*CHAPTER 8*

# ON THE THEORY OF DELEGATION

BENGT HOLMSTRÖM*
*Northwestern University*

## 1. Introduction

This chapter concerns decentralized decision-making in a nonmarket organization. A principal has to make a decision under uncertainty. He has available agents whom he may consult in the process of making the decision, because the agents possess some relevant information. The problem the principal faces in using agents is that the agents may have different objectives than he has. I am interested in the principal's problem of designing a decision mechanism, which optimally exploits the agents' cooperation.

My specific concern is with delegation as a means for organizing decision-making. In delegation the agent is given freedom to make the decision subject to constraints specified by the principal. A central issue is the relationship between the agent's freedom of choice, the information he possesses, and the divergence in preferences. This question is addressed in the context of simple one-dimensional quantity controls. For a stylized model of economic planning, it is shown that the agent will be given more freedom when he gets more informed or when his preferences come closer to the principal's. The model extends Weitzman's (1974) work on prices versus quantities and gives an indication of the relationship between the tightness of economic control and information.

The result on information and the agent's freedom of choice is highly model specific, whereas the result on the effects of preference divergence can be extended. I do so by considering the general one-dimensional control problem with the principal using interval (quantity) controls. In this case agents can be partially ordered by their unrestricted response functions, so that the principal always prefers and gives more freedom to an agent whose response function is uniformly closer than another agent's.

* This chapter is a much revised version of "On Decentralization, Incentives and Control", RP 366, Graduate School of Business, Stanford University, March 1977. I have benefited from suggestions by Joel Demski, Froystein Gjesdal, Sanford Grossman, Charles Holloway, Richard Kihlstrom, David Kreps, Takao Kobayashi, and Robert Wilson to whom I express my thanks.

*Bayesian Models in Economic Theory, edited by M. Boyer and R.E. Kihlstrom*

Without specifying probability distributions for the agents' information, no stronger ordering is possible. For higher-dimensional control problems (like nonlinear pricing) one has to be content with even weaker orderings; generally, only agents whose preferences lie on a ray originating from the principal's can be compared independently of probability distributions. Though one finds that some agents can be compared, these results are rather negative for the prospects of modelling agency markets.

A basic question of interest in delegation is when it will be a useful tool for cooperation; that is, when is it of value for the principal? This question is of relevance for understanding the emergence of economic institutions that are based on agency relationships. I give two sufficient conditions for a positive value of delegation, which show that the value depends on how coherent the principal's and the agent's preferences are and what opportunities there are to control the agent so that his induced preferences align more closely to the principal's.

The rest of the chapter is organized as follows. Section 2 presents a quite general formulation of the decentralization problem, which specializes to delegation when different agents' information do not get coordinated in the decision process. The delegation problem is reformulated in section 3. A theorem on the existence of optimal control sets is proved and examples are presented to indicate the general applicability of delegation. Section 4 treats the earlier mentioned extension of Weitzman's planning problem and section 5 contains the general results on the use of interval controls. The value of delegation is studied in section 6 and conclusions with a comparison to some of the literature are contained in section 7.

## 2. A general formulation of the decentralization problem

Suppose there is a decision $d$ to be made such that it belongs to a prespecified set of alternatives $D$, called the *decision space*. A *decision process* can quite generally be described as a mapping $d: M \to D$, from inputs $m \in M$ to final decisions $d(m) \in D$.

We will here be concerned with decision processes that have $n$ participants and can be described as follows. Each participant $i$ chooses a *message* $m_i$ from his set of alternative messages $M_i$, called *i's message space*. The final decision is determined by the message $n$-tuple $m = (m_1, ..., m_n)$ via a *decision function* $d: M \to D$, where $M = \times_{i=1}^{n} M_i$ is called the *joint message space*. If the decision function $d$ depends on more than one $m_i$ we have a *decentralized decision process*. The pair $N = (d, M)$ is called a *decision mechanism*.

We are looking at the problem of a single decision-maker, called the principal, who has to choose $d \in D$ facing some uncertainty. The uncertainty is described by a probability space $(Z, \mathcal{F}, P)$, where $z \in Z$ is the *state of nature* and $P$ is the principal's subjective probability measure on events in $\mathcal{F}$. The principal's objective is to maximize the expected value of his *preference function* $F_0: D \times Z \to \mathbb{R}^1$, which for each $d \in D$ is a measurable function on $(Z, \mathcal{F})$.

The key feature of our problem structure is that the principal has available $n$ agents (indexed by $i = 1, ..., n$), who possess some private information about the state of nature. Each agent $i$ has observed the outcome of a random variable $\tilde{y}_i$ on the probability space $(Z, \mathcal{F}, P)$, which provides information about $\tilde{z}$. For this reason the principal is interested in using the agents in the decision process by defining a decision mechanism $N = (d, M)$ which decentralizes the decision as described above. The agents' messages will normally correspond to part or all of their private information.

What decision mechanism $N$ should the principal use? The problem that the principal faces is that each agent will act in his own self-interest according to a preference function $F_i$: $D \times Z \to \mathbb{R}^1$, $i = 1, \ldots, n$.[1] Any particular choice of decision mechanism will result in a noncooperative game between agents. I will model this as a game of incomplete information (see Harsanyi, 1967–1968) so that the agents' strategies are message functions $m_i(y_i)$ of their private information. It is assumed that a Nash equilibrium (or a Bayesian equilibrium) is the appropriate solution concept. This gives the principal the basis for evaluating different decision mechanisms.

Let me now formulate precisely the principal's problem. For this we have to assume that the agents know each others preference functions $F_i$ and the functional form of each other's private information (not the outcomes). They should also agree on the specification of the probability space $(Z, \mathcal{F}, P)$. These assumptions can be defended as in Harsanyi (1967–1968), and rest on the idea that the probability space can be augmented sufficiently to incorporate all differences in information. Given a decision mechanism $N = (d, M)$ a Nash equilibrium is defined as a set of functions $\{\bar{m}_i(y_i; N)\}_{i=1}^n$ satisfying:

$$E[F_i(d(\bar{m}(y;N)), z)] \geqslant E[F_i(d(\bar{m}^i(y;N), m_i(y_i)), z)], \tag{1}$$

for every alternative strategy $m_i(y_i)$ and for every $i = 1, \ldots, n$. Here $\bar{m}(y;N) = (\bar{m}_1(y; N) \ldots, \bar{m}_n(y;N))$, $\bar{m}^i = (\bar{m}_1, \ldots, \bar{m}_{i-1}, \bar{m}_{i+1}, \ldots, \bar{m}_n)$, and $(m^i, m_i) \equiv m$. The existence of a Nash equilibrium for each $N$ is assumed.

In choosing among different decision mechanisms $N$, the principal is restricted to a *set of admissible decision mechanisms* $\mathcal{N}$. I will comment on this set below. The principal's problem can now be stated as follows:

**Decentralization problem.** Choose a decision mechanism $N \in \mathcal{N}$ such that it maximizes

$$E[F_0(d(\bar{m}(y;N), z))], \tag{2}$$

where $\bar{m}(y;N)$ is a Nash equilibrium for each $N$, defined by (1).

If there is more than one Nash equilibrium, assume the principal will choose one of them as a basis for his optimization problem. The existence of a solution to the decentralization problem will only be discussed in connection with delegation.

There is another way of viewing the principal's problem defined above. Let $\bar{d}(y) =$

[1] It is a notational short-cut to write preferences directly over decisions. In later examples we see how these functions may be derived from more primitive concepts.

$d(\bar{m}(y))$. The function $\bar{d}(\cdot)$ describes what eventually will be realized in equilibrium given the outcomes of the agent's signals. It is called the *outcome function*. An outcome function is said to be *attainable* if there exists a decision mechanism $N \in \mathcal{N}$ which yields that outcome function at a Nash equilibrium. Denote the set of attainable outcome functions by $\mathcal{K}$. The principal's problem can then equivalently be phrased as:

Find $\bar{d}(\cdot) \in \mathcal{K}$ which maximizes

$$E[F_0(\bar{d}(y), z)]. \tag{3}$$

Or, verbally, find the best attainable outcome function.

Let the range of $y_i$ be $Y_i$ and let $Y = \mathrm{X}_{i=1}^{n} Y_i$. A simple but useful result is the following:[2,3]

**Revelation principle.** An outcome function $\bar{d}(\cdot)$ is attainable if and only if the decision mechanism $N = (\bar{d}(\cdot), Y)$ has a Nash equilibrium such that $\bar{m}_i(y_i) = y_i$.

This principle shows that the principal need not look at general message spaces $M$ in the design problem. He can always take $M = Y$ and concentrate on the choice of the decision function alone. However, he can only consider decision functions that yield truth-telling Nash equilibria.

A few remarks about the formulation are in order.

(1) We have avoided a further complication of the problem by assuming that the principal has no private information of interest to the agents. Otherwise we would have had to include the principal as a player in the game of incomplete information with his strategy being the choice of $N$ as a function of his information.

(2) Randomized strategies can be included in the formulation by augmenting the agents' signal spaces $Y$, by independent and uniformly distributed random variables. Likewise, the principal could use the outcome of a uniformly distributed random variable to randomize his decision rule. All of this, of course, could be embodied in the specification of $z$.

(3) By appropriately specifying $\mathcal{N}$ we could allow agents to take private actions (as part of $d$), which cannot be controlled by the principal (e.g. levels of effort).[4]

(4) Rather than maximizing the principal's objective, we could look for efficient

[2] Gibbard (1973) was apparently the first to use the revelation principle in his study of game form design. Since then, many authors have rediscovered the principle. Dasgupta, Hammond and Maskin (1979) provide a detailed study of the principle under various solution concepts. Myerson was the first to use the term revelation principle, which I find descriptive.

