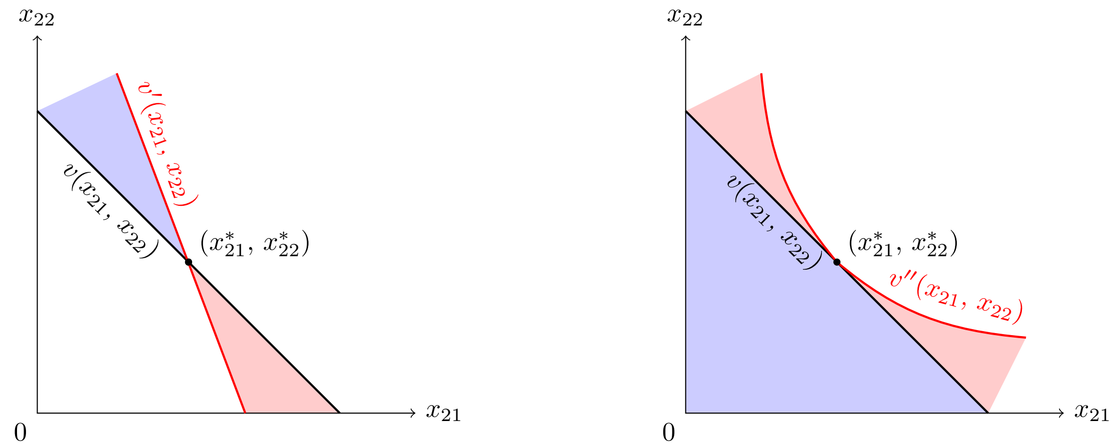


Figure 4.8: Sufficient condition for truthful implementability.

preferred to the Walrasian allocation of trader 2, (x_{21}^*, x_{22}^*) under preferences $v(\cdot, \cdot)$. However, when their preferences have changed to $v'(\cdot, \cdot)$, those bundles became more preferred to (x_{21}^*, x_{22}^*), which induces trader 2 to mis-report their preferences. So, for truthful implementability of the competitive allocation, we require a sort of preference reversal, i.e., for trader 2, (x_{21}^*, x_{22}^*) must be weakly preferred to (x'_{21}, x'_{22}) under the utility function $v(\cdot, \cdot)$; but (x'_{21}, x'_{22}) must be weakly preferred to (x_{21}^*, x_{22}^*) under the utility function $v'(\cdot, \cdot)$. This is to say that (x'_{21}, x'_{22}) must lie in the blue-shaded region in the left panel of Figure 4.8

Suppose that the alternative preferences of trader 2 is given by $v''(x_{21}, x_{22}) = x_{21}x_{22}$ instead of $v'(x_{21}, x_{22})$. In this case we have $p'' = p^*$ and $x'' = x^*$. Under preferences $v''(\cdot, \cdot)$, the set of bundles that are less preferred to (x_{21}^*, x_{22}^*), i.e., the lower contour set of (x_{21}^*, x_{22}^*), expands (the red-shaded region in the right panel of Figure 4.8). In other words, if the bundles that were inferior to (x_{21}^*, x_{22}^*) under preferences $v(\cdot, \cdot)$ are still inferior for trader 2 under the changed preferences $v''(\cdot, \cdot)$, then the same allocation x^* must be chosen under the new preferences. This monotone way of changing preferences is called the Maskin-monotonicity which guarantees truthful implementation of the competitive allocation in the Edgeworth box.

4.2.2 Allocation of objects among several buyers

Consider an economy with one seller (agent S) of single object, and two potential buyers (agents 1 and 2). An allocation of the economy is to allocate the object to a single buyer (think of the sale of a single unit of an indivisible private good). An allocation of the economy is denoted by a. The set of possible allocations is given by:

$$A = \left\{ (x_S, x_1, x_2, t_S, t_1, t_2) \in \{0, 1\}^3 \times \mathbb{R}^3 \mid \sum_i x_i = 1 \text{ and } \sum_i t_i \leq 0 \right\}.$$

In words, in an allocation, $x_i = 1$ if agent $i = S, 1, 2$ gets the object, and $x_i = 0$ if they do not get the object. We allow for monetary transfers in that t_i denotes the transfer received by agent $i = S, 1, 2$. Note that some t_i can be negative. The feasibility restrictions we impose that the aggregate allocation is 1 (as there is a single unit to allocate), and the aggregate transfer must be non-positive. The agents will be characterized by their valuation for the good. In particular, the seller has a valuation θ_S and the buyers have valuations θ_1 and θ_2. The utility function of an agent with valuation or "type" θ_i given by:

$$u_i(a, \theta_i) = \theta_i x_i + t_i.$$

We shall assume that the seller's valuation θ_S is common knowledge, and thus normalize it to 0; however, θ_1 and θ_2 are private information. We assume that both θ_1 and θ_2 are drawn from uniform distribution on $[0, 1]$

Definition 4.7: Ex-post efficiency

An allocation $a \in A$ is ex-post efficient if it allocates the object to the highest-valuation buyer and if it involves no waste of money, i.e., for all $\theta = (\theta_1, \theta_2)$,

$$x_i(\theta)(\theta_i - \max\{\theta_1, \theta_2\}) = 0 \quad \text{for all } i, \quad \text{and} \quad \sum_i t_i = 0.$$

The first part of the above definition reads as (a) if $\max\{\theta_1, \theta_2\} = \theta_1$, then $(x_1, x_2) = (1, 0)$, and hence, $x_1(\theta_1 - \theta_1) = 0$ and $x_2(\theta_2 - \theta_1) = 0$, and (b) if $\max\{\theta_1, \theta_2\} = \theta_2$, then $(x_1, x_2) = (0, 1)$, and hence, $x_1(\theta_1 - \theta_2) = 0$ and $x_2(\theta_2 - \theta_2) = 0$.

