*** Don Lavoie is Professor of Economics at the Program on Social and Organizational Learning, George Mason University. He received his PhD. in Economics in 1981 at New York University, working with Austrian economists Israel Kirzner and Ludwig Lachmann. He is best known as author of two books published: *Rivalry and Central Planning* (1985) and *National Economic Planning: What is left?* (1985), which elaborate on the Austrian school's critique of centralized economic planning, and for his work in philosophy of economics, including an edited collection of essays entitled *Economics and Hermeneutics* (Routledge, 1991), which explores the implications of Hermeneutics philosophy for the social sciences. He has also been exploring the implications of recent developments in 'object-oriented' software engineering, and co-authored [with Howard Baetjer, William Tulloh, and Richard Langlois] a study of the software industry called *Component Software: A Market Perspective on the Coming Revolution in Software Development* (Patricia Seybold Group, 1993). His most recent work is an edited collection of essays by Ludwig Lachmann, *Expectations and the Meaning of Institutions* (Routledge, 1994).

1 Mises formulated the calculation argument in his 1920 essay, 'Die Wirtschaftsrechnung im sozialistischen Gemeinwesen,' and his 1922 book Die Gemeinwirtschaft, works which were only translated into English in the mid-1930s.

2 For a comprehensive overview of the calculation debate, see Lavoie (1985a). For applications to economic policy issues, see Lavoie (1985b).

3 This interpretation of the Soviet type economy was first articulated by the philosopher, Michael Polanyi (1951).

4 The writings of hermeneutical philosophers (for a good summary see Warnke, 1987) on the notion of openness explain why it is central to the process of verbal understanding.

5 At first glance, the large difference between Japanese and European companies might seem to stem from the fact that more European companies produce high-performance sports cars and luxury cars, which might require more labour, materials, and floor space. The results in Table 9.2, however, cover only the high-volume auto producers. In a cross-national comparison of luxury car manufacturers, Womack, Jones, and Roos (1990, 88-91) discovered that Japanese companies again enjoy an advantage, assembling each car with less than half the labour hours needed in the most productive European plants.

CHAPTER 9
THE PRINCIPAL-AGENT RELATIONSHIP IN ORGANIZATIONS*

*Trond Petersen***

9.1 Introduction

Industrial sociology was one of the strongest fields in the sociology discipline during the 1950s. Its focus was primarily on the internal organization of firms: on reward systems, on social rewards between workers, on the culture of the workplace, and in particular on the relationship between formal and informal organization. A central feature of this literature was that it bordered on being empiricist: through careful ethnographies researchers gave juicy accounts of workplace behaviour, accounts that often were outstanding in their descriptive richness, but which usually lacked the theoretical ambition expected during the last decade or so (see e.g. Roy, 1952; Gouldner, 1954; Whyte, 1955; Hughes, 1958; Crozier, 1967). At present, the field of industrial sociology is weak. In the U.S. it is quite marginal, except for a resurgence among Marxists in the 1970s (see esp. Burawoy, 1979). In Britain it is somewhat stronger (see e.g. Edwards, 1986), but is not well integrated with central concerns in sociology more broadly.[1] This is unfortunate as it is a field with considerable promise, on all accounts: conceptually, empirically, and practically. It ought to hold a central place in the sociology discipline.

Currently, two fields outside sociology address many of the issues central to industrial sociology and both provide new and useful materials. One is labour history, which, like industrial sociology, often is empiricist and conceptually weak, but which instead of doing ethnographies, using participant observation and intensive interviews, investigates historical records.[2] Thompson (1963) is a well-known fore-

runner to the presently intense activity in this field. The other field is the economics of organization or the theory of the firm, which in contrast is fairly weak on empirical studies but rich in conceptual and theoretical investigations. In economics this is a recent development, of the last 15 years or so, and it has opened up the previously black box of the firm. The field has two important strands: Williamson's (1985) transaction cost economics and the principal-agent literature (e.g. Milgrom and Roberts, 1992, chs. 5–6). The latter is the focus of this essay.[3] My aim is not merely to review an important literature. It goes further, to hopefully bring into better contact the field of industrial sociology and the set of more conceptual ideas found in the principal-agent literature but that have relevance to one of the most impressive research endeavours in sociology.

A principal-agent relationship arises when a principal contracts with an agent to perform some tasks on behalf of the principal. In executing the tasks, the agent chooses an action. The action in turn has some consequences, that is, an outcome, and the outcome affects the welfare of both the principal and the agent.

Examples of such relationships abound. In a worker-employer relationship, the employer is the principal and the worker is the agent. When a house owner contracts with an estate agent to sell her house, the owner is the principal and the real-estate salesperson the agent. In a doctor-patient relationship the patient is the principal and the doctor the agent. In an insurer-insuree relationship, the insurer is the principal the insurance seeker the agent. In an owner-manager relationship, the owner (or owners) of an organization is the principal, while the manager (or managers) of the organization is the agent. Often in these relationships, the agent possesses skills and abilities that are needed for performing the tasks. The principal either lacks these, or is less effective in performing the tasks than the agent, or can spend his time more productively on other tasks.

The problem for the principal, at least in some of these cases, is first to choose an agent and second to choose a procedure for rewarding the agent. The selection procedure may rely on interviews, letters of recommendation, tests, or hearsay from other principals that have been in contact with the agent, etc. The procedure for rewarding the agent often bases the reward on the observed outcome of the agent's action, but may also just be based on the action itself, say, a flat fee that is paid to the agent for performing the relevant action. The reward is customarily money, but could be any other thing or host of things valued by the agent.