[3] There is one restriction of the revelation principle which is important. The mechanism $N = (\bar{d}(\cdot), Y)$, sometimes called a direct revelation mechanism (Green and Laffont, 1979), has to belong to the set $\mathcal{N}$. This need not always be the case, but is assumed here.

[4] Holmström (1977) and Myerson (1982) provide an explicit formulation with nonobservable actions. Notice that restrictions on $\mathcal{N}$ will in general be quite complicated. For instance, they would include the requirement that nonobservable actions form a Nash equilibrium given the information of the agents.

decision mechanisms by having the principal act according to the preference function:

$$\sum_{i=1}^{n} E[\lambda_i(y_i)\, F_i(d, z)] + E[F_0(d, z)]. \tag{4}$$

(See Harris and Townsend, 1981; Holmström, 1977, section 2.5; Myerson, 1979.)

A few examples may illustrate the general decentralization formulation.

If all $F_i$ are identical we have a team (Marschak and Radner, 1972). Each agent $i$ is in charge of a decision $d_i$. Their private information can be communicated through an information system $I$, which may be the null system. The center's problem is to find the optimal decision rules to be employed. For this problem, we can take $F_0 = F_i$, $d = (d_1, \ldots, d_n)$. The restrictions on the decision rule will be determined by the information system $I$ and can be incorporated through a definition of $\mathcal{N}$. Since we have a team we are not making any effective changes by letting the (auxiliary) principal design the decision mechanism, so the team problem is indeed subsumed in the general formulation.

Another example is the revelation of preferences for public goods (Groves, 1973; Green and Laffont, 1979). The public good decision is $d_0$. Individuals have valuation functions $V_i(d_0, y_i)$ over the public good and their preferences are linear in money. Let $(d_1, \ldots, d_n)$ denote transfer payments to the principal. Then $F_i(d, z) = V_i(d_0, y_i) - d_i$, $F_0(d, z) = \Sigma_i\, V_i(d_0, y_i)$, and the problem fits the general framework.

A final example is the design of auctions (Harris and Raviv, 1981; Myerson, 1981; Maskin and Riley, 1980). An object is set up for seal-bid auction by a seller (principal). Buyers (agents) have information about their preferences and the quality of the object is given by $(y_1, \ldots, y_n)$. Let $d_0$ be the award decision and $(d_1, \ldots, d_n)$ the payments from buyers to sellers. Let $y_0$ be a randomizing signal for the seller. Preferences are given by $F_0(d_0, d_1, \ldots, d_n, y)$ and $F_i(d_0, d_i, y)$. The seller's problem is to find the best rule $d_0: Y \to \{1, \ldots, n\}$, $d_i: Y \to \mathbb{R}$, which yields truth-telling as a Nash equilibrium as described in the general formulation.

## 3. The delegation problem

An important subclass of the decentralization problem is arrived at by restricting the decision function to the form $d(m) = (d_1(m_1), \ldots, d_n(m_n))$. In that case no information gets coordinated in the decision process, and the process becomes equivalent to delegating the decision $d_i$ to agent $i$ subject to the constraint that

$$d_i \in C_i = \{d_i \mid \exists\, m_i \in M_i \quad \text{s.t.} \quad d_i = d_i(m_i)\}. \tag{5}$$

Rather than choosing decision functions $d_i(\cdot)$, the principal can be viewed as choosing *control sets* $C_i$. A reason for not coordinating information may be the transactions costs saved by the simpler decision process. In any case, it appears that delegation is a very prevalent mode of decentralization.

The central features of delegation can best be studied by analyzing the case of a single agent. Then the delegation process proceeds simply as follows: the agent observes a signal $y$, the principal chooses a control set $C \subseteq D$, and the agent determines the final decision by choosing $d \in C$ so as to maximize his own benefit. Formally, given $C$ and the private signal $y$, the agent solves the program (from now on superscript A refers to agent, P to principal):

$$\max_{d \in C} E[F^{\mathrm{A}}(d, z) \mid y]. \tag{6}$$

$D$ will be taken to be compact and $C$ closed so a solution always exists. Let $d(y, C)$ be the set of maximizers in (6). A selection from $d(y, C)$ for each $y$ will be denoted $d(y \mid C)$ and called a *response function*. Notice that we can define a response function for the principal as well even though he does not observe $y$. When necessary, $d_{\mathrm{A}}(\cdot)$ will identify the agent's response function, $d_{\mathrm{P}}(\cdot)$ the principal's (particular selections when necessary).

It is essential how the agent resolves ties. One possibility is the rule:

$$d_{\max}(y \mid C) = \operatorname*{argmax}_{d \in d(y, C)} E[F^{\mathrm{P}}(d, z) \mid y]; \tag{7}$$

i.e. the agent acts according to the principal's interests in case of a tie. This will ensure existence of an optimal control set, as seen below. For the moment though, let $d(y \mid C)$ be the response function that the principal perceives as the agent's behavior. Then, we may define the principal's problem as:

**Delegation problem:** Choose $C \in \mathcal{N}$ such that it maximizes

$$E[F^{\mathrm{P}}(d(y \mid C), z)]. \tag{8}$$

Here $\mathcal{N}$ is the class of feasible control sets. (We use $\mathcal{N}$ again, since a control set is equivalent to a decision mechanism.)

It deserves to be mentioned that there is no reason to delegate if the agent does not have any private information. More generally, the control set $C$ has the cardinality of the signal space $Y$. Of course, if $F^{\mathrm{P}} = F^{\mathrm{A}}$, setting $C = D$ is optimal.

One of the virtues of the delegation formulation is the following result:[5]

[5] An example of nonexistence is the following. Let $D = \{0,2\}$. Let $y = z$ and $Y = \{0,2\}$; Prob $\{y = 0\}$ = Prob $\{y = 2\} = 1/2$. Finally, $F^{\mathrm{A}}(d, z) = -(d-z-1)^2$ and $F^{\mathrm{P}}(d, z) = -(d-z)^2$. The supremum of the control problem is 0. It can be approximated arbitrarily closely by $C = \{0, 2 + \epsilon\}$, $\epsilon > 0$. However, if the agent in case of a tie chooses the principal's least preferred action, $C = \{0,2\}$ will make him choose $d = 0$ both when $y = 0$ and $y = 2$. Thus, the supremum cannot be achieved. On the other hand, if he acts according to $d_{\max}(\cdot)$ in (7), $C = \{0,2\}$ will achieve the supremum. Likewise, one can show that any small (continuous) uncertainty about $F^{\mathrm{A}}$ will guarantee existence of a solution. These observations suggested the conditions in theorem 1.

**Theorem 1.** Assume $D$ is compact, $\mathcal{N}$ is a closed subset of the set of all closed subsets of $D$, $F^{\mathrm{P}}$ and $F^{\mathrm{A}}$ are continuous in $d$ and uniformly bounded, and either (i) or (ii) holds:

(i) the agents act according to $d_{\max}(\cdot)$ defined in (7), or

(ii) Prob $\{d(y, C)$ is not singleton$\} = 0$, $\forall C \in \mathcal{N}$.

Then the delegation problem has a solution.

*Proof.* See appendix A.

The theorem gives two simple conditions under which existence is guaranteed. The latter condition is usually satisfied if there is any uncertainty about the agent's preferences.

Let me next turn to some examples of delegation. The first one represents the simplest one-dimensional control problem.

### 3.1. *Economic planning* (Weitzman, 1974)

A central planner tries to control the production of a single good of which an amount $d$ can be produced at a cost $C(d, z)$ for a benefit of $B(d)$. The production unit knows $z$, the center knows $B(d)$. Weitzman considers two modes of control: setting centrally a quantity $d$, or setting a price $p$ and delegating the choice of $d$ to the production unit, which will subsequently maximize its profits $pd\text{-}C(d, z)$. To get a comparison of the two modes one may approximate $B(d)$ and $C(d, z)$ by quadratic functions:

$$B(d) = b + B'(d-d^*) + \tfrac{1}{2} B''(d-d^*)^2, \tag{9}$$

$$C(d, z) = c(z) + (C' - h(z))\ (d-d^*) + \tfrac{1}{2} C''(d-d^*)^2, \tag{10}$$

where $d^*$ is the best centralized action.[6] Without loss of generality, take $h(z) = z$ and $E(z) = 0$, so that $C' = B'$ by optimality of $d^*$.

Weitzman shows that the best price to set in the price mode is $p^* = B' = C'$, and that the price mode is better than the quantity mode exactly when $B'' + C'' > 0$, independently of the distribution of $z$.

I will consider a more elaborate form of delegation, which mixes a policy of prices and quantities. Specifically, the center can set a price and in addition require that $d \in [d_{\mathrm{L}}, d_{\mathrm{U}}]$. If $z$ has a symmetric distribution and the approximation in (9) and (10) is used, it is evident that the optimal price to set will again be $p^* = B' = C'$. Supressing the price decision, what remains is a one-dimensional control problem, with $D \subset \mathbb{R}^1$ and preference functions:

$$F^{\mathrm{P}}(d, z) = B(d) - C(d, z), \tag{11}$$

[6] The action $d^*$ is what the center would choose on its own, not knowing $y$.

$$F^{A}(d, z) = p^{*}d - C(d, z). \tag{12}$$

I will return to this model in the next section.

### 3.2. An insurance model

Since the path-breaking work by Mirrlees (1971) and Spence (1973), many signalling and screening models have appeared in the literature (e.g. Rothschild and Stiglitz, 1976, and Riley, 1979). Though these models deal with more than two persons, they normally involve only bilateral trade, which can be put in the framework of delegation. In order to be specific, I will look at the case of an insurance market (Rothschild and Stiglitz, 1976).