Example 4.4: Winner pays the highest valuation

Suppose we want to implement the following allocation $a(\theta)$: For $i, j = 1, 2$, and $i \neq j$,

$$x_i(\theta) = \begin{cases} 1 & \text{if } \theta_i \geq \theta_j, \\ 0 & \text{if } \theta_i < \theta_j. \end{cases} \quad \text{and} \quad t_i(\theta) = -\theta_i x_i(\theta);$$

$$x_S(\theta) = 0, \quad \text{and} \quad t_S(\theta) = -(t_1(\theta) + t_2(\theta)).$$

The above allocation rule allocates the object to the highest-valuation buyer, and the seller receives a transfer equal to the highest valuation. The allocation rule is not only ex-post efficient, but also it is very attractive for the seller in that the seller extracts the entire consumer surplus generated by the trade if the allocation is implemented.
Suppose the buyers are expected utility maximizers. The question is: if buyer 2 always announces their true valuation, will it be optimal for buyer 1 to do the same? For each θ_1, buyer 1's problem is to announce a type $\tilde{\theta}_1$ so as to solve

$$\max_{\tilde{\theta}_1} \mathbb{E}[\theta_1 x_1(\tilde{\theta}_1, \theta_2) + t_1(\tilde{\theta}_1, \theta_2)] = \text{Prob.}(\theta_2 \leq \tilde{\theta}_1)[\theta_1 \cdot 1 - \tilde{\theta}_1 \cdot 1] + \text{Prob.}(\theta_2 > \tilde{\theta}_1)[\theta_1 \cdot 0 - 0] = (\theta_1 - \tilde{\theta}_1)\tilde{\theta}_1.$$

The above maximization yields the optimum announcement $\tilde{\theta}_1 = \theta_1/2$. Likewise, for buyer 2 we have $\tilde{\theta}_2 = \theta_2/2$. The buyers have incentives to under-report their valuation in order to lower the transfer they must make to the seller in case they are the winner. Of course, this increases the probability of not obtaining the object. However, each buyer would exploit this trade-off, at least to some extent.

So, does there exist some other allocation rule that can be truthfully implemented? The answer is yes.

Example 4.5: Winner pays the second-highest valuation

Suppose we want to implement the following allocation $\hat{a}(\theta)$: For $i, j = 1, 2$, and $i \neq j$,

$$x_i(\theta) = \begin{cases} 1 & \text{if } \theta_i \geq \theta_j, \\ 0 & \text{if } \theta_i < \theta_j. \end{cases} \quad \text{and} \quad t_i(\theta) = -\theta_j x_i(\theta);$$

$$x_S(\theta) = 0, \quad \text{and} \quad t_S(\theta) = -(t_1(\theta) + t_2(\theta)).$$

The above rule implies that if buyer i wins the object (i.e., i is the highest-valuation buyer), they pay the lower valuation.

Now, let us analyze the incentive of buyer 1 to reveal truthfully their valuation when buyer 2 announces $\tilde{\theta}_2$ (the case of buyer 2 is symmetric). First, consider the case when $\tilde{\theta}_2 \leq \theta_1$. By announcing

$\tilde{\theta}_1 = \theta_1$ (the true valuation), buyer 1 obtains the object, and their utility is given by $\theta_1 - \tilde{\theta}_2 \geq 0$. If they announce $\tilde{\theta}_1 \neq \theta_1$, they obtain the object as long as $\tilde{\theta}_1 \geq \tilde{\theta}_2$. In this case, buyer 1's utility is $\theta_1 - \tilde{\theta}_2$ which is the same as that under truthful revelation. On the other hand, if $\tilde{\theta}_1 < \tilde{\theta}_2$, buyer 1 does not obtain the object, and consumes 0 utility. So, for $\tilde{\theta}_2 \leq \theta_1$, buyer 1 does not have incentive to falsify their valuation. Next, suppose that $\tilde{\theta}_2 > \theta_1$. In this case, buyer 1 obtains 0 (buyer 2 gets the object) whether or not they report truthfully as long as $\tilde{\theta}_1 < \tilde{\theta}_2$. By contrast, if buyer 1 announces $\tilde{\theta}_1 > \tilde{\theta}_2$, they obtain the object; however, buyer 1 gets a utility equal to $\theta_1 - \tilde{\theta}_2 < 0$. Therefore, in this case too, buyer 1 does not have incentive to falsify their type. So, the optimal announcement of buyer 1 is $\tilde{\theta}_1 = \theta_1$ regardless of what the other buyer announces. Formally, telling truth is a weakly dominant strategy for both buyers. Thus, the allocation rule can be implemented even if buyers' valuations are private information—it suffices to ask each buyer to announce their type, and then choose $\hat{a}(\theta)$.

A mechanism wherein each buyer is asked to report their type is a direct mechanism. What can we say about the implementability of an allocation rule when agents are asked to announce a function of their type (an indirect mechanism)?

Example 4.6: First-price sealed-bid auction

In a first-price sealed-bid auction, each buyer (bidder) i submits a sealed-bid $b_i \geq 0$. The bids are then opened, and the buyer with the highest bid gets the object and pays their bid to the seller (auctioneer). Because buyer valuations are private information, our solution concept would be the Bayesian Nash equilibrium (BNE). In a Bayesian game, the strategy (bid) of each buyer i is a function of their type θ_i. For the interest of simplicity, suppose buyers follow a linear strategy, $b_i(\theta_i) = \beta_i \theta_i$ with $\beta_i \in [0, 1]$. Consider bidder 1's problem who solves

$$\max_{0 \leq b_1 \leq \beta_2} (\theta_1 - b_1) \cdot \text{Prob.}(b_2(\theta_2) \leq b_1) = (\theta_1 - b_1) \cdot (b_1/\beta_2).$$

The upper limit β_2 on b_1 is because β_2 is buyer 2's maximum bid (when $\theta_2 = 1$), and hence, buyer 1 should never bid more than β_2. Buyer 2 solves a similar problem. So, the optimal bidding functions are given by:

$$b_1(\theta_1) = \min\left\{\frac{1}{2}\theta_1, \beta_2\right\} \quad \text{and} \quad b_2(\theta_2) = \min\left\{\frac{1}{2}\theta_2, \beta_1\right\}.$$

If $\beta_1 = \beta_2 = \frac{1}{2}$, we have $\min\{\theta_i/2, \beta_j\} = \theta_i/2$. In this case, the first-price auction yields the same outcome as Example 4.4.