The principal-agent relationship is interesting when (a) there is some uncertainty in the way the agent's action gets transformed into the output, and (b) there is asymmetrical information, for example, the agent (e.g., a worker) observes her own action, but the principal (e.g., an employer) does not observe the action of the agent. The principal can then not be sure whether the agent acts in the principal's best interest.

In reviewing this relationship and its associated literature I take the viewpoint of a sociologist studying organizations, with an emphasis on how this literature can be

used to inform concerns in industrial sociology in particular, but also the sociology of organizations and professions more broadly. I focus on how the principal-agent literature can be used to analyze three organizational situations: (i) selection and hiring of employees, (ii) how to motivate employees once they are hired, and (iii) motivation in work groups. When relevant, I will make evaluative comments from the viewpoint of organizational sociology. The remainder of the paper is organized as follows. Section 9.2 outlines the basic ideas in the principal-agent relationship. Section 9.3 applies those ideas to analyzing the three specific substantive domains just identified. Section 9.4 discusses some empirical studies based on ideas from principal-agent theory. Section 9.5 discusses criticism. Section 9.6 concludes the paper.

9.2 Basic concepts

There are five central elements that describe the principal-agent setting.

First, agents may differ in their types. Type may refer to such things as whether the agent is careful versus careless (e.g., in insurance settings), whether the agent is industrious versus lazy, whether the agent is reliable versus unreliable, whether the agent is trustworthy versus untrustworthy, whether the agent is able versus unable. In short, it says something about the agent's capacity to perform the tasks contracted for by the principal. For example, when visiting the dentist, the patient (the principal) would like both an able and trustworthy dentist (the agent), one who can perform the work correctly and who does not perform more work than needed. Similar issues arise when contracting with car mechanics, electricians, and many other kinds of agents. Recommendations are often a means of identifying able and trustworthy agents in such settings.

Second, the agent's action influences the desired outcome of the relationship. The agent's action can be the effort a worker exerts at a job, the care with which a dentist does the work, and so forth. There are usually some costs to the agent from performing the action. For example, in an employment relationship, it is usually costly to the worker, physically or mentally, to exert effort. For a car mechanic, to carefully repair a car takes time away from performing other tasks.

Third, there are usually random factors that influence the outcome in addition to the agent's action and type. These random factors are customarily beyond the agent's and principal's control. For instance, the success of a salesperson may depend both on her type and her action, where the type may indicate the extent to which one has what it takes to be a salesperson and the action may be the sales effort. But given the type and the effort, the success will also depend on factors over which the agent has no control, namely the number of customers that arrive at the store, which in turn might depend on the weather, special sales in other stores, etc.

Fourth, there is the outcome. It depends, as already discussed, on the agent's type, her action, and the random factors beyond her control. The outcome is usually observable both to the agent and the principal. The outcome could be the number of shoes produced by a factory worker, the amount of sales of a department-store salesperson, the success of a surgical procedure, and so forth. The outcome could also have many facets, such as quantity and quality, or the quantity and quality of several relevant factors.

Fifth, there is the concept of asymmetrical information. Both parties observe the outcome. Usually, with asymmetrical information, only the agent observes her own action and type. The principal may be able to observe the action in some circumstances, but then at a cost. Sometimes the agent also observes the random factors prior to choosing her action, but the principal does not. To describe these issues of asymmetrical information, Arrow (1985) introduced the suitable distinction between the *hidden action* and the *hidden information* model. In the hidden action model, the principal does not observe the action taken by the agent. He observes only the outcome of the action. The agent observes her own action and type, and may also observe the random factor prior to choosing the action. In the hidden information model, the principal observes the agent's action, but not the random factor influencing the outcome, or in the case of agents of differing types, the type of the agent. The agent in the hidden information model, observes the random factor influencing the outcome before choosing her action, in which case the realization of the random factor can be taken into account in choosing the action. But since the principal does not observe the random factor, he cannot assess the extent to which the agent chose an action that coincided with the principal's best interests.

These five points describe the setting under which interaction between the principal and the agent takes place. They need not all be simultaneously present except that the fifth point, about asymmetrical information, always is. This description can be coupled with any assumptions about human nature. Below I outline the particular behavioural assumptions chosen in agency theory. But I stress that the essential description of the principal-agent relationship itself is found in the five features above. In the legal literature, where these terms and particular descriptions often are used, the more specific assumptions about human nature outlined below are not necessarily made.

There are two assumptions about human nature, at least in the formal mathematically oriented principal-agent literature. *First*, humans are hyperrational. That is, they are able to make very difficult deliberations and computations very fast.[4] Unlike Williamson's (e.g., 1985, p. 51) transaction cost economics framework, they are not boundedly rational.

Second, the actors behave selfishly and do so with guile. In concordance with Williamson's (1985) transaction cost economics framework, they are opportunistic. For example, if the agent observes the random factor either prior to or after choosing her action, she will willingly misrepresent or lie about its actual value if that serves her interests best. For example, if the outcome of the action is less desirable than hoped for, she may try to blame it all on the random factors over which she had no control.