The insurance company is the principal and the insured the agent. In the simplest case, the agent faces a risk of having an accident, described by a random variable $\tilde{x}$ (= 0 if no accident, = 1 if accident). The agent knows his probability of having an accident, denoted by $y$, whereas the insurance company is uncertain about $y$. We can then take $\tilde{z} = (\tilde{x}, \tilde{y})$. The decision is what insurance policy $d$ should be offered to the agent. Let $d = (a, t)$, where $a$ is the level of insurance (payment in case of an accident net of premium) and $-t$ is the premium. The company is assumed to be risk-neutral, and has a preference function, given by:

$$\begin{aligned} F^{P}(d,z) &= -t \quad \text{if} \quad \tilde{x} = 0, \\ &= -a \quad \text{if} \quad \tilde{x} = 1. \end{aligned}$$

The agent is risk-averse, with a utility function $U$ and consequently a preference function:

$$\begin{aligned} F^{A}(d,z) &= U(t) \quad \text{if} \quad \tilde{x} = 0, \\ &= U(a) \quad \text{if} \quad \tilde{x} = 1. \end{aligned}$$

Since the company does not know $y$, it turns out that it is beneficial to delegate the choice of an insurance policy to the agent by offering him a set $C$ of insurance policies, rather than a single one. The agent prefers, for any fixed level of insurance $a$, smaller premiums to larger, so the company only needs to consider delegation of sets of the form $C = \{(a, S(a))\}$. This can be seen from fig. 8.1. $S(a)$ is the negative of the price the agent has to pay for the level $a$ of insurance. We see that the company's optimal pricing problem is one of optimal delegation.

Remark. It is of interest to note that closed subsets of $\mathbb{R}^2$ correspond to closed epigraphs of the price function in screening models. This means that theorem 1 guarantees the existence of an optimal solution for these problems if price functions are upper semi-continuous.

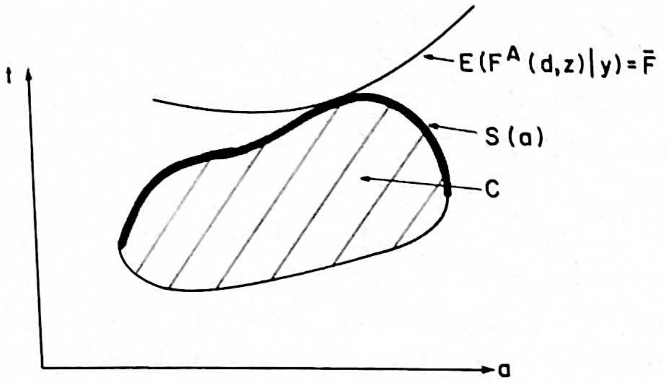


Figure 8.1

*3.3. Moral hazard* (Mirrlees, 1976; Holmström, 1979, 1982)

The agent's utility function is $U^A(w, a)$, defined over wealth $w$ and effort $a$; the principal's is $U^P(w)$, defined over wealth alone. Effort affects the distribution of the monetary outcome $x$. The agent has observed a signal $y$ before the principal determines the sharing rule $s(x)$. Therefore, it generally pays the principal to let the agent participate in the choice of $s(x)$. In the delegation framework, $d = (s(\cdot), a)$, but only $s(\cdot)$ can be controlled by the principal.

Let $f(x \mid y, a)$ be the conditional density of $x$ given the signal $y$ and effort $a$ and let $g(y)$ be the density of $y$. The principal's problem can be stated as:

$$\max_{s(\cdot)} \iint U^P(x - s(x))\, f(x \mid y, a(y))\, g(y)\, \mathrm{d}x\, \mathrm{d}y$$

s.t.

$$a(y) = \underset{a}{\operatorname{argmax}} \int U^A(s(x), a)\, f(x \mid y, a)\, \mathrm{d}x, \quad \forall y.$$

It can be reformulated as a delegation problem, by writing out the derived preference functions:

$$F^P(s(\cdot), a, y) = \int U^P(x - s(x))\, f(x \mid y, a)\, \mathrm{d}x,$$

$$F^A(s(\cdot), a, y) = \int U^A(s(x), a)\, f(x \mid y, a)\, \mathrm{d}x.$$

The examples above serve to indicate the generality of the delegation formulation as well as the structural relationship between rather different looking problems.

## 4. Quantity controls in economic planning

In this section I return to the economic planning problem described above in order to illustrate the use of interval controls. The objective is to find out how economic control depends on the curvature of the benefit and cost functions, and the information gap.

With the quadratic approximations in (9) and (10) the preference functions in (11) and (12) can be written (ignore constants and recall that $p^* = B' = C'$) as the loss functions:

$$L^{\mathrm{P}}(d, z) = \left(d - d^* - \frac{z}{C'' - B''}\right)^2, \tag{13}$$

$$L^{\mathrm{A}}(d, z) = \left(d - d^* - \frac{z}{C''}\right)^2. \tag{14}$$

Define $a_{\mathrm{P}} = 1/(C'' - B'') > 0$, $a_{\mathrm{A}} = 1/C'' > 0$, and $b_{\mathrm{P}} = b_{\mathrm{A}} = d^*$. Then (13) and (14) are of the general form:

$$L^i(d, z) = (d - (a_i z + b_i))^2, \quad i = \mathrm{A,P}. \tag{15}$$

For the analysis of information change it will be useful to add some generality by letting $z = y + x$ and assume that the agent observes $y$ but not $x$. Assume further that $x$ and $y$ are independent so that $x$ represents pure noise in the signal. Because of (15) there is no loss in generality to let $Ez = Ex = Ey = 0$. If necessary $a_i$ and $b_i$ have to be redefined, but not in the planning problem considered, since $Ez = 0$ was already assumed. Let $\mathrm{Var}\, x = s_x^2$ and $\mathrm{Var}\, y = s_y^2$. Conditional on $y$, expected losses are (dropping constants):

$$\ell_i(d, y) = (d - (a_i y + b_i))^2, \quad i = \mathrm{A,P}. \tag{16}$$

Consequently, the response functions are linear:

$$d_i(y) = a_i y + b_i, \quad i = \mathrm{A,P}. \tag{17}$$

Let $(d_{\mathrm{L}}, d_{\mathrm{U}})$ be a control interval. The expected loss for the principal when he uses this control is:

$$E_{\mathrm{P}}(d_{\mathrm{L}}, d_{\mathrm{U}}) = \int_{-\infty}^{y_{\mathrm{L}}} (d_{\mathrm{L}} - d_{\mathrm{P}}(y))^2 \,\mathrm{d}G(y) + \int_{y_{\mathrm{L}}}^{y_{\mathrm{U}}} (d_{\mathrm{A}}(y) - d_{\mathrm{P}}(y))^2 \,\mathrm{d}G(y)$$
$$+ \int_{y_{\mathrm{U}}}^{\infty} (d_{\mathrm{U}} - d_{\mathrm{P}}(y))^2 \,\mathrm{d}G(y) + s_x^2, \tag{18}$$

where $y_{\mathrm{L}}$ and $y_{\mathrm{U}}$ are defined by $d_{\mathrm{A}}(y_{\mathrm{L}}) = d_{\mathrm{L}}$, $d_{\mathrm{A}}(y_{\mathrm{U}}) = d_{\mathrm{U}}$. Since $d_{\mathrm{A}}$ is increasing we can look for optimal values $y_{\mathrm{L}}^*, y_{\mathrm{U}}^*$ instead. It will suffice to analyze the upper limit because the analysis of the lower limit is symmetric.

Differentiating $E_P$ with respect to $y_U$ gives the first-order condition for $y_U^*$:

$$\int_{y_U^*}^{\infty} (a_A y_U^* + b_A - a_P y - b_P)\, dG(y) = 0. \tag{19}$$

This is the condition for $y_U^*$ when we have general quadratic loss functions with linear responses as given by (15) and (16). To get further I will assume that $x$ and $y$ (hence $z$) are normally distributed and return to the specific cases of central planning in which, recall, $a_A > a_P > 0$ and $b_A = b_P$. In that case (19) reduces to:

$$\left(\frac{a_A}{a_P}\right) y_U^* = \frac{g(y_U^*)\, s_y^2}{1 - G(y_U^*)} \quad . \tag{20}$$

Notice that the function $h(y) = g(y)/(1-G(y))$ is the hazard rate for the Normal distribution – a well-known function from reliability theory (see Barlow and Proschan, 1975). In appendix B the following properties of the hazard rate are established:

$$h(y) > y/s_y^2, \quad \forall y, \tag{21}$$

$$0 < h'(y)\, s_y^2 < 1, \quad \forall y, \tag{22}$$

$$\lim_{y \to \infty} \frac{h(y)}{y} = \frac{1}{s_y^2} . \tag{23}$$

Thus, the RHS in (20) has $y$ as an asymptote. Since $a_A > a_P > 0$. (21)–(23) imply that (20) has a unique solution $0 < y_U^* < \infty$. Thus,

$$d_U^* = a_A y_U^* + b_A > b_A = d^*,$$

where $d^*$ is the principal's best centralized act. Symmetrically $d_L^* < d^*$ so $d^*$ is contained in the optimal delegation interval $(d_L^*, d_U^*)$.