Example 4.7: Second-price sealed-bid auction

In a second-price sealed-bid auction, each buyer (bidder) i submits a sealed-bid $b_i \geq 0$. The bids are then opened, and the buyer with the highest bid gets the object, but pays the second-highest bid to the seller (auctioneer). Let us solve the equilibrium for more than 2 buyers as the argument is very similar to Example 4.5. Let $N = \{1, \ldots, n\}$ be the set of $n \geq 2$ potential buyers of a single indivisible object. Buyer i's payoff is given by:

$$u_i(\theta_i) = \begin{cases} \theta_i - \max_{j \neq i} b_j & \text{if } b_i > \max_{j \neq i} b_j, \\ 0 & \text{if } b_i < \max_{j \neq i} b_j. \end{cases}$$

We also assume that if there is a tie, i.e., $b_i = \max_{j \neq i} b_j$, the object goes to each winning bidder with equal probability. We now show that $b_i = \theta_i$ for all $i \in N$ is a dominant strategy equilibrium. Consider, say buyer 1, and let $\hat{b}_1 \equiv \max_{j \neq 1} b_j$ be the highest competing bid.

By bidding θ_1, buyer 1 will win if $\theta_1 > \hat{b}_1$, and not if $\theta_1 < \hat{b}_1$ (if $\theta_1 = \hat{b}_1$, buyer 1 is indifferent between winning and losing). Suppose, however, that buyer 1 bids $b_1 < \theta_1$. If $\theta_1 > b_1 \geq \hat{b}_1$, buyer 1 still wins and their utility is still $\theta_1 - \hat{b}_1$. If $\hat{b}_1 > \theta_1 > b_1$, buyer 1 still loses. However, if $\theta_1 > \hat{b}_1 > b_1$, then buyer 1 loses whereas if they had bid $b_1 = \theta_1$, they would have consumed a positive utility. Thus, bidding less than θ_1 can never increase buyer 1's utility but in some circumstances may actually decrease it. A similar argument shows that it is not profitable to bid more than θ_1.

It is not difficult to see that, when there are two buyers, i.e., $n = 2$, the second-price sealed-bid auction yields the same outcome as Example 4.5.

4.3 Direct mechanism and the revelation principle

In the previous section we have seen that the principal, to maximize her expected utility, offers a direct incentive compatible mechanism $g(\theta) = (q(\theta),\, t(\theta))$ to the agent without knowing his true type, and the agent truthfully reveals his type. A direct mechanism, which we will define formally below, is a mechanism in which the agent is asked to report his type directly to the principal on the basis of which the principal proposes an allocation. A direct mechanism is the simplest form of mechanisms that can be offered to the agent. In general a mechanism may be much more complex than just asking the agent to announce his type.

Definition 4.8: Mechanism

A mechanism is a game form $\Gamma = (\mathcal{M},\, \tilde{g})$ that consists of a message or strategy space $\mathcal{M}$ of the agent, and a mapping $\tilde{g} : \mathcal{M} \to \mathcal{A}$, writes $\tilde{g}(m) = (\tilde{q}(m),\, \tilde{t}(m))$ for all $m \in \mathcal{M}$, where

$$\mathcal{A} = \{(q,\, t) \mid q \in \mathbb{R}_+ \text{ and } t \in \mathbb{R}\}$$

is the set of possible allocations.

When facing such a mechanism, the agent with type θ chooses a best response $m^*(\theta)$ which is implicitly defined as

$$V(\tilde{q}(m^*(\theta)),\, \theta) - \tilde{t}(m^*(\theta)) \geq V(\tilde{q}(m),\, \theta) - \tilde{t}(m) \quad \text{for all} \quad m \in \mathcal{M}. \tag{4.1}$$

The mechanism $(\mathcal{M},\, \tilde{g}(\cdot))$ induces therefore an allocation rule $a(\theta) = (\tilde{q}(m^*(\theta)),\, \tilde{t}(m^*(\theta)))$ mapping the set of types Θ into the set of allocations $\mathcal{A}$.

Definition 4.9: Direct mechanism

A direct revelation mechanism is a mapping $g : \Theta \to \mathcal{A}$ which writes as $g(\theta) = (q(\theta),\, t(\theta))$ for all $\theta \in \Theta$. The principal commits to offer the transfer $t(\theta')$ and the production level $q(\theta')$ if the agent announces the value θ' for any $\theta' \in \Theta$. Moreover, a direct mechanism $g(\cdot)$ is 'truthful' if it is incentive compatible for the agent to announce his true type, for any type, i.e.,

$$V(q(\theta),\, \theta) - t(\theta) \geq V(q(\theta'),\, \theta) - t(\theta') \quad \text{for all} \quad \theta,\, \theta' \in \Theta. \tag{4.2}$$

Notice that the message space $\mathcal{M}$ of the agent may be very complex which may consist of a whole gamut of messages including report of own types. First, one may wonder if a better outcome could be achieved with a more complex contract allowing the agent to possibly choose among more options. Second, one may also wonder whether some sort of communication device between the agent and the principal could be used to transmit information to the principal so that the latter can recommend outputs and payments as a function of transmitted information. This is not the case. Indeed, the *Revelation Principle* ensures that there is no loss of generality in restricting the principal to offer simple menus having at most as many options as the cardinality of the type space Θ.