The principal-agent literature analyzes what happens when actors who behave according to the two assumptions just identified face the principal-agent setting described above. This literature does not recommend that humans act according to such assumptions. It only studies what happens when they do. Furthermore, one may fully believe that the five features that characterize the principal-agent relationship capture something essential about several situations of interest without accepting the assumptions about human nature under which the analysis usually is carried out.

This apparatus can be used to discuss several things, but four are of interest here. First, it can be used to discuss the screening and selection devices used by principals before entering into contracts with agents. For example, in employment relationships the employer would like to hire workers who are industrious rather than lazy. In insurance contracts, the insurance company would like to insure people who are not particularly accident prone, or at least, try to charge higher rates for those that are.

Second, once a contract has been entered, the literature attempts to analyze what kind of reward rules should be used in the relationship. Should it be one based on the agent's action or one based on the observed outcome? When the reward is based on the action, it is to some extent based on a variable the principal does not observe in its entirety, because one rarely can fully observe what an agent actually did. What gets rewarded is whether the agent presumably did the job. For example, in doctor-patient relationships the reward to the doctor is based on her action, that is, the procedure chosen, but not on the success of the procedure nor on the care or skills with which the procedure was performed. In a lawyer-client relationship, at least in the U.S., the reward to the lawyer often depends on the outcome, whether the lawyer won and the size of any settlement if there was one.

Third, if the parties settle on an outcome-based reward, the apparatus can be used to explain or perhaps rather recommend the form the reward rule should take. Should it be based exclusively on the outcome, as in pure commission systems, or should it be based partly on the action and partly on the outcome, as when a salary plus commission is paid.

Fourth, the apparatus can be used to explain the existence of long-term contracts, such as internal labour markets in firms.

9.3 Applications of framework

Application 1: adverse selection

Job applicants usually differ in their types, for example, how good or suitable they are for a job. Employers do not observe the type of job applicants, and may hence end up hiring bad employees.

Similarly, people seeking insurance differ in their types, for example, in how accident prone they are. Insurance companies do not observe the types of those seeking insurance, and may hence insure people who are bad risks.

As a solution to the problem of selecting good employees, employers may erect screening procedures: letters of recommendation, job interviews, probationary periods, etc., in order to learn about employees. Or, in the case where high-level managerial jobs are to be filled, such as company executives and university presidents, there will be elaborate search procedures, utilizing professional search companies and involving a large number of parties.

As a solution to the problem of avoiding bad risks, insurance companies may use screening devices based on statistical discrimination and experience ratings based on past accident histories (Heimer, 1985). Young drivers are charged higher car insurance rates than older drivers as are those who have experienced more accidents in the past, etc.

In such situations, the principal must deal with two separate problems: (1) To select the right kind of agent, that is, the more desirable type, and (2) to induce that agent to choose the right action. There seems to be two main strategies the principal can use to accomplish these tasks. To focus ideas consider a case where there are only two types of agents, say, high and low producers.

The first strategy for dealing with these two problems is a two-stage procedure. First, the principal screens and selects agents, hoping to get agents that are, say, high producers. Then, having hopefully chosen such agents the principal needs to select a reward scheme by which to reward the agents in such a way that the agents will perform well. Some examples of reward rules are given in Figure 9.1 below. The principal will not know whether he in fact has chosen an agent of the right type, but can be more confident of having done so than if no screening procedure had been used.

The second strategy is a one-step procedure. There is no initial selection or screening of agents. In the one-step procedure, the principal offers each agent two reward systems, giving the reward as the function of the outcome. One of the reward systems is constructed so that all agents that are high producers will choose it. The other reward system is constructed in a manner so that all agents that are low producers will choose it. So, with offering such a menu of reward schemes no initial

screening or selection is needed, rather, the agents hopefully self-select themselves into the scheme most appropriate for them. Moreover, the agent's choice of reward scheme, together with the observation of the outcome of the agent's action, provide information about the agent's type that sometimes can be used in future interactions. All it takes for the scheme to work well is that the agents understand it and that they act in their own best interest.

An example of such one-step self-selection schemes occurs in insurance. There are often two contracts available: those with no deductibles and those with deductibles. The rates for the former are much higher than the rates for the latter. The idea is then that agents who are particularly accident prone or particularly risk averse will self-select themselves into contracts with no deductibles.

Another example of one-step self-selection schemes is in employment contracts. Let there be two variables that describe a worker's type: the degree to which she is risk averse or not and whether she is a low or high producer. There can now be a choice of compensation form, straight salary versus piece rates or some output-based reward, as illustrated in Figure 9.1. The workers that are risk averse might choose the former while those who are not the latter, or the low producers might choose the former and the high producers the latter (Lazear, 1986). The particular combination of being a low producer and risk averse most likely would lead to the choice of a straight salary, while the combination of being a high producer and risk neutral or risk loving, most likely would lead to the choice of a straight piece rate, if one were available.

In these examples, the independent or explanatory variable has been the agent's type. The dependent or endogenous variable, the arrangement that is to be explained, is the selection procedure used to hopefully select agents of the right type.