Now look at changes in the information gap $\tilde{y}$. Only the RHS of (20) changes with $s_y^2$. In appendix B I show that:

$$\frac{\partial}{\partial s_y} [h(y)\, s_y^2] > 0 \quad \text{for all} \quad y. \tag{24}$$

Since $a_A$, $a_P > 0$, it follows from (20) (by total differentiation w.r.t. $s_y$) that $\partial y_U^*/\partial s_y > 0$, which implies $\partial d_U^*/\partial s_y > 0$. A symmetric analysis shows that $\partial d_L^*/\partial s_y < 0$. Thus, the agent is given more freedom as $s_y^2$, that is, the information gap, increases. Notice that $s_x^2$ does not influence the amount of freedom the agent receives. Therefore, since $s_z^2 = s_y^2 + s_x^2$, an increase in $s_y^2$ could be interpreted either so that the principal gets less informed ($s_z^2$ increases with $s_y^2$) or that the agent gets more informed ($s_z^2$ stays the same and the increase in $s_y^2$ is offset by a decrease in $s_x^2$). In the latter case it is straightforward to compute that the principal gets better off when the agent becomes more informed.

Finally, consider the effect of a change in $C''$ and $B''$. From the expressions for $a_A$ and $a_P$, (20) gives directly that either an increase in $B''$ or a decrease in $C''$ will decrease the freedom of the agent. This accords with Weitzman's findings that when $B'' + C''$ increases, prices tend to be better than quantities.

The preceding discussion is summarized as:

**Proposition.**[7,8] In the economic planning problem approximated by quadratic loss functions (13) and (14) and with a Normally distributed information gap $\tilde{y}$:

(i) The production unit will be given a finite degree of freedom $[d_L^*, d_U^*]$ which will include the best centralized act $d^*$. Neither prices nor quantities are optimal.

(ii) The production unit is given more freedom with an increase in the information gap $\tilde{y}$. If the production unit becomes more informed, the central planner becomes better off.

(iii) The production unit is given more freedom the closer are $a_A$ and $a_P$, that is, when there is either a decrease in the curvature of the benefit function or an increase in the curvature of the cost function.

Part (ii) above should be contrasted with Weitzman's finding that the relative benefit of the price and quantity mode is independent of the information gap. Actually, one can show that $y_U^* \to \infty$ and $y_L^* \to -\infty$ as $s_y^2 \to \infty$, so in the limit the price mode becomes optimal independent of $B''$ and $C''$.[9] Though the information structure is quite special, the analysis suggests that economic control (e.g. pollution control) should be exercised as a function of differences in information and specifically so that control is less tight with a less informed center.

## 5. A general comparison of preferences

In the economic planning problem things come out according to intuition: the agent gets more freedom with better information and with preferences closer to the principal's.

[7] Note that similar results apply for general $a_P$, $a_A$, $b_P$, and $b_A$, not just for the specific ones in the economic planning problem. The proposition is valid (with some obvious modifications) for the general case of Normal uncertainty, quadratic loss functions and linear responses.

[8] An alternative parameterization for which one obtains an explicit solution is the following (Holmström, 1977, section 2.3.2): $a_A = a_P = 1$, $b_A > b_P$, and $y = z = \text{Unif}(-\delta, \delta)$. The optimal control interval has only an upper bound given by $d_U^* = b_P + \max(0, \delta - (b_A - b_P))$. Again, the closer the preferences are ($b_A$ closer to $b_P$) or the bigger the information gap (measured by $\delta$), the more freedom the agent is given.

This example has been further analyzed in Crawford and Sobel (1982) under a different equilibrium concept (see section 7).

[9] We have, $g(0)s_y^2/(1 - G(0)) = 2\varphi(0)s_y \to \infty$ as $s_y \to \infty$, where $\varphi$ is the density of a standardized Normal distribution. Since the hazard rate is increasing by (22), the expression above implies that the RHS of (20) goes pointwise to infinity with $s_y$ for positive values of the argument. Hence, $d_U^* \to \infty$ as $s_y \to \infty$. Similarly, $d_L^* \to -\infty$.

Unfortunately, such is not the case in general. Improved information may reduce the agent's freedom and make both parties worse off. An example is given in Holmström (1977) and the issue has recently been studied in more detail by Green and Stokey (1980a), with the conclusion that very little can in general be said. Only a quite special form of information improvement, called success-enhancing, will guarantee that both parties get better off.[10] The purpose of this section then is to study what can be said about the other results in the proposition.

First, let us study when the optimal control interval contains the best centralized act $d^*$ as was the case in the economic planning problem. Define the *level sets*,

$$Y_i(d) = \{y \in Y \mid d_i(y) \geqslant d\}, \quad i = \mathrm{A},\mathrm{P}, \tag{25}$$

of an agent A and the principal P. $Y_{\mathrm{A}}(d)$ is the set of signal outcomes under which the agent would prefer a higher action than $d$. Obviously, $Y_i(d) \subset Y_i(d')$, when $d' \leqslant d$. The agent's and the principal's preferences (or preference functions) are said to be *coherent* if, for every $d \in D$:

$$Y_{\mathrm{A}}(d) \cap Y_{\mathrm{P}}(d) = Y_{\mathrm{A}}(d) \quad \text{or} \quad Y_{\mathrm{P}}(d).$$

Coherence expresses the requirement that if for one signal $y$ the agent prefers $d_1$ to $d_2$ and the principal $d_2$ to $d_1$, then there cannot be another signal for which the reverse is true. Coherence turns out to be an essential assumption that loosely speaking implies that the principal's and the agent's preferences move in the same direction as $y$ changes.[11] Without such conformity, delegation will frequently be valueless.

I wish to emphasize that the ensuing analysis does not require any assumptions about the dimensionality of $y$ as is commonly (and regrettably) the case in much of information economics. On the other hand, coherence is the vehicle by which our one-dimensional intuition carries over to the general case.

**Theorem 2.** Assume that preferences are concave, coherent, and continuous. Then there exists an optimal interval which contains the principal's best centralized decision.

*Proof.* Let the optimal interval control $[a, b]$ lie above the principal's best decision $d^*$. The other possibility can be treated similarly. Furthermore, assume $[a, b]$ is an optimal interval with the lowest lower bound. Such an interval exists by the continuity assumption and because the set of optimal lower bounds above $d^*$ is closed. I will show that there exists an $\epsilon > 0$ s.t. $[a\text{-}\epsilon, b]$ is at least as good as $[a, b]$, unless $a = d^*$. This will prove the claim.

Suppose $a > d^*$, contrary to the assertion in the theorem. By coherence, two possibilities exist: (i) $Y_{\mathrm{P}}(a) \subseteq Y_{\mathrm{A}}(a)$ and (ii) $Y_{\mathrm{A}}(a) \subseteq Y_{\mathrm{P}}(a)$. They are pictured in figs. 8.2

[10] The information improvement corresponding to a smaller $s_x^2$ in the example above is, however, not success-enhancing.

[11] Notice that coherence is *not* equivalent to $d_{\mathrm{A}}(y) \geqslant d_{\mathrm{P}}(y)$ or $d_{\mathrm{A}}(y) \leqslant d_{\mathrm{P}}(y)$ for all $y$, since then unrestricted responses could not intersect as they do for instance in the planning example.

and 8.3 (the heavy lines represent the agent's controlled response functions). It could be that $Y_P(a) = Y_A(a)$. I will for the moment assume this is not the case, and come back to it later.

*Case 1.* $Y_P(a) \subseteq Y_A(a)$ and $Y_P(a) \neq Y_A(a)$. Let superscript c denote complement. By coherence and continuity of preferences, there exists an $\epsilon > 0$ s.t. $d_P(y) < d_A(y)$, for $y \in Y_A^c(a\text{-}\epsilon) - Y_A(a)$. By definition, $d_P(y) \leqslant a - \epsilon$, for $y \in Y_A^c(a - \epsilon)$. By concavity of the agent's preference function,

$$a - \epsilon \leqslant d_A(y \mid [a - \epsilon, b]) \leqslant d_A(y \mid [a, b]), \quad \text{for} \quad y \in Y_A^c(a),$$

and

$$d_A(y \mid [a - \epsilon, b]) = d_A(y \mid [a, b]), \quad \text{for} \quad y \in Y_A(a).$$

It follows that under the control $[a - \epsilon, b]$ the agent's response will be pointwise as close to the principal's as under the control $[a, b]$. Since the principal's preference function is concave, $[a - \epsilon, b]$ is at least as good as $[a, b]$, contradicting the minimality of $a$.

*Case 2.* $Y_A(a) \subseteq Y_P(a)$ and $Y_A(a) \neq Y_P(a)$. In this case, when one lowers the bound to $a - \epsilon$, it is not true that the agent's response will become pointwise closer (see fig. 8.3). We have to argue differently. By continuity and coherence of preferences there exists an $\epsilon > 0$ s.t.