Proposition 4.1: The revelation principle

Any allocation rule $a(\theta)$ that can be implemented by a mechanism $(\mathcal{M}, \tilde{g}(\cdot))$ can also be implemented by a direct incentive compatible mechanism $g(\theta)$.

Proof. The indirect mechanism $(\mathcal{M}, \tilde{g}(\cdot))$ induces an allocation rule $a(\theta) = (\tilde{q}(m^*(\theta)), \tilde{t}(m^*(\theta)))$. Now construct a direct mechanism $g(\theta)$ such that $g := \tilde{g} \circ m^* : \Theta \to \mathcal{A}$. To prove the proposition we have to check that the direct mechanism that we have constructed is incentive compatible. Since condition (4.1) holds for all $m \in \mathcal{M}$, it holds in particular for $m^*(\theta')$ for all $\theta' \in \Theta$, i.e.,

$$V(\tilde{q}(m^*(\theta)), \theta) - \tilde{t}(m^*(\theta)) \geq V(\tilde{q}(m^*(\theta')), \theta) - \tilde{t}(m^*(\theta'))\theta \quad \text{for all} \quad \theta, \theta' \in \Theta. \tag{4.3}$$

Then, using the definition of g, we get the incentive compatibility condition (4.2) for the direct mechanism $g(\theta)$, which completes the proof. $\square$

The above result is special case of the general result proven by Myerson (1981) in the context of a mechanism design problem with many agents. Analysis of mechanism design with many agents is beyond the scope of this course.

Chapter 5

Auctions

Markets or institutions comprise a set of rules in order to allocate "individuals" or "objects" among "individuals" or "organizations". In a market (e.g., the Edgeworth box), the equilibrium price of an object depend on the willingness-to-pay of the buyers, which are public knowledge. We now relax the assumption of public information regarding willingness-to-pay. In such environments, prices will still depend on willingness-to-pay; however, we have to elicit willingness-to-pay in order to determine prices. Thus, Auctions are institutions that are not only meant to set allocation rules, but also they comprise payment rules, i.e., how much market participants must pay to obtain a single object or several objects. Moreover, auctions are indirect mechanisms that are used to reveal private information of the market participants.

We shall consider a simple problem of market design wherein a single object to be allocated to one of many potential buyers whose valuations for the object are private information. This is a special case of a market where more than one objects are allocated among the individuals. However, the objective of a mechanism would be to elicit the true willingness-to-pay of the market participants as we have seen Chapter 4. Allocating multiple objects or multiple units of a single objets are also of great interest; however, such models are considerably more complicated.

Auctions are designed to accomplish several goals. First, we would like to understand the behavior of bidders. For this, we will adopt an appropriate notion of equilibrium and analyze equilibrium behavior of bidders. Second, we would like to compare auction formats in terms of their equilibrium outcomes. When comparing auction formats, we usually take two criteria into account—(a) efficiency and (b) expected revenue to the seller. The notion of efficiency is the standard one of ex-post Pareto efficiency. In the context of single-object auction, the concept of efficiency is simple—an auction is efficient if the object is allocated to the highest-valuation bidder. Expected revenue, on the other hand, reflects an ex-ante objective of the seller, which is to maximize expected revenue across auction formats. Under reasonable conditions, we are able to rank standard auctions in terms of efficiency and expected revenue.

5.1 Auction formats

There are two main popular auction formats—open-bid auction and sealed-bid auction. In an open-bid auction, the bids are announced publicly, and hence, all bids are observable. In a sealed-bid auction, on the other hand, potential buyers of an object submit their bids in sealed envelopes, and thus, bids are not observed by others. Examples of open-bid format includes

- Ascending price auction or English auction. In this format, the auctioneer announces the price starting from a given price (called the reserve price) which increases as the auction advances. The potential buyers announces whether they want the object at the last announced price. The auction stops when only one bidder is left who wins the object and pays their bid. Artworks, antiques, etc. are sold in this format.

- Descending price auction or Dutch auction. The price starts at a high level, and starts to drop as the auction goes on. The potential buyers call out "mine". The first individual to call out gets the object, and pays their bid. Fresh food items, flowers, etc. are sold in this format.

On the other hand, examples of sealed-bid auction are

- First-price auction. In this format, all bidders simultaneously submit bids in sealed envelopes. The highest bid wins and the winner pays their bid.
- Second-price auction or Vickrey auction. Similar to the first-price auction. However, the winner pays the highest losing bid or the second-highest bid.
- Third-price auction. Winner pays the third-highest bid.
- All-pay auction. Everybody pays their bid irrespective of winning or losing.

Another categorization of auctions is based on informational aspects—private value auction and common value auction. In a private value auction, bidder valuation is private information. For example, the willingness-to-pay for a valuable painting is highly subjective. Bidders do not know each others' valuations. By contrast, in a common value auction, all bidders have similar objective valuation for an object. For example, in a corporate takeover bid, all potential buyers of a company value the target more or less the same as the target company's financial statement is publicly available.

Example 5.1: First- and second-price auctions

There are four potential buyers of a car, $i = 1, 2, 3, 4$. The bids are $b_1 = \$7$, $b_2 = \$9$, $b_3 = \$4$ and $b_4 = \$3$. In the first-price auction, bidder 2, the highest bidder wins the object, and pays \$9. In the second price auction also bidder 2 wins, but they pay \$7.