Application 2: contingent rewards and moral hazard

Once an agent has been chosen to perform a task, the next problem is to get her to perform, that is, to choose the right action, which usually is costly to the agent. This can be accomplished by tying the agent's reward to the outcome of the action or by monitoring the action. The choice between the two forms, rewarding the outcome or monitoring the action, depends on several factors. One factor is the cost and precision with which outcomes versus actions can be observed and measured. In situations where outcomes are easy to observe, measure, and assess in terms of their value, outcome-based rewards are more likely. When actions are easy and cheap to observe and measure, one might base rewards on the actions taken, rather than the outcomes of the action. Another factor is the extent to which actions impact the outcome. If variation in, say, work effort has little impact on the outcome, there is no point in basing rewards on outcomes, as the extra effort needed to produce a small increase in output may be quite large. Such work situations characterize toll

collectors at bridges, salespeople of small-ticket items, some jobs in automated factories, etc. A third factor determining the choice between the two compensation forms is the degree to which the outcome or the action consists of many facets. For example, the work of a secretary may involve typing manuscripts, taking telephone calls, arranging meetings and travels, and more. In such situations it might be hard to construct an outcome-based reward scheme, as there are so many aspects of the outcome. Instead, one might remunerate the employee with a fixed salary, but consider her for promotions and merit increases at regular intervals.

Some simple cases are now considered, cases rewarding the outcome alone, the action alone, and a mixture of the outcome and the action. In the first case, the reward to the agent depends exclusively on the outcome x, not the action. A very simple way to do this would be to construct the reward scheme as follows. The agent pays the principal a fee, say, x_0, and gets to receive the rest of the output $x-x_0$ to herself. The agent is then a so-called residual claimant. This is the way franchising is organized as well as some taxi companies, at least in the U.S. (see, e.g., Russell, 1985, p. 225). Some agricultural production is structured this way, as when farmers rent the land from landowners, as in the French 18th century *fermage* system (see Stinchcombe, 1983, pp. 179–181; Le Roy Ladurie, 1974). The arrangement provides strong incentives. The drawback with this procedure is that the agent gets to bear all the risks of the relationship, because the principal always receives the fixed fee, x_0. Partly for that reason, one might surmise, workers rarely rent the machinery or other means of production from management.[5]

Another way to organize the relationship would be to pay the agent a fixed fee and let the principal keep whatever is left. This is the typical arrangement among employees who receive a fixed salary. The problem with that strategy is that this leaves the agent with few incentives for choosing the right action, because her reward is independent of the outcome.[6] Of course, if the principal can observe the action of the agent at low cost, it might make sense to pay the agent a fixed fee (e.g., salary) and then observe the action or part of it. The relationship will be terminated if the action is not satisfactory.

Due to the problems encountered by these two schemes, no risk sharing versus no incentives, what one sometimes sees therefore is that the principal pays the agent a small fee, which is independent of output, plus a reward which depends on the output. The small fee protects the agent against some of the uncertainty, that is, it provides insurance, whereas the output-related reward provides her with incentives to work hard. Thus the reward scheme provides first so-called coinsurance between the principal and the agent and second incentives to the agent. The insurance is not total or full, as in the case when the agent receives a fixed fee independent of output. Also, the incentives are lower than in the case where the agent pays a fixed fee to the principal and gets to keep the surplus of her work herself (see Stinchcombe, 1983, p. 180).

Figure 9.1 illustrates the three arrangements. The terminology in Figure 9.1 corresponds to that of piece-rate contracts, but the reward systems illustrated are applicable to several other contexts. Output is denoted x while the reward is denoted w. Part (a) illustrates the contract where the agent's reward is independent of the output produced by the agent, where w_s is the fixed reward the agent receives. Part (b) illustrates the case where the agent's reward depends only on the output produced, where x_p denotes a quota or minimum output that must be produced before the agent starts receiving a reward, and two cases are illustrated, the quota is zero and a positive quota. The reward per unit output above the quota x_p is given by the slope of the 45-degree line. Part (c) illustrates the case where part of the reward is independent of output and another part depends on output, where w_{s+p} denotes the part of the reward that is independent of output, and x_{s+p} is the quota or minimum output that must be produced before the worker starts receiving the part of the reward dependent on output. The reward per unit output above the quota x_{s+p} is given by the slope of the 45-degree line. In a piece-rate contract, w_{s+p} would be the guaranteed baseline wage.

Examples of schemes with coinsurance and incentives occur in piece-rate work. There, one usually has a quota plus a piece-rate, as illustrated in Figure 9.1, part (c).

The piece-rate setting can be used to illustrate the so-called moral hazard problem.[7] Since the principal does not observe the action, the agent may have an incentive to shirk. The incentive to shirk arises in part because a fixed wage is received irrespective of the outcome. That problem has been abundantly illustrated in the sociology literature dealing with piece-rate systems.[8] The incentive-scheme alleviates part of this problem, but not all of it.[9] The idea is that if the principal could tell the agent exactly what to do, that is, to choose the action for the agent, then *both* the principal and the agent can be made better off. The fact that the principal cannot tell the agent what to do, he can only provide a reward scheme for the agent, causes the agent to take the action that is in her best interest, given the reward scheme. But this action is often different from the one that would have made both parties better off. This difference between the two actions, the one that the principal would like the agent to take and the one that it is in the agent's best interest to take, is due to what is called *moral hazard*. The term refers to the possibility that the agent will take the action which is in her best interest rather than the one that the principal would like the agent to take. Laffont (1989, p. 180) explains the somewhat awkward term moral hazard as follows: 'The agent could commit himself morally to pursue an action that is determined jointly with the principal. Then it is the principal's doubt concerning the morality of the agent that creates the problem.'