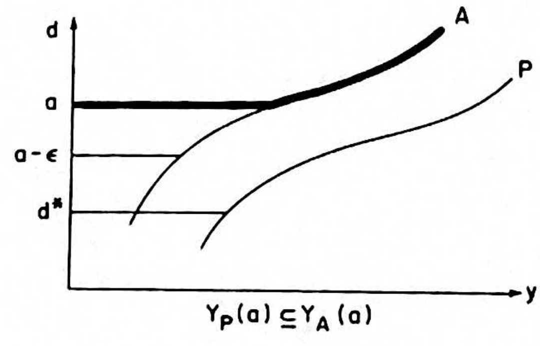


Figure 8.2

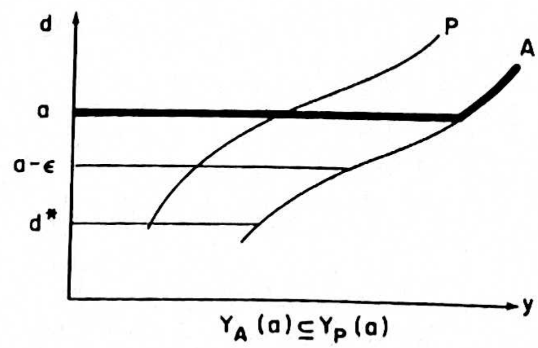


Figure 8.3

$$d_P(y) \geq d_A(y) \quad \text{for} \quad y \in Y_A(a-\epsilon) - Y_A(a). \tag{26}$$

Furthermore, $Y_A(a) \subseteq Y_P(a)$ implies

$$d_P(y) \geq a \quad \text{for} \quad y \in Y_A(a). \tag{27}$$

By concavity of the principal's preference function, he considers $d = a - \epsilon$ at least as good a constant response as $d = a$ (since the integral of a pointwise concave function is concave and $a - \epsilon$ is closer to $d^*$ than $a$). It follows that the principal prefers $d = a - \epsilon$ to $d = a$ on $Y_A^c(a)$, since he prefers $d = a$ to $d = a - \epsilon$ on $Y_A(a)$ by (27). By (26) the principal prefers (weakly) $d_A(y \mid [a-\epsilon, b])$ to $d = a - \epsilon$ on $Y_A^c(a)$, since the former is pointwise closer in this region. Combined we get that the principal finds $d_A(y \mid [a - \epsilon, b])$ at least as good as $d_A(y \mid [a, b])$ on $Y_A^c(a)$. On $Y_A(a)$ these responses coincide so we have shown that $[a - \epsilon, b]$ is at least as good as $[a, b]$, contradicting the minimality of $a$.

*Case 3.* $Y_A(a) = Y_P(a)$. By coherence of preferences and continuity we get either into case 1 or case 2 above, when we lower the bound to $a - \epsilon$, and we get a contradiction as before.

Consequently, $a > d^*$ is not possible. An analogous proof shows $b < d^*$ is not possible and the theorem is proved. ■

Next, I turn to the issue of preference comparisons. Consider two agents who have the same information, but one agent's preferences are in some sense closer to the principal. Does this imply he will be given more freedom as in the planning problem? The answer is in the affirmative with the appropriate notion of closeness.

**Definition.** Let A and A′ be two agents with the same information $\tilde{y}$. We say that agent *A's preferences are uniformly closer than agent A′'s with respect to the principal's*, if for every $d \in D_0$ = the range of $d_P(y)$:

$$Y_{A'}(d) \subseteq Y_A(d) \subseteq Y_P(d),$$

or (28)

$$Y_{A'}(d) \supseteq Y_A(d) \supseteq Y_P(d).$$

I define uniform closeness only over the set of decisions which the principal may take ($D_0$), because the optimal interval controls will always be subsets of $D_0$.

Note that pairwise coherence is implied by the definition of uniform closeness.

The following lemma will clarify the meaning of uniform closeness.

**Lemma 1.** Suppose agent A's preferences are uniformly closer than A′'s w.r.t. the principal's, and that all preference functions are strictly unimodal. Then, given any interval control $C$, we have for every $y \in Y$:

$$d_{A'}(y \mid C) \leqslant d_A(y \mid C) \leqslant d_P(y),$$

or

$$d_{A'}(y \mid C) \geqslant d_A(y \mid C) \geqslant d_P(y).$$

*Proof.* Make the contrapositive assumption, say,

$$d_A(y \mid C) < d_{A'}(y \mid C) \leqslant d_P(y) \tag{29}$$

for some $y$. The other possibility is symmetric. Let $d' = d_{A'}(y \mid C)$. By (29) and unimodality, $y \in Y_P(d')$ and $y \notin Y_A(d')$. This implies, by uniform closeness, $y \notin Y_{A'}(d')$, so $d_{A'}(y) < d'$. Since $C$ is an interval, it contains all points between $d_A(y \mid C)$ and $d'$. By strict unimodality of agent $A'$'s preferences he should then choose $d_{A'}(y \mid C) < d'$, since $d_{A'}(y) < d'$, which contradicts the definition of $d'$. ■

Notice that we have to restrict attention to interval controls. For controls with gaps the claims would be false in general. In that case we would need more information about the specific preference structure of the agents. With interval controls the relevant information is carried in the uncontrolled response functions alone.

The lemma shows that uniform closeness implies that agents A's response function always lies *between* P's and A''s. An example of uniform closeness would be a situation where agents determine the scale of an investment and all three parties have exponential utility functions with A's risk-aversion coefficient lying between P's and A''s.

It should be noted that it is possible that A is uniformly closer in preferences than A' for one information signal $\tilde{y}$, but not for another. Hence, one cannot define uniform closeness directly over $z$.

The following is the main theorem on comparison of preferences.

**Theorem 3.** Assume preferences are concave and continuous. If agent A is uniformly closer than A' in preferences w.r.t. the principal, then the principal will be no worse off with agent A than with A', and he will give A at least as much freedom as A', regardless of the distribution of $\tilde{y}$.

*Proof.* The claim that the principal is no worse off with A than A' follows directly from lemma 1, since A's controlled response function is pointwise closer than A''s, regardless of the control $C$.

For the second claim I need to show that if $[a', b']$ is an optimal control for A', then there exists an optimal control $[a, b]$ for A, which contains $[a', b']$. By lemma 1 the optimal intervals overlap, since they contain the principal's best centralized act (theorem 2). Hence, if $[a, b]$ does not contain $[a', b']$, then either $a' < a$ or $b' > b$.

Assume $b' > b$. I will show that $[a', b]$ is a strictly better control of A' than $[a', b']$, contradicting the optimality of $[a', b']$. Two cases are possible by uniform closeness.

*Case 1.* $Y_{A'}(b) \supseteq Y_A(b) \supseteq Y_P(b)$. An illustration of the situation is given in fig.

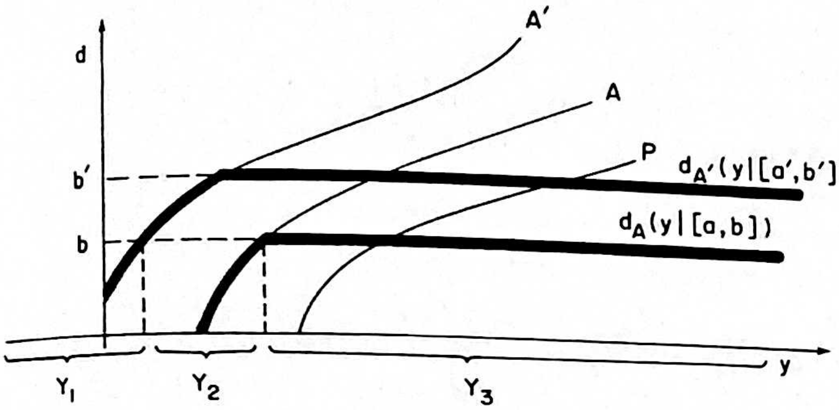


Figure 8.4

8.4, where I have written $Y_1 = Y^c_{A'}(b)$, $Y_2 = Y_{A'}(b) - Y_A(b)$, and $Y_3 = Y_A(b)$.

Let $f_i(d, y) = E(F_i(d, z) \mid y)$, $i = \mathrm{P, A, A'}$. Since $F_\mathrm{P}$ is concave, so is $f_\mathrm{P}$, and correspondingly for $f_\mathrm{A}$ and $f_{\mathrm{A}'}$.

$$\int f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b']), y)\, \mathrm{d}P(y) = \int_{Y_1} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b']), y)\, \mathrm{d}P(y) + \int_{Y_2} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b']), y)\, \mathrm{d}P(y) + \int_{Y_3} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b']), y)\, \mathrm{d}P(y). \tag{30}$$

By definition of $Y_{\mathrm{A}'}(b)$,

$$\int_{Y_1} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b']), y)\, \mathrm{d}P(y) = \int_{Y_1} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b]), y)\, \mathrm{d}P(y). \tag{31}$$

On $Y_{\mathrm{A}'}(b) - Y_\mathrm{A}(b)$, $d_{\mathrm{A}'}(y \mid [a', b']) \geqslant b$, whereas $d_\mathrm{A}(y) \leqslant b$. Consequently, by uniform closeness and using lemma 1, $d_\mathrm{P}(y) \leqslant d_\mathrm{A}(y) \leqslant b$. Since the principal's preference function is concave, he would prefer $d = b$ to $d_{\mathrm{A}'}(y \mid [a', b'])$ on $Y_{\mathrm{A}'}(b) - Y_\mathrm{A}(b)$. But on this set, $d_{\mathrm{A}'}(y \mid [a', b]) = b$, so

$$\int_{Y_2} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b']), y)\, \mathrm{d}P(y) \leqslant \int_{Y_2} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b]), y)\, \mathrm{d}P(y). \tag{32}$$

On $Y_\mathrm{A}(b)$, $d_{\mathrm{A}'}(y \mid [a', b'])$ is no better than $d_\mathrm{A}(y \mid [a, b'])$ for the principal, by lemma 1. On the other hand,

$$\int_{Y_3} f_\mathrm{P}(d_\mathrm{A}(y \mid [a, b']), y)\, \mathrm{d}P(y) < \int_{Y_3} f_\mathrm{P}(d_\mathrm{A}(y \mid [a, b]), y)\, \mathrm{d}P(y),$$

since $[a, b']$ was assumed suboptimal, and $d_\mathrm{A}(y \mid [a, b']) = d_\mathrm{A}(y \mid [a, b])$ on $Y_3^c$. Consequently,

$$\int_{Y_3} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b']), y)\, \mathrm{d}P(y) < \int_{Y_3} f_\mathrm{P}(d_{\mathrm{A}'}(y \mid [a', b]), y)\, \mathrm{d}P(y). \tag{33}$$

Combining (30)–(33):

$$\int f_P(d_{A'}(y \mid [a', b']), y)\, dP(y) < \int f_P(d_{A'}(y \mid [a', b]), y)\, dP(y).$$

This contradicts the optimality of $[a', b']$. Hence, $b' \leqslant b$ in case 1.