The revenue for the auctioneer in the second price-auction is clearly lower as they obtain \$7 as opposed to \$9 in the first-price auction. Then, why does the auctioneer settle for the second-highest bid? As we shall see that bidding strategies change according to the payment rule. However, the expected revenues of first- and second-price auctions are the same, and hence, the auctioneer is indifferent between the two formats. In most of the auction formats, the allocation rules are the same—the highest bidder wins the object. The formats mostly differ in payment rules. Of course, there are exceptions. In Chinese auction, the object is allocated probabilistically. In particular,

$$\text{Prob. } \{\text{bidder } i \text{ wins by bidding } b_i\} = \frac{b_i}{\sum_{j=1}^{n} b_j}.$$

Example 5.2: English auction

There are four potential buyers of a car, $i = 1, 2, 3, 4$. $\theta_1 = \$8$, $\theta_2 = \$12$, $\theta_3 = \$5$ and $\theta_5 = \$2$. Price starts at 0 and increases. Bidders interested in purchasing at current price press button to indicate. Price stops when all but one bidder (say bidder 2) drops out. Bidder 2 wins at the stopped price.

We shall later learn that the English auction and the second price auctions are strategically equivalent. On the other hand, the Dutch and the first-price auctions are strategically equivalent.

5.2 A formal model of independent private value auction

We describe a formal canonical model of private value auctions. There is a set $I = \{1, \ldots, n\}$ bidders with $n \geq 2$. There is a single object to be sold. Bidder i's valuation or willingness-to-pay, θ_i is a random variable which is distributed according to the cumulative distribution function $F_i(\theta_i)$ on a support $[0, v]$. Valuations of any two bidders are independent of each other. We shall also assume that valuations are identically distributed, i.e., $F_i(\cdot) = F(\cdot)$ for all $i \in I$. In other words, all bidders draw valuations from the same probability distribution. Assume that $F(\theta_i)$ is continuous with the pdf f. Also, all bidders are risk neutral.

5.2.1 First-price sealed-bid auction

Recall from Chapter 4 that we have solved for first-price auction for 2 bidders with linear bidding strategies where valuations are drawn from a uniform distribution. Our objective here is to generalize to $n \geq 2$ bidders who can submit any non-linear bidding function and valuations are drawn from the distribution function $F(\cdot)$. The game is as follows. All n bidders submit simultaneously (in sealed envelopes) their bids, $b_1, \ldots, b_n$. The highest bidder wins and pays their valuation.

We shall analyze a symmetric Bayesian Nash equilibrium (BNE) of the bidding game, i.e., $b_i(\theta_i) = b(\theta_i)$ for all $i \in I$.[1] Recall that in a Bayesian game, strategy of any player is a function of their type. We shall assume that $b(\theta_i)$ is increasing, continuous and differentiable on $[0, v]$. To solve for the symmetric BNE, let all bidders j, different from i, submit the identical bidding function, i.e., $b_j(\theta_j) = b(\theta_j)$ for all $j \neq i$. Then, bidder i's expected payoff (as a function of their bid b_i and valuation θ_i) is given by

$$u_i(b_i, b_{-i}, \theta_i) = (\theta_i - b_i) \times \text{Prob.}\{b(\theta_j) \leq b_i, \text{ for all } j \neq i\} = (\theta_i - b_i)[F(b^{-1}(b_i))]^{n-1}.$$

Thus, bidder i chooses b_i to solve

$$\max_{b_i} \ (\theta_i - b_i)[F(b^{-1}(b_i))]^{n-1}.$$

The first-order condition of the above maximization problem is given by

$$-(F(b^{-1}(b_i)))^{n-1} + (\theta_i - b_i)(n-1)(F(b^{-1}(b_i)))^{n-2} f(b^{-1}(b_i))(b^{-1})'(b_i) = 0.$$

Because we want to show that $b_i = b(\theta_i)$ is the best response of bidder i against $b_j = b(\theta_j)$ for all $j \neq i$, replace b_i by $b(\theta_i)$ in the above expression. Note that $(b^{-1})'(b(\theta_i)) = 1/b'(\theta_i)$. Finally, ignoring the subscript i, the above first-order condition gives rise to the following linear differential equation

$$b'(\theta) = (n-1)(\theta - b(\theta)) \cdot \frac{f(\theta)}{F(\theta)}. \tag{5.1}$$

There are several ways to solve the above differential equation. One way to do is the following. Write (5.1) as

$$\underbrace{b'(\theta)(F(\theta))^{n-1} + b(\theta)\{(n-1)(F(\theta))^{n-2} f(\theta)\}}_{\frac{d}{d\theta}[b(\theta)(F(\theta))^{n-1}]} = \underbrace{\theta(n-1)(F(\theta))^{n-2} f(\theta)}_{\theta \cdot \frac{d}{d\theta}[F(\theta)^{n-1}]}. \tag{5.2}$$

Taking integrals on both sides of (5.2), we obtain

$$\int_0^\theta d[b(x)(F(x))^{n-1}] = \int_0^\theta x d[(F(x))^{n-1}]. \tag{5.3}$$

The left-hand-side of (5.3) is $b(\theta)(F(\theta))^{n-1}$ because $F(0) = 0$. On the other hand, the right-hand-side is given by $\theta(F(\theta))^{n-1} - \int_0^\theta (F(x))^{n-1} dx$. Then, it follows from (5.3) that

[1] See Krishna (2010, Appendix G) for the general existence result for first-price auction.