This framework and in particular the idea of moral hazard may explain: (i) long-term relationships: when a relationship is repeated the uncertainty might be alleviated, because the probability of getting a low (or a high) realization of the uncertain factor in every period is small and hence the principal will get a better idea of whether or not the agent performs in his interest, (ii) monitoring and supervision in

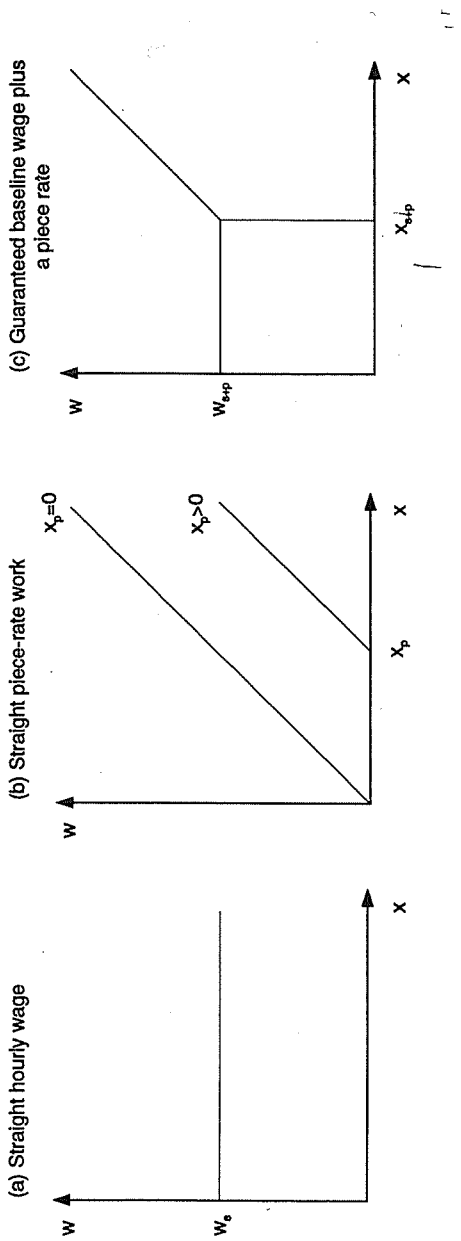

Figure 9.1: *Graphic illustration of three compensation schedules. The symbols in the figure denote the following: x is the amount of output produced and w is the wage paid as a function of the amount of output; w_s and w_{s+p} are the guaranteed wages which are paid under a straight wage and under a guaranteed wage plus a piece rate respectively; and x_p and x_{s+p} are amounts of output that must be made before the worker starts earning the piece rate, under straight piece-rate work and under a guaranteed wage plus a piece rate respectively.*

relationships, (iii) why incentive schemes are used, instead of straight fees or hourly wages, as incentive schemes provide incentives and sometimes coinsurance if there is a baseline guarantee.

In the examples above there have been several independent variables. One is the cost to the agent of expending effort, which is a feature of the agent's psychology. Another has been the relationship between effort expended and the outcome produced, which is a feature of the technology of the workplace. A third has been the degree to which the agent's action can be measured easily and at low cost. The dependent variables have been two. First, the kind of compensation scheme used, one based partly or exclusively on the outcome or one based on monitoring the action leading to the outcome. Second, it has been the amount effort or the action the compensation scheme induces.

Application 3: team work

The setting is where two workers do better by working together than alone, but each has an incentive to overstate her contribution to the joint outcome since it is difficult to verify vis-a-vis the other what she did. The problem here is that a manager usually only can observe the joint outcome, but not the individual inputs. Hence, each worker has an incentive not to work (see Alchian and Demsetz, 1972), because her individual contribution is not identifiable. This is also the case under a group piece-rate scheme, because the contribution of each worker to the group output is *1/n*, where *n* is the group size.[10] This is a situation of moral hazard. It arises not due to random factors, but due to the fact that output depends not only on own effort but also on efforts of other workers (see also Holmström, 1982). In this setting, there is still only one principal, but there are two or more agents with which he contracts.

The solution to the problem of shirking is to hire a supervisor who monitors the effort levels, fires shirkers and hires new employees. The monitor takes the surplus from the relationship between the two workers, which is her incentive to be efficient at monitoring. But this, that is, the supervisor above, is, according to Alchian and Demsetz, just capitalism and hierarchy. Capitalism is explained by the need to keep shirkers in line and by the gains from cooperation.

Holmström (1982) discussed an alternative solution to the problem of shirking in which a target-rate scheme is used, where a high reward is paid if a production target is reached and a low otherwise. This scheme can overcome free-rider problems of group piece-rate schemes under team production. Perrow's (1986, ch. 7) criticizes Alchian and Demsetz' (1972) analysis. He claims that employer's might also shirk, which may be as important, but ignored here. My evaluation is that the problem of shirking by employees is not trivial. For example, worker-owned factories sometimes call in supervisors to monitor their actions on a day-to-day basis (see, e.g., Bernstein, 1980, p. 15). Shirking by employers has been dealt with to some degrees by unions, in the sense that union representatives on the board is a

means of getting to know the financial situation of the company. This institutional arrangement is very important in Germany, called comanagement (see, e.g., Berghahn and Karsten, 1987, ch. 4). A company will always want to say that it had a bad year, and therefore wants to cut wages. However, it only wants to cut employment in bad years. Unions therefore consent to employment cuts, but resent wage cuts.