*Case 2.* $Y_{A'}(b) \subseteq Y_A(b) \subseteq Y_P(b)$. Proceeding analogously to case 1, one can show that $b' > b$ leads to a contradiction of the optimality of $[a', b']$. Hence, $b' \leqslant b$ also in this case.

Finally, a similar argument shows that $a' \geqslant a$, completing the proof. ■

If we call two agents *similar* provided they always take acts on the same side of the principal's response function, we have the following partial converse of theorem 3 (proof omitted):

**Theorem 4.** Assume agents are similar and the agents' and the principal's preferences are concave and pairwise coherent. If the agents cannot be ordered by uniform closeness, then there exist two distributions of $\tilde{y}$ such that for one, agent A is strictly preferred to and given as much freedom as A$'$, and for the other, the reverse holds true.

As a consequence of theorems 3 and 4 we get:

**Corollary.** Assume preferences are concave and pairwise coherent. Among similar agents, A is preferred to A$'$ for all distributions of $\tilde{y}$ if and only if A is uniformly closer than A$'$ in preferences w.r.t. the principal.

What these results show is that only a very weak partial ordering of agents (from a principal's point of view) can be obtained even in the simplest of situations. The relevance of being able to order agents lies in an extended theory of market equilibrium. Orderings will play a role for how agents get allocated among firms.

Of course, when the decision space is of higher dimension, even less complete orderings can be expected. What appears true is that agents can be ordered (for all possible information structures) only if they lie on a ray in preference space that originates from the principal's preference profile. That it is sufficient can easily be seen. Say that $F^A = \lambda F^P + (1-\lambda) F^{A'}$, $\lambda \in (0, 1)$, so that agent A's preferences are a convex combination of P and A$'$. Then P always prefers the choice of A to A$'$ because:[12]

$$E[F^{A'}(d_2, z) \mid y] \geqslant E[F^{A'}(d_1, z) \mid y],$$

$$E[F^P(d_2, z) \mid y] > E[F^P(d_1, z) \mid y], \quad \text{implies}$$

$$E[F^A(d_2, z) \mid y] = \lambda E[F^P(d_2, z) \mid y] + (1-\lambda) E[F^{A'}(d_2, z) \mid y] > E[F^A(d_1, z) \mid y].$$

12 This obviously implies that A is uniformly closer than A$'$. Thus, the convexity condition yields a less complete ordering than uniform closeness, which only has force with interval controls.

If preference profiles are linear in $d$(say, because we are considering randomized decisions), then the converse holds true; that is, A is preferred to A$'$ for all probability distributions only if A is a convex combination of P and A$'$. This follows by a standard variant of Farkas' lemma; (Ross, 1980, is to be credited for the discovery of these techniques in a somewhat more specific context).

## 6. Value of delegation

When are there gains to delegation? This is a basic question of interest because it provides an economic rationale for why a principal and an agent should cooperate. From the value of delegation one may derive a choice theoretic explanation for observed institutions. For instance, in Holmström (1982) it is shown that in a centrally planned economy, target delegation of the type contemplated for the Soviet Union (Weitzman, 1976) leads to Pareto improvements because plant managers know better their production possibilities. Similarily, in Baron and Holmström (1980) the demand for investment banking services is derived from the value of delegation.

I will provide two simple sufficient conditions for when it pays to decentralize. The first one states that if the agent takes acts both below and above the principal's best centralized act and preferences are coherent, then the principal should give the agent some freedom. The second one concerns the use of transfer payments in coaxing the agent's preferences closer to the principal's so as to make delegation beneficial.

The following lemma, though obvious, is the main principle behind delegation.

**Lemma 2.** Let $d^*$ be the best centralized act. Assume there exists a control $C \in \mathcal{N}$, containing $d^*$, and a set $Y_0 \subseteq d_A^{-1}(C)$ for which:

(i) the principal prefers weakly $d_A(y \mid C)$ to $d^*$ for every $y$;
(ii) the principal prefers strictly $d_A(y)$ to $d^*$ for $y \in Y_0$; and
(iii) Prob $\{y \in Y_0\} > 0$.

Then the control $C$ is strictly preferred to $d^*$ both by the agent and the principal.

*Proof.* Obvious.

The point is that the agent's decision reveals information about $\tilde{y}$ (sometimes perfect information) and the lemma says that if the principal, *conditional on the agent's decision*, prefers it to the best centralized decision $d^*$, then the agent should be allowed to make the decision. Notice that if there were a set of decisions among which the principal is indifferent, a Pareto improvement would not be guaranteed by delegating this set to the agent (as is the case in a world of symmetric information). Conditional on the agent's choice, the indifference would generally change; possibly so that the agent's choice always becomes worse than a particular best centralized decision. This phenomenon is akin to adverse selection.

**Theorem 5.** Assume:

(i) the expected preference functions $f_P$ and $f_A$ are unimodal, coherent, and continuous;

(ii) the distribution of $\tilde{y}$ is nonatomic;

(iii) the agent takes acts both below and above the best centralized act $d^*$; and

(iv) $d_A(y) \neq d_P(y)$ for $y \in \{y \mid d_P(y) = d^*\}$.

Then there exists an interval control $C$ such that both principal and agent are strictly better off with $C$ than with $d^*$; i.e. it is Pareto improving to delegate.

*Proof.* I will show that we can find $C$ and $Y_0$ as required in lemma 2, with $C$ an interval. Let $Y(\epsilon) = \{y \mid d^* - \epsilon \leqslant d_A(y) \leqslant d^* + \epsilon\}$. By continuity, coherence, and (iv), $\epsilon > 0$ can be chosen small enough so that either $d_P(y) > d_A(y)$ or $d_P(y) < d_A(y)$ on $Y(\epsilon)$.

Take the first case, $d_P(y) > d_A(y)$ on $Y(\epsilon)$, since the second case is symmetric. Let $Y^+(\epsilon) = \{y \in Y(\epsilon) \mid d_A(y) \geqslant d^*\}$. By (iii) and continuity, $Y^+(\epsilon)$ is of full dimension and so by (ii) $P(Y^+(\epsilon)) > 0$. By strict unimodality the principal prefers strictly $d_A(y)$ to $d^*$ on $Y^+(\epsilon)$. Take $d_1 = d^*$, $d_2 = \sup\{d_A(y) \mid y \in Y^+(\epsilon)\}$, and define $C = [d_1, d_2]$. Let $y$ by arbitrary. Three possibilities arise:

(1) $y \in Y^+(\epsilon)$. Then $d_A(y)$ is strictly preferred to $d^*$ by the principal.

(2) $d_A(y \mid C) = d^*$.

(3) $d_A(y \mid C) = d_2$.

By coherence the third case implies $d_P(y) > d_2$ and so by unimodality the principal prefers $d_2$ to $d^*$. Hence, $Y^+(\epsilon)$ and $C$ satisfy the conditions of lemma 2. ■

Let me emphasize that coherence is a crucial assumption in the theorem, as I already argued in connection with the definition. Indeed, if we say that the principal and the agent are *noncoherent* when $Y_A^c(d) \subseteq Y_P(d)$ or $Y_P^c(d) \subseteq Y_A(d)$ for every $d$, then the agent is of no value for the principal. This follows from the observation that if $C$ is an arbitrary control, then the principal prefers any constant decision in $C$ to the use of the control $C$.

It should be clear from lemma 2 that condition (ii) could be replaced by a number of variants, including some which apply for discrete distributions.

We can get a weak converse of the theorem. Since it is trivially true that a necessary condition for decentralization by intervals is that the ranges of the principal's and the agent's response functions are not disjoint, we have:

**Theorem 6.** The agent is of value (i.e. it pays to delegate) for all nonatomic distributions of $\tilde{y}$, if and only if preferences are coherent and the range of his response function contains the range of the principal's.

If transfer payments are allowed, as in the screening models, delegation will be valuable under less stringent conditions. In fact, the opportunities to cooperate are increased rather substantially, since now the response functions need not be close to each other in a metric sense. What counts is how the agent's rate of substitution between acts and compensation changes with $y$ (which unfortunately has to be assumed one-dimensional).

Let $d = (a, w)$, where $a$ is the action to be taken and $w$ is the transfer from principal to agent. Given $y$, let conditional preferences be given by $f^{\mathrm{P}}(a, w, y), f^{\mathrm{A}}(a, w, y)$.[13] Partial derivatives are denoted by subscripts.

**Theorem 7.** Assume $[\partial(f_1^{\mathrm{A}}/f_2^{\mathrm{A}})/\partial y] \cdot [\partial(f_1^{\mathrm{P}}/f_2^{\mathrm{P}})/\partial y] < 0$, and that preferences are convex (level sets convex). Then it pays to delegate using a price scheme combined with an interval control.