Proposition 5.1: Equilibrium bidding in first-price auction

In first-price sealed-bid auction with n bidders, and identically and independently distributed bidder valuations with cdf $F(\cdot)$, the symmetric BNE bidding strategy is given by

$$b(\theta) = \theta - \frac{\int_0^\theta (F(x))^{n-1} dx}{(F(\theta))^{n-1}}. \tag{5.4}$$

Clearly, all bidders bid below their true valuation, i.e., $b(\theta) < \theta$. The reason is simple. If a bidder bids their valuation and wins the auction, their payoff would be zero which is exactly equal to that if they do not win. Thus, the objective is to obtain a strictly positive expected payoff in case of winning the object. This is in contrast with the second price auction wherein the highest bidder wins but pays the highest losing bid, and hence, bidding own valuation does not induces zero payoffs in case of winning. Note that the degree of bid shading is given by

$$\theta - b(\theta) = \int_0^\theta \left[\frac{F(x)}{F(\theta)}\right]^{n-1} dx,$$

which depends on the number of competing bidders in that as n increases the above quantity approaches zero, and hence, $b(\theta)$ approaches θ. However, the above expression cannot be computed without knowing the exact functional form of $F(\theta)$. The following example analyzes the equilibrium bidding function for uniform distribution.

Example 5.3: First-price auction with uniform distribution

Let $\theta \sim U[0,\,1]$. Then $F(\theta) = \theta$. In this case, we have

$$\theta - b(\theta) = \int_0^\theta \left(\frac{x}{\theta}\right)^{n-1} dx = \frac{\theta}{n} \quad \Longleftrightarrow \quad b(\theta) = \left(1 - \frac{1}{n}\right)\theta.$$

Thus, all bidders adopt a linearly increasing bidding strategy. When, $n = 2$, we have $b(\theta) = \theta/2$.

Note in first-price auction (as well as in many "standard" auctions) that the highest-bidder wins the object. So, it is important to analyze the second-highest bid. Fix a bidder, say i, and let the random variable $Y_i = \max\{\theta_1,\,\ldots,\,\theta_{i-1},\,\theta_{i+1},\,\ldots,\,\theta_n\}$, i.e., Y_i be the second-highest valuation among the remaining $n-1$ bidders, $I \setminus \{i\}$. Let G denote the distribution function of Y_i, i.e., $G(y) = \text{Prob.}\{Y_i \le y\}$. Note that

$$\begin{aligned} G(y) &= \text{Prob.}\{Y_i \le y\} \\ &= \underbrace{\text{Prob.}\{\theta_1 \le y\}}_{F(y)} \times \ldots \underbrace{\text{Prob.}\{\theta_{i-1} \le y\}}_{F(y)} \times \underbrace{\text{Prob.}\{\theta_{i+1}}_{F(y)} \le y\} \times \ldots \times \underbrace{\text{Prob.}\{\theta_n \le y\}}_{F(y)} \\ &= (F(y))^{n-1}. \end{aligned}$$

If $g(y)$ is the associated density function, it is given by

$$g(y) = (n-1)(F(y))^{n-2} f(y).$$

Note that bidder i with valuation θ wins a first-price auction if $b(Y_i) \le b(\theta)$. Because $b(\cdot)$ is an increasing function, the winning probability in first-price auction is given by

$$\text{Prob.}\{b(Y_i) \le b(\theta)\} = \text{Prob.}\{Y_i \le \theta\} = G(\theta) = (F(\theta))^{n-1}.$$

With the above modification, (5.2) can be written as

$$b'(\theta)G(\theta) + b(\theta)g(\theta) = \theta g(\theta). \tag{5.5}$$

Integrating the above, we obtain

$$b(\theta) = \frac{\int_0^\theta y g(y) dy}{G(\theta)} = \mathbb{E}(Y_i \mid Y_i \leq \theta). \tag{5.6}$$

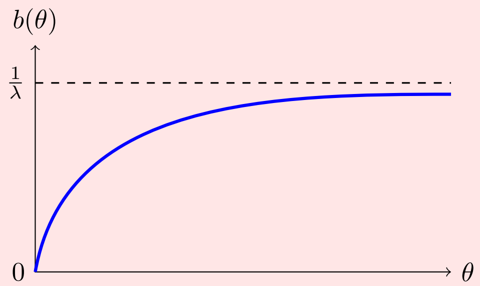
The symmetric BNE bidding strategy is thus a conditional expectation of the highest competing valuation. Now, consider the following example.

Example 5.4: First price auction with exponential distribution

Let $n = 2$ and valuations be exponentially distributed on $[0, \infty)$ with parameter $\lambda > 0$, i.e., $F(\theta) = 1 - e^{-\lambda\theta}$. In this case, the symmetric BNE bidding functions are given by

$$b(\theta) = \frac{1}{\lambda} - \frac{\theta e^{-\lambda\theta}}{1 - e^{-\lambda\theta}}.$$

The equilibrium bidding function is increasing and concave with $\lim_{\theta\to\infty} b(\theta) = \frac{1}{\lambda} = \mathbb{E}(\theta)$. The following figure depicts $b(\theta)$ for $\lambda = 1$.



An interesting point to note is that the bidders bid less aggressively as their valuation grows. For example, when $\theta = \$1$, the symmetric equilibrium bid is given by $b(\$1) = \$0.42$. However, a bidder with valuation \$100 would not bid more than \$1, i.e., $b(\$100) \leq \$1$. In other words, a 9,900% increase in valuation explains a maximum of 19.048% increase in the equilibrium bids. The reason is that there is a very small chance that a bidder with high valuation would lose in equilibrium. In fact, for $\lambda = 1$, a bidder with valuation \$1 does not win with probability $e^{-1} = 0.367879$, whereas a bidder with valuation \$100,000 loses with probability $e^{-100} = 3.72008^{-44}$.