In this example of team work, the independent variables have been first the agents' costs of expending effort and second the costs or even inability of the principal to observe or meter each agent's contribution to the team output. The dependent variables have been two. First, the compensation form, one based on a straight hourly pay coupled with supervision of work effort, or one based on the team output, either a group piece rate, which is susceptible to free-riders, or one based on a group target rate, which may solve these. Second, given the compensation form, the dependent variable has been the amount of effort expended.

9.4 Empirical studies

Empirical investigation 1: adverse selection

The most famous aspect of Japanese management is that of life-time employment. About 30% of the male workforce is employed in firms practicing so-called Japanese management. According to Aoki (1988, ch. 3, esp. p. 65) and to some extent to Abegglen and Stalk (1985, ch. 9), firms screen employees quite carefully at the time of initial employment. Once offered a job, there is an expectation that the employee will remain with the firm for the rest of his (usually men) career. Aoki (1988, p. 65) claims that in the initial career stage a substantial portion of the employees quit for other jobs, hoping to obtain a better match. Among those who survive the initial career stage, some will do very well and climb at a steady pace within the company, others will do less well. At mid-career, say, ages 30–40, those who do less well will often quit, partly due to pressures on them to do so, whereas those who do well will stay on to the end of their career. Retirement is at age 55. This screening of high-performing employees and weeding-out of low performers occur through differentiation in promotion rates.

There are high costs to mid-career separation. The first of the costs is unique to Japanese firms. A separation payment is paid at the time one leaves. Its size depends on seniority. At the time of retirement it will be quite high if one stays with the same firm. For those who quit at mid-career, the separation payment will be low, and they will not be able to build up sufficient seniority with the next employer, so the separation payment will be low there at the time of retirement. The second cost is that the position one can expect to get with the next employer typically is lower than the position with the current employer.

In conclusion, the companies practicing so-called Japanese management seem to have a very effective way of screening and retaining high producers. It does take time, however, and seems to be dependent on relatively low turnover rates.[11] The careful screening of employees at time of hire helps increase the quality of those hired initially. The retention procedures once hired facilitate keeping the best employees. These two procedures also may lead to some self-selection among employees: Only the most able will apply for jobs in the sector practicing the so-called Japanese employment system. Less able and less industrious workers know or suspect that even if hired, they will not be able to make it at mid-career and will then face the embarrassment of not getting ahead.

Empirical investigation 2: compensation design

Eisenhardt (1985) studied the choice of compensation form, salary versus commission, in 54 specialty retail stores in the San Francisco area (California). She relates the choice of compensation system to several factors: the sales price of the items sold, the extent to which a job is programmable, that is, whether it is easily definable, the extent to which work behaviour (or work input) can be easily metered, the cost of metering outcome, the uncertainty of the outcome, and the sex composition of labour force in the store.

She finds that the use of salaries rather than commissions increase in each of the variables. This means that reliance on a salary rather than a commission, increases the more programmable a job is, the easier it is to meter behaviour, the costlier it is to meter outcomes, and the more uncertain the outcome is. This is consistent with what one should expect from an agency-theoretic perspective.[12]

Some of the measures used in the study are rather indirect. This is the case with the measure of outcome uncertainty, which obtained from industry characteristics rather than from individual stores. The cost of metering of outcome is also measured rather indirectly, by (1) the number of stores in the retail chain, and (2) the ownership type of the store, private or family owned versus public.

Eisenhardt (1988) updates the earlier study (Eisenhardt, 1985). As in the 1985 study, she investigates the choice between salary and commission, using data on the same 54 specialty retail stores. She juxtaposes explanations from the agency perspective and institutional theory in sociology (e.g., Meyer and Rowan, 1977).

According to the agency perspective, the easier it is to meter or observe behaviour, that is, labour inputs, the more likely it is that a straight salary is used. Conversely, the harder it is to meter behaviour, the more reliance will be placed on output-related rewards, that is, commissions. Similarly, the lower the span of control, that is, the lower the ratio of workers to supervisors, the more reliance there should be on salaries because behaviours are then better observed. Stores selling big-ticket items should rely more on commissions because these items are harder to sell. Two additional hypotheses of this type were developed from the agency perspective.

From the institutional perspective it was predicted that the older the store the higher the reliance on commissions. The reason for this is that commissions were more common, say, 40 years ago than today. According to the institutional perspective, firms to some degree tend to reflect the legitimate environment at the time of their founding (see Stinchcombe, 1965). Hence, younger stores rely less on commission and older stores more.

In the empirical analysis all hypotheses were supported. I find the theoretically most relevant finding to be that stores where behaviours are easy to observe or meter rely more on salaries. This relates to the choice of control system: output versus input based, that is, direct incentives (e.g., commissions) versus supervision.

I make one comment on the hypothesis and finding that the lower the span of control the higher the reliance on salaries. This is presented as a 'causal' hypothesis. However, I think the following is a more correct analysis. A store faces a choice between an output-based control system such as commissions, in which case the span of control will be large, or an input-based control system based on salary and supervision of behaviour and work effort, in which case the span of control is low. What happens here is that a store chooses simultaneously the reward system and the span of control: either an input-based reward system, which is a salary and considerable supervision of effort, or a commission and little supervision of effort, a large span of control. The sequence is not that the span of control determines the reward system. Both are chosen jointly.

Petersen (1992a) studied a sample of 63,000 salespersons in 178 U.S. department stores. Three types of compensation systems were addressed: straight salary, straight commission, and salary plus commission. It is argued that straight commission is, from the employee's viewpoint, a more risky payment system than salary plus commission. If that is the case, an agency-theoretic perspective would predict that straight commission workers on the average would earn more than salary-plus-commission workers. This is so because the former are compensated in average wages for the extra risks they face.