*Proof.* There are many cases to consider all of which are similar, so I will only focus on one of them. Assume $\partial(f_1^{\mathrm{A}}/f_2^{\mathrm{A}})/\partial y < 0$, $\partial(f_1^{\mathrm{P}}/f_2^{\mathrm{P}})/\partial y > 0$, and let $a^*$ be the best centralized act. If the agent is not used, $w = 0$ by definition. Let $y^*$ be a signal outcome such that $a^*$ maximizes $f^{\mathrm{P}}(a, 0, y^*)$. Assume $k \equiv -f_1^{\mathrm{A}}(a^*, 0, y^*)/f_2^{\mathrm{A}}(a^*, 0, y^*) < 0$ (the other cases can be treated symmetrically). Let the price schedule be $w(a) = ka$ and let the agent choose freely from the set $C = \{(a, w(a)) \mid a \geqslant a^*\}$. Because $\partial(f_1^{\mathrm{A}}/f_2^{\mathrm{A}})/\partial y < 0$, the agent will choose $a > a^*$ when $y > y^*$. Because $\partial(f_1^{\mathrm{P}}/f_2^{\mathrm{P}})/\partial y > 0$ so would the principal (note that $f_2^{\mathrm{P}} < 0$). Thus, whatever $a \geqslant a^*$ the agent chooses, the principal is better off.[14] ■

The conditions $\partial(f_1^{\mathrm{A}}/f_2^{\mathrm{A}})/\partial y < 0$ and $\partial(f_1^{\mathrm{P}}/f_2^{\mathrm{P}})/\partial y > 0$ state that for higher $y$ values the agent is willing to give up more money to take higher actions, whereas the principal is willing to accept higher actions for a lower price. It is in this sense that their preferences move in the same direction as $y$ increases and this makes delegation gains possible. The condition of theorem 7 applies to many screening problems (though not to the earlier discussed insurance example).

## 7. Concluding remarks

Part of the purpose of this chapter has been to formalize the notion of delegation, which is one of the most common forms of decentralization in organizations. Starting from a general formulation of the decentralization problem, I showed that delegation corresponds to the special case where information from different agents does not get coordinated in the decision process. Therefore, delegation could be viewed as a process where the agent is given the authority to make the final decision, but subject to constraints (the control set) set by the principal. Though perhaps not immediately apparent, a number of recent models on incentives and screening fall into the category of delegation.

Two factors influence the optimal delegation decision: the relationship between the principal's and the agent's preferences and the degree of the agent's expertice. Intuition

13 Note that $f_2^{\mathrm{P}} < 0$, since $w$ is a transfer from principal to agent.

14 The reader is urged to draw a simple graph, noting that $-f_1^{\mathrm{A}}/f_2^{\mathrm{A}}$ is the tangent of the agent's indifference curve in $(a, w)$ space.

might suggest that the agent should be delegated more authority the closer the preferences of the two parties are and the more expert information he possesses. Indeed, in the context of a much analyzed economic planning model introduced by Weitzman (1974), I found the intuition confirmed. The result suggests that with a less informed government (or more informed firms), firms should be less stringently regulated.

In general, however, without specific parameterizations, surprisingly weak comparative statics results obtain. Preferences have to be closer in a quite strong sense to guarantee an increase in the agent's freedom, even when control sets are restricted to be intervals.

In contrast, the requirements for delegation to be of any value were substantially weaker. Loosely speaking, the two conditions provided in this chapter require that preferences move in the same direction with changes in information. Furthermore, with more dimensions of exchange between the agent and the principal, these conditions are easier to meet, as illustrated by a comparison of the one-dimensional and the two-dimensional cases. Both conditions are useful in explaining the emergence of institutions based on a principal–agent relationship.

Delegation has two advantages over more complex decision processes involving information transmission between the agent and the principal. Its simple structure saves on administrative costs and so, where gains from information coordination are small, we expect delegation to be an efficient mode of decentralization. A second advantage stems from the strong form of precommitment that delegation involves. It is clear that the principal is better off when he can precommit himself to a mechanism (as modelled here) than when he cannot and therefore is expected to respond optimally against the agents' messages. It appears that delegating authority to an agent, as opposed to asking the agent for information and promising to act on the information in a particular way, is a more convincing form of precommitment (though they are formally equivalent).

In some situations it may be difficult to make precommitments or delegate authority. Crawford and Sobel (1982) have recently analyzed the model presented here under the assumption that the principal cannot precommit himself. In the ensuing equilibrium the agent will disguise his information by reporting a coarser partition, which prevents the principal from exploiting him completely. Crawford and Sobel address the same questions as here: How do changes in the agent's preferences affect cooperation? Qualitatively, their conclusions are similar. Preferences have to be sufficiently coherent for cooperation to be of value and as preferences become uniformly closer, more of the agent's information can be exploited (by a refined partition equilibrium).

Two papers by Green and Stokey (1980a, 1980b) complement the above analyses on preference changes. They look at the effects of improved information on cooperation both in the case of precommitment (i.e. delegation) and no precommitment. The reported results are analogous to the ones regarding preference changes. Generally, a more informed agent need not be more valuable for the principal. A weak ordering is obtained by considering success-enhancing improvements in information, which appears closely related to the convexity condition given in this chapter. A unified theory for preference and information changes might be conjectured on these grounds.

The three papers mentioned above and this chapter, all point to the very limited results one can obtain in general regarding preference and information changes. Illuminating as this knowledge is, it is rather negative for the prospects of a general market theory of agents, since tastes for agents have no simple ordering. Apparently, we have to be content with highly parameterized models, looking for simultaneous restrictions on preference and information structures.

## Appendix A: Existence of optimal controls

The following assumptions are made:

**Assumption A1.** $D$ is a compact subset of a complete, separable metric space.

**Assumption A2.** $\mathcal{N}$ is a closed subset of $2^D$, w.r.t. the Hausdorff-metric (see below).

**Assumption A3.** $F^{\mathrm{A}}$ and $F^{\mathrm{P}}$ are continuous and uniformly bounded.

Let $f^i(d, y) = E[F^i(d, z) \mid y]$ for $i = \mathrm{A,P}$ and all $y$. In proving existence I will take the standard approach of showing that our problem is one of finding the supremum of a continuous function over a compact set. Since the argument of the objective function is a set, we have to define a suitable metric to get the objective function continuous. For this purpose I will use the Hausdorff ($H$-) metric (see Munkres, 1975), which is defined for two sets $A, B \in 2^D$ as:

$$H(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} m(a, b), \ \sup_{b \in B} \inf_{a \in A} m(a, b) \right\},$$

where $m$ is the metric of the space containing $D$.[15] The crucial result, which gives us compactness is the following (see Munkers, 1975):

**Lemma A1.** If $D$ is a compact set in the metric $m$, then the set of all nonempty bounded closed subsets of $D$ is compact in the Hausdorff-metric $H$.

Since we assumed $\mathcal{N}$ is closed in the $H$-metric, lemma A1 tells us that $\mathcal{N}$ is compact, because $2^D$ is metrizable.

**Lemma A2.** $d(y, C)$ is u.s.c. in $C$ for every $y$.

*Proof.* Fix $y$. Let:
(i) $C_n \to C$ (in the $H$-metric), $C_n, C \in \mathcal{N}$.
(ii) $d_n \in d(y, C_n)$.
(iii) $d_n \to \hat{d}$.

15 I am indebted to David Kreps for suggesting the use of the Hausdorff-metric.

I need to show that $\hat{d} \in d(y, C)$. Since $C$ is compact, $d(y, C)$ is nonempty. Let $\bar{d} \in d(y, C)$ be arbitrary. I show that $\hat{d} \in C$ and $f^A(\hat{d}, y) = f^A(\bar{d}, y)$, implying $\hat{d} \in d(y, C)$.

Let $\delta > 0$ be arbitrary. $C_n \to C \Rightarrow \exists n_1$ and a sequence $\{d(n)\}$, $d(n) \in C\ \forall n$, s.t. $m(d(n), d_n) < \delta$ when $n \geqslant n_1 \cdot d_n \to \hat{d} \Rightarrow \exists n_2$ s.t. $m(\hat{d}, d_n) < \delta$ when $n \geqslant n_2$. From the triangle inequality we get $m(d(n), \hat{d}) < 2\delta$ when $n \geqslant n_0 = \max(n_1, n_2)$. Since $\delta$ was arbitrary, $\hat{d}$ is a limit point of a sequence in the closed set $C$ and so $\hat{d} \in C$.

$\hat{d} \in C \Rightarrow f^A(\hat{d}, y) \leqslant f^A(\bar{d}, y)$. If, moreover, $f^A(\hat{d}, y) \geqslant f^A(\bar{d}, y)$, we are done. $C_n \to C \Rightarrow \exists$ a sequence $\{d'(n)\}$, $d'(n) \in C_n$, $\forall n$, s.t. $d'(n) \to \bar{d}$. We have that $f^A(d'(n), y) \to f^A(\bar{d}, y)$, since $f^A$ is continuous. We also have $f^A(d_n, y) \to f^A(\hat{d}, y)$ by (iii). $d'(n) \in C_n$ implies $f^A(d'(n), y) \leqslant f^A(d_n, y)\ \forall n$, by (ii). Taking limits on both sides gives $f^A(\bar{d}, y) \leqslant f^A(\hat{d}, y)$, concluding the proof. ■

**Lemma A3.** If Prob $\{d(y, C)$ is not a singleton$\} = 0\ \forall C \in \mathcal{N}$, then $E\{f^P(d(y \mid C), y)\}$ is continuous in $C$ for any specific choice of response function from the response correspondence.

*Proof.* Let $C_n \to C$. Let $\Lambda = \{y \mid d(y \mid C)$ is not a singleton$\}$. By assumption $P(y \in \Lambda) = 0$.

$$E\{f^P(d(y \mid C_n), y)\} = E\{f^P(d(y \mid C_n), y); y \in \Lambda\} + E\{f^P(d(y \mid C_n), y); y \in \Lambda^c\}.$$

The first term in the RHS is 0. Let me show that $f^P(d(y \mid C_n), y) \to f^P(d(y \mid C), y)$ for $y \in \Lambda^c$. Then, since $f^P$ is continuous and bounded, the integral will converge to the desired limit by the bounded convergence theorem.