Now, we would relax the assumption of risk-neutral bidders. Let all bidders have identical utility functions $v(\cdot)$ with $v'(\cdot) > 0$ and $v''(\cdot) < 0$. We continue analyzing a symmetric BNE wherein $b_i(\theta_i) = b(\theta_i)$ for all $i \in I$. Thus, bidder i chooses b_i to solve

$$\max_{b_i} v(\theta_i - b_i) \times \text{Prob.}\{b(\theta_j) \leq b_i, \text{ for all } j \neq i\} = v(\theta_i - b_i)[F(b^{-1}(b_i))]^{n-1}.$$

Notice that, relative to the case of risk-neutral bidders, the ex-ante probability that bidder i wins does not change under risk-aversion. The only difference arises from the utility function of each bidder, $v(\cdot)$. Solving the following exercise would give us some idea about the behavior of risk-averse bidders in first-price auction.

Exercise 5.1: Risk-averse bidders in first-price auction

Consider a first-price auction with two bidders, i.e., $n = 2$ wherein each bidder has CRRA utility function, $v(\omega) = \omega^\alpha$ with $0 < \alpha < 1$, and valuations are distributed according to the cdf $F(\theta_i)$ on $[0, v]$ for all $i \in I$. The two extreme cases correspond to $\alpha = 1$ that implies risk-neutral bidders, and $\alpha = 0$ meaning extreme risk aversion. Find the symmetric BNE bidding function $b(\theta)$. Now, assume that valuations are uniformly distributed on $[0, 1]$. Show that the bidder bid more aggressively (i.e., lesser bid shading) as they become more risk-averse (i.e., α decreases) with the limiting case that

$\lim_{\alpha \to 0} b(\theta) = \theta$.

Finally, we discuss the relation between first-price auction and other auction formats, in particular, the Dutch auction.

Proposition 5.2: Strategic equivalence

The first-price sealed-bid auction is strategically equivalent to the descending-price or the Dutch auction.

The intuition behind the above result is simple. Descending price auction is a dynamic process where the price starts at a higher level, and starts dropping over time. If a bidder calls out "mine" at any prevailing price, the auction stops, and that bidder gets the object and pays the prevailing price. Therefore, the bids are not history-dependent in that the prevailing price contains all the information. If the auction is still on at some price, say $50, this means that potential buyer has not called out when the price was higher than $50. Because a bidder just needs to decide at what price he will shout "mine", he can decide it just before the auction starts. This implies that we can perfectly study bidders' behavior in a Dutch auction by assuming that they indeed choose their bids (i.e., their strategies) before the auction starts. But then, because the winner is the bidder with the highest bid, and the price is the winner's bid, the Dutch auction is equivalent to the first-price sealed-bid auction.

5.2.2 Second-price sealed-bid auction

The equilibrium bidding behavior in a second-price sealed-bid auction has already been analyzed in Chapter 4 [see Example 4.7]. In what follows, we shall compare second-price auction with other auction formats.

Proposition 5.3: Strategic equivalence

The second-price sealed-bid auction is strategically equivalent to the ascending-price or the English auction.

The English or ascending-price auction is a dynamic auction, and a bidder must decide at every point whether to continue or not. A potential buyer's bidding strategy is to pick a price at which to drop out (dynamically). Let bidder i has valuation θ_i. Should the bidder be willing to buy the object when the price exceeds valuation? If $p > \theta_i$, bidder i drops out of the auction. By contrast, if $p < \theta_i$, then bidder i stays in the auction until the price reaches their valuation. If θ_i is the highest valuation, then they pay $p = \theta_i$. In other words, the optimal bidding strategy in English auction is $b_{EA}(\theta_i) = \theta_i$ for all $i \in I$. Thus, English auction is strategically equivalent to second-price sealed-bid auction.

5.3 Revenue equivalence

In this section, we analyze a striking result in auction theory—under very general conditions, an auctioneer is indifferent among all "standard" auction formats. An auction is said to be standard if the allocation rule dictates that the highest-bidder is awarded the object.

Theorem 5.1: The revenue equivalence theorem

Suppose that valuations are independently and identically distributed, and all bidders are risk-neutral. Then, any symmetric and increasing equilibrium of any standard auction, such that the lowest-valuation bidder makes zero payment in expectation, yields the same expected revenue to the seller.

Proof. Consider a standard auction format, and denote it by a. Fix a symmetric equilibrium increasing bidding function $b(\cdot)$ of a. Let $m^a(\theta)$ denote the equilibrium expected payment a bidder with valuation θ makes in the auction. Suppose that $b(\cdot)$ is such that $m^a(0) = 0$. Consider a given bidder, say bidder i with valuation θ who bids $b(\theta')$ instead of the equilibrium bid $b(\theta)$. Bidder i wins the auction if $b(Y_i) \leq b(\theta')$ where Y_i is the second-highest valuation. This is equivalent to $Y_i \leq \theta'$. Bidder i's expected payoff is

$$\pi^a(\theta', \theta) = G(\theta')\theta - m^a(\theta'), \tag{5.7}$$

where $G(y) = (F(y))^{n-1}$ is the distribution of the second-highest valuation, Y_i. Note that $m^a(\theta')$ depends on the other players' strategy b through $G(\theta')$ and θ', but not on the true valuation θ. The first-order condition associated with (5.7) is

$$\frac{\partial \pi^a(\theta', \theta)}{\partial \theta'} = g(\theta')\theta - \frac{dm^a(\theta')}{d\theta'} = 0.$$

At an equilibrium, it is optimal to report $\theta' = \theta$, and hence, the above first-order condition becomes

$$\frac{dm^a(\theta)}{d\theta} = g(\theta)\theta$$

for all $\theta \in [0, v]$. Thus,

$$m^a(\theta) = m^a(0) + \int_0^\theta yg(y)dy = \int_0^\theta yg(y)dy = G(\theta) \times \mathbb{E}(Y_i \mid Y_i \leq \theta) \tag{5.8}$$

because, by assumption, $m^a(0) = 0$. Note that $m^a(\theta)$ is a random variable because θ is a random variable. The ex-ante expected payment of any given bidder (before they picked their valuation from $F(\cdot)$, and hence, the terminology, *ex-ante*) is given by

$$\mathbb{E}[m^a(\theta)] = \int_0^v m^a(\theta)f(\theta)d\theta = \int_0^v \left(\int_0^\theta yg(y)dy\right) f(\theta)d\theta.$$

The above is how much the seller will receive in expectation from a given bidder. Because there are n bidders, the expected revenue for the seller from auction a is

$$\mathbb{E}[R^a] = n \times \mathbb{E}[m^a(\theta)] = n \times \mathbb{E}[G(\theta) \times \mathbb{E}(Y_i \mid Y_i \leq \theta)]. \tag{5.9}$$

Because the right-hand-side of (5.9) is independent of the auction format, a, the theorem holds. □

In the following example, we compute the expected revenue of a standard auction for a specific distribution function of valuations.