An alternative interpretation developed in Petersen (1992a), drawing on historical research on payment systems, suggests that the two schemes need not primarily reflect different risk-sharing arrangements. Instead, they may reflect the outcome of bargaining situations where those obtaining a salary-plus-commission system rather than a straight-commission system, do so due to their stronger bargaining power. Historically, workers have always fought for the existence and size of baseline wages, while employers have fought against these. Even up to this day, employers in retail trade prefer straight commission systems to straight salaries or salary-plus-commission, whereas employees tend to prefer one of the two commission systems to a straight salary (see Petersen, 1992a, sec. 2). If this bargaining interpretation is correct, then one should expect that salary-plus-commission workers on the average earn more than straight commission workers. Furthermore, one should not find both systems used side-by-side in the same occupation and firm. This will be feasible only when they reflect the outcomes of differing risk-sharing arrangements, not differing bargaining strengths.

Three central findings are reported in the empirical analysis (Petersen, 1992a, sec. 4). First, salary-plus-commission workers on the average earn more than straight commission workers in most occupations, when one makes no distinction between intra- and inter-establishment differences. Second, at the establishment level, when both systems are used in the same occupation, straight commission workers on the average earn more than salary-plus-commission workers. So, the bargaining interpretation seems to have most predictive power in explaining differences between establishments, whereas the risk-sharing interpretation is more successful in explaining within-establishment differences. Third, the two payment systems, straight commission and salary plus commission, are rarely used side-by-side within the same establishment and occupation, namely in less than three percent of the 952 possible establishment-occupation matches. This means that the intra-establishment differences to be explained are small in comparison to the between establishment differences. The observed pattern of mixing the two schemes is more consistent with a bargaining than a risk-sharing interpretation. So, even though the risk-sharing interpretation accounts well for the observed intra-establishment differences in wages, such intra-establishment differences account for a very small portion of the wage differences found in the industry. The upshot of all this is that the agency literature may have placed too much weight on the risk-sharing aspects of output-related employment contracts, and perhaps too little weight on other issues.[13]

Empirical investigation 3: team work

Petersen (1992b) studied reward systems in work groups. It is well known that group-based piece-rate systems are susceptible to free-rider problems. Such problems may, however, be overcome by the construction of appropriate target rates, where a high wage is received by every team member if a production target is reached and otherwise a low wage (Holmström, 1982). So, a pure market-type relationship, rather than one relying on a supervisor, as proposed by Alchian and Demsetz (1972), may solve the free-rider problem.

Petersen (1992c) discussed some alternative solutions to the free-rider problem under group piece-rate systems, in particular those relying on social rewards, altruism, and moral commitment. Social reward mechanisms are effective in curtailing shirkers, but are, from the workers' viewpoint, also costly: someone must hand out the sanctions. Altruism can fully alleviate the problem only when there is no conflict whatsoever between the interests of an individual worker and the collective of workers, that is, when each worker puts as much weight on the welfare of the other workers as on her own, a situation not likely to occur. Moral commitment is a very efficient solution. Its main problem is that little is known about how to bring it about.

In an empirical analysis of production workers in the U.S. wood household furniture industry and the nonferrous foundries industry it was shown that workers paid a group target rate on the average earn higher wages than those paid a group piece rate. This finding is consistent with the claim that group target rates can solve free-rider problems whereas group piece rates are quite susceptible to these.

9.5 Criticism and extensions

Criticism

The most well-known critique of agency theory in sociology is that of Perrow (1986, ch. 7). He raises two points. Below, I discuss each in turn. Then I make my own comments on Perrow's criticisms.

His first criticism amounts to the claim that the problems of moral hazard and adverse selections are not empirically important. With respect to adverse selection, employers look for satisfactory employees, but not necessarily the best. Presumably, according to Perrow, satisfactory employees are easy to come by.

With respect to moral hazard, Perrow claims that this is only important in situations where (i) the effect of effort on output is big, and (ii) the random factor is very uncertain. According to Perrow, then, not too many situations possess these two features. Furthermore, Perrow claims that bureaucratic organizations make moral hazard and adverse selection empirically unimportant. But, note that agency theory purports to explain bureaucratic organizations as one means of solving moral hazard and adverse selection, because they alleviate the problems when repeated interaction occurs. This point was brought up in the discussion of the Japanese firm in Section 9.4 above.

His second criticism is that the assumption of selfish or self-regarding behaviour used in economic theory is not correct. The degree of self-regarding versus other-regarding behaviour depends on the structure of the situation, that is, human nature is malleable and can be manipulated by the organization. In particular, self-regarding (i.e., selfish) behaviour is more likely when (i) the interaction is not repeated, (ii) individual reward is stressed as opposed to group reward, (iii) individual performance is measurable, (iv) few interdependencies between workers, (v) authority is preferred, and (vi) hierarchies are tall.