Write $d_n = d(y \mid C_n)$; $d_n \in d(y, C_n)$. I claim $d_n \to d(y \mid C)$. This is true if and only if every subsequence $\{d_{n'}\}$ has a refinement $d_{n''} \to d(y \mid C)$. Since $D$ is compact, any subsequence $\{d_{n'}\}$ has a convergent subsequence $d_{n''} \to \hat{d}$. Since $d(y, C)$ is u.s.c. by lemma A2, $\hat{d} \in d(y, C)$, which is a singleton (because $y \in \Lambda^c$) and so $\hat{d} = d(y \mid C)$. Hence, $d_n \to d(y \mid C)$. This completes the proof. ■

**Theorem 1 (ii).** Assume assumptions A1–A3 and in addition Prob $\{d(y, C)$ is not a singleton$\} = 0$, $\forall C \in \mathcal{N}$, then there exists an optimal solution $C^* \in \mathcal{N}$ to the delegation problem regardless of the maximizing response function the agent uses.

*Proof.* The delegation problem is to find a maximizing $C^* \in \mathcal{N}$, if it exists, for the function $E\{f^P(d(y \mid C), y)\}$, where $d(y \mid C)$ is a particular maximizing response function of the agent. By lemma A1 $\mathcal{N}$ is compact and by lemma A3 the objective function is continuous in $C$. This implies there exists an optimal control $C^* \in \mathcal{N}$. ■

Let me now study the particular response function $d_{\max}(y \mid C)$.

**Lemma A4.** If (i) $C_n \to C$ and (ii) $E\{f^P(d_{\max}(y \mid C_n), y)\} \to \bar{E}$, then $\bar{E} \leqslant E\{f^P(d_{\max}(y \mid C), y)\}$. In other words, $E\{f^P(d_{\max}(y \mid C), y)\}$ is u.s.c. in $C$.

*Proof.* Write $d_n = d_{\max}(y \mid C_n)$, $E_n = E\{f^{\mathrm{P}}(d_n, y)\}$:

$$\begin{aligned}\bar{E} &= \lim_{n\to\infty} E_n = \limsup E\{f^{\mathrm{P}}(d_n, y)\} \leqslant \lim_{k\to\infty} E \sup_{n \geqslant k} \{f^{\mathrm{P}}(d_k, y)\}\\ &= E\{\limsup f^{\mathrm{P}}(d_n, y)\}\end{aligned} \tag{A.1}$$

(the inequality holds for sup's and taking limits, which exist since the sequences decrease, we get the limiting inequality; the last equality follows by bounded convergence). I claim: $\limsup f^{\mathrm{P}}(d_n, y) \leqslant f^{\mathrm{P}}(d_{\max}(y \mid C), y)$ for each $y$.

It is easy to construct a subsequence $\{d_{n'}\}$ s.t. $f^{\mathrm{P}}(d_{n'}, y) \to \limsup f^{\mathrm{P}}(d_n, y)$. Since $D$ is compact there is a converging refinement $d_{n''} \to \dot{d}$. Of course, $f^{\mathrm{P}}(d_{n''}, y) \to \limsup f^{\mathrm{P}}(d_n, y)$. Since $f^{\mathrm{P}}$ is continuous, we also have $f^{\mathrm{P}}(d_{n''}, y) \to f^{\mathrm{P}}(\dot{d}, y)$. Hence, $\limsup f^{\mathrm{P}}(d_n, y) = f^{\mathrm{P}}(\dot{d}, y)$. By lemma A2, $d(y, \cdot)$ is u.s.c. for every $y$. This implies $\dot{d} \in d(y, C)$ and so $f^{\mathrm{P}}(\dot{d}, y) \leqslant f^{\mathrm{P}}(d_{\max}(y \mid C), y)$ by definition of the $d_{\max}$-function. This proves the claim and the lemma follows directly from (A.1). ■

**Theorem 1(i).** For the response function $d_{\max}(y \mid C)$, there exists an optimal control $C^* \in \mathcal{N}$.

*Proof.* The theorem is a direct consequence of the u.s.c. of $E\{f^{\mathrm{P}}(d_{\max}(y \mid C), y)\}$ and the compactness of $\mathcal{N}$. ■

## Appendix B: The hazard rate for a normal distribution

The hazard rate is defined as:

$$h(y) = \frac{g(y)}{1 - G(y)}, \quad y \sim N(0, s^2),$$

$$h'(y) = \frac{g'(y)}{1 - G(y)} + \frac{g^2(y)}{(1 - G(y))^2} = h(y)\,(h(y) - s^2 y), \tag{B.1}$$

$$h''(y) = h(y)\,[(h(y) - y)^2 + h'(y) - 1]. \tag{B.2}$$

It is well known that $h'(y) > 0$ (see Barlow and Proschan, 1975). Consequently, (21) follows from (B.1)

Let $\varphi$, $\Phi$ be the density and distribution functions of a standardized Normal distribution. Using l'Hôpital's rule twice gives:

$$\frac{\varphi(y)}{y(1 - \Phi(y))} \to 1, \text{ as } y \to \infty.$$

Hence,

$$\frac{g(y)}{y(1-G(y))}=\frac{\varphi(y/s)}{y(1-\Phi(y/s))\,s}\to\frac{1}{s^2}\quad\text{as}\quad y\to\infty, \tag{B.3}$$

which is (23).

Let $h_0$ be the hazard rate for a standardized Normal distribution. Suppose $h_0'(\bar{y}) > 1$ for some $\bar{y}$, contrary to the claim in (22).

By (B.2) $h_0''(\bar{y}) > 0$, implying $h_0'(y) > 1$ for all $y > \bar{y}$. This contradicts (B.3), since $h_0(y) > y$ for all $y$. Hence, $h_0'(y) < 1$ for all $y$. Eq. (22) follows then from the fact that $h'(y) = h_0'(y)/s^2$, since I already stated that $h'(y) > 0$.

To prove (24) we have:

$$\frac{\partial}{\partial s}[h(y)s^2]=\frac{\partial}{\partial s}[h_0(y/s)\,s]=h_0\left(\frac{y}{s}\right)+sh_0'\left(\frac{y}{s}\right)\left(\frac{-y}{s^2}\right).$$

If $y \leqslant 0$, this expression is certainly $> 0$, and when $y > 0$, we can minorize it by using (21) and (22) and conclude:

$$\frac{\partial}{\partial s}[h(y)s^2] > \frac{y}{s}+s\,\frac{-y}{s^2}=0.$$

This establishes (24).

## References

Barlow and Proschan (1975), *Statistical Theory of Reliability and Life Testing* (Holt, Rinehart and Winston, Inc).

Baron, D. and B. Holmstrom (1980), "The Investment Banking Contract for New Issues under Asymmetric Information: Delegation and the Incentive Problem", *Journal of Finance*.

Crawford, V. and J. Sobel (1982), "Strategic Information Transmission", *Econometrica* 50.

Dasgupta, P., P. Hammond and E. Maskin (1979), "The Implementation of Social Choice Rules: Some Results on Incentive Compatibility", *Review of Economic Studies* 46.

Gibbard, A. (1973), "Manipulation of Voting Schemes: A General Result", *Econometrica* 41.

Green, J. and J.J. Laffont (1979), *Incentives in Public Decision Making* (Amsterdam: North-Holland).

Green, J. and N. Stokey (1980a), "The Value of Information in the Delegation Problem", HIER DP776, Harvard University (revised September 1981).

Green, J. and Stokey, N. (1980b), "A Two-Person Game of Information Transmission", HIER DP751, Harvard University (revised December 1981).

Groves, T. (1973), "Incentives in Teams", *Econometrica*.

Harris, M. and A. Raviv (1981), "Allocation Mechanisms and the Design of Auctions", *Econometrica*.

Harris, M. and R. Townsend (1981), "Resource Allocation with Asymmetric Information", *Econometrica*.

Harsanyi, J. (1967–1968), "Games of Incomplete Information Played by Bayesian Players, Parts I-III", *Management Science* 14.

Holmström, B. (1977), "On Incentives and Control in Organizations", unpublished Ph.D. thesis, Graduate School of Business, Stanford University.

Holmström, B. (1979), "Moral Hazard and Observability", *The Bell Journal of Economics.*

Holmström, B. (1982), "The Design of Incentive Schemes and the New Soviet Incentive Model", *European Economic Review* 17.

Marschak, J. and R. Radner (1972), *Economic Theory of Teams* (Yale: Yale University Press).

Maskin, E. and J. Riley (1980), "Auctioning an Indivisible Object", DP87D, J.F. Kennedy School of Government, Harvard University.

Mirrlees, J. (1971), "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies* 38.

Mirrlees, J. (1976), "The Optimal Structure of Incentives and Authority within an Organization", *Bell Journal of Economics.*

Munkres, J. (1975), *Topology: A First Course* (Englewood Cliffs, N.J.: Prentice-Hall, Inc.).

Myerson, R. (1979), "Incentive Compatibility and the Bargaining Problem", *Econometrica* 47.

Myerson, R. (1982), "Optimal Coordination Mechanisms in the Principal–Agent Problems", *Journal of Mathematical Economics* 10.

Myerson, R. (1981), "Optimal Auction Design", *Mathematics of Operations Research* 6.

Riley, J. (1979), "Informational Equilibrium", *Econometrica* 47.

Ross, S. (1980), "Equilibrium and Agency-Inadmissible Agents in the Public Agency Problem", *American Economic Review.*

Rothschild, M. and J. Stiglitz (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *Quarterly Journal of Economics.*

Spence, M. (1973), "Job Market Signalling", *Quarterly Journal of Economics.*

Weitzman, M. (1974), "Prices vs. Quantities", *Review of Economic Studies.*

Weitzman, M. (1976), "The New Soviet Incentive Model", *The Bell Journal of Economics.*