Example 5.5: Expected revenue under power distribution

Let valuations follow a power distribution with parameter $\alpha > 0$ on $[0, 1]$, i.e., $F(\theta) = \theta^\alpha$. The corresponding density function is $f(\theta) = \alpha\theta^{\alpha-1}$. Thus, $G(\theta) = \theta^{\alpha(n-1)}$, and $g(\theta) = \alpha(n-1)\theta^{\alpha(n-1)-1}$ which implies $\theta g(\theta) = \alpha(n-1)\theta^{\alpha(n-1)}$. The expected payment of a bidder with valuation θ is given by

$$m^a(\theta) = \int_0^\theta yg(y)dy = \int_0^\theta \alpha(n-1)y^{\alpha(n-1)}dy = \frac{\alpha(n-1)\theta^{\alpha(n-1)+1}}{\alpha(n-1)+1}.$$

Therefore,

$$\mathbb{E}[m^a(\theta); \alpha] = \int_0^1 m^a(\theta)f(\theta)d\theta = \frac{\alpha^2(n-1)}{\alpha(n-1)+1}\int_0^1 \theta^{\alpha n}d\theta = \frac{\alpha^2(n-1)}{(\alpha(n-1)+1)(\alpha n+1)}.$$

The expected revenue thus is

$$\mathbb{E}[R^a; \alpha] = n \times \mathbb{E}[m^a(\theta)] = \frac{\alpha^2 n(n-1)}{(\alpha(n-1)+1)(\alpha n+1)}\,.$$

When $\alpha = 1$, we have uniform distribution on $[0,\,1]$. The expected revenue is thus given by

$$\mathbb{E}[R^a; 1] = \frac{n-1}{n+1}\,.$$

From the Revenue equivalence theorem 5.1, it follows that first- and second-price auctions are revenue-equivalent. We should also be careful in interpreting the revenue equivalence principle. This result holds under several assumptions—valuations are private and they are independently and identically distributed, and bidders are risk-neutral. The revenue equivalence theorem thus serves as a benchmark; relaxing some of the assumptions would allow us to rank various auction formats in terms of seller's expected revenues. For example, if bidders were risk-averse, first-price auction yields higher expected revenue than second-price auction (why?). We have so far also ignored the analysis of efficiency. We have already mentioned that in a single-object auction wherein the highest-bidder wins, the auction is ex-post efficient. So, all auction formats we have studied above are efficient.

Bibliography

Aumann, Robert J. (1964), "Markets with a continuum of traders." *Econometrica*, 32, 39–50.

Baron, David and Roger Myerson (1982), "Regulating a Monopolist with Unknown Cost." *Econometrica*, 50, 911–930.

Bolton, Patrick and Mathias Dewatripont (2005), *Contract Theory*. The MIT Press.

Fuchs, William (2007), "Contracting with Repeated Moral Hazard and Private Evaluations." *The American Economic Review*, 97, 1432–1448.

Grossman, Stanford and Oliver Hart (1983), "An Analysis of the Principal-Agent Problem." *Econometrica*, 51, 7–45.

Hölmstrom, Bengt (1979), "Moral Hazard and Observability." *Bell Journal of Economics*, 10, 74–91.

Hölmstrom, Bengt (1982), "Moral Hazard in Teams." *Bell Journal of Economics*, 13, 324–340.

Hölmstrom, Bengt and Paul Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics and Organization*, 7, 24–52.

Innes, Robert (1990), "Limited Liability and Incentive Contracting with Ex-Ante Action Choices." *Journal of Economic Theory*, 52, 45–67.

Khalil, Fahad, Jacques Lawarrée, and Troy Scott (2015), "Private Monitoring, Collusion, and the Timing of Information." *The RAND Journal of Economics*, 46, 872–890.

Krishna, Vijay (2002), *Auction Theory*, second edition. Academic Press.

Krishna, Vijay (2010), *Auction Theory*, 2nd edition. Academic Press, San Diego, CA.

Laffont, Jean-Jacques and Jean Tirole (1986), "Using Cost Observation to Regulate Firms." *Journal of Political Economy*, 94.

Lazear, Edward and Sherwin Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89, 841–864.

Legros, Patrick and Stephen Matthews (1993), "Efficient and Nearly Efficient Partnerships." *The Review of Economic Studies*, 60, 599–611.

Levin, Jonathan (2003), "Relational Incentive Contracts." *The American Economic Review*, 93, 835–857.

MacLeod, Bentley (2003), "Optimal Contracting with Subjective Evaluation." *The American Economic Review*, 93, 216–240.

Maskin, Eric and John Riley (1984), "Monopoly with Incomplete Information." *RAND Journal of Economics*, 15, 171–196.

Mookherjee, Dilip (1984), "Optimal Incentive Schemes with Many Agents." *The Review of Economic Studies*, 51, 433–446.

Myerson, Roger (1981), "Optimal Auction Design." *Mathematics of Operations Research*, 6, 58–63.