My assessment of Perrow's criticisms is first that with respect to the importance of the phenomena, I think Perrow is wrong. I believe that adverse selection is quite important. Companies rely on letters of recommendation, interviews, probationary periods, search and evaluation committees, etc. Even if all that is sought is a satisfactory employee, finding one need not be easy. It is easier than finding a perfect employee, but not guaranteed. I also believe that moral hazard is important. About

25% of production workers in the U.S. manufacturing industries are paid on a piece-rate scheme with some quota, thereby providing insurance and incentives (see, e.g., Petersen, 1992b, Table 1). This is not trivial. All executive compensation has severe elements or moral hazard. The issues there are to align the interests of the executive to those of stock-holders, how to weigh long-term versus short-term interests, etc.

Turning then to his criticism of the assumptions about human nature, Perrow no doubt has a valid point. Nevertheless selfish behaviour is clearly important and needs careful analysis. Perrow's conditions for when selfish behaviour is likely to be widespread pertain to workers or work groups, as, for example, the condition of repeated interaction, but not necessarily to the relationship between workers and employers. It is the latter relationship that agency theory primarily treats. Petersen (1992b) discusses how selfish behaviour that leads to severe free-rider problems may be alleviated in situations where interaction is repeated, as in work groups with low turnover rates. Some evidence, albeit circumstantial, was offered to this effect (Petersen, 1992b, sec. 6).

Extensions

To conclude this section I add some reflections of my own, issues that also have been discussed in part in Nalbantian (1987). First, I believe that agency theory to date has suffered somewhat from relying on models that are narrow in terms of (1) the conceptions of inputs or actions the agent can take, and (2) the nature of the rewards the agent reaps from the relationship.

Usually, in an agency model, the agent chooses a single action variable, for example, effort. But it is rarely this simple in real-life settings. Even in industrial settings, such as shoe production, the agent's action consists of at least two aspects: the effort for quantity and the care for quality. In other settings, the action usually consists of an array of activities. For example, in teaching in schools or universities, teachers must impart both basic skills, that can be evaluated in tests, and elements of deep thinking and culture, which are not easy to test whether they have been acquired or not. When it is taken into account that the action usually has many aspects, and the attention to more than one aspect is beneficial for an organization, then the recommendation often made in agency theory to use output-related rewards need no longer make sense. Instead, it seems that career incentives where employees are evaluated for promotion and merit raises at infrequent intervals is a better means of motivating employees for choosing actions that are in the organization's interest (see Stinchcombe, 1983, p. 182).

Similarly, with respect to the nature of the reward, this has usually been money. But in most situations the goals of people are broader than acquiring money. No doubt money is important, perhaps the most important motivator in employment relationships, although that is disputed by some psychologists. But it is by no means the only motivator. As has been documented by a long series of sociological studies,

social rewards among workers are important for regulating workplace behaviour, the quest for social status may be important, and so forth (see, e.g., Roy, 1952; Whyte, 1955; Burawoy, 1979; Edwards, 1986).[14]

The point about narrow conceptions of the actions that can be taken and the rewards agents care about is not inconsistent with an agency-theoretic perspective. However, incorporating these factors into the formal models, of which there has been a very large number, might well lead to conclusions that are quite different from those obtained to date. I venture the guess that these conclusions will be more in line with observed practices than the conclusions derived from the narrower framework. Some such efforts are in progress (Holmström and Milgrom, 1991).

Second, I believe the agency framework would be strengthened by taking the notion of bounded rationality seriously. The problem with the latter is that it is very difficult to model formally, and a formal model is often desirable, as it may give some precise insights. The problem with the precise insights from many agency models is, however, that they are insights that sometimes do not seem to be worth having. Very complicated formulae are derived for the optimal contract between a principal and an agent. Unfortunately, few principals or agents would contemplate taking the advice offered by the agency-theoretic models. It is just too complicated. Few principals are applied mathematicians, and even if they were, their agents would not be, and for a contract to work, it has to be understood also by the agent. As a delegate to a union congress in the Soviet Union in 1927 commented on wall-charts in one of the city's mills displaying the piece-rate valuation system, 'to the workers it's all Chinese' (see Ward, 1990, p. 171). Needless to say, this all boils down to the fact that neither principals nor agents are hyperrational. Taking into account the fact of bounded rationality, with its present unfortunate consequences for formal modelling, might move theorizing within the agency framework toward more realistic conclusions.

But in spite of these weaknesses, my overall assessment is that the agency framework is an unusually rich and relevant framework for studying behaviour in organizations. It alerts one to some of the core problems in organizations, such as adverse selection, incentives, and risk-sharing. However, I believe it could improve greatly as a tool for understanding behaviour in organizations by paying more attention to the two points discussed above. I believe the first point can and will be incorporated with a few years of research. The second point may prove more difficult to accommodate, but perhaps is more important.

9.6 Concluding remarks

Agency theory has much to offer in the study of organizations. It illuminates, albeit in rather simple ways, although the formal modelling may be forbidding, selection of agents (e.g., employees), motivation devices, output-based versus input-based

reward systems, direct incentives, such as piece-rates, versus indirect incentives, such as bureaucratic career incentives, risk-sharing, and more. It is a framework that appears not to have generated as large an empirical literature as the other major strand of the economics of organization, Williamson's (e.g., 1985) transaction cost framework. This may be due to the fact that theorizing in the agency-theoretic literature tends to be formal with use of complicated mathematics, and is hence not easily accessible. But there should be ample opportunities for empirical work using this framework. There are also some obvious differences in foci between principal-agent theory and the transaction cost economics (TCE) framework. In TCE the central dependent variable is the governance structure for transactions. That framework has been most succesful for explaining exchanges of physical goods, such as automobile parts, computer chips, and so forth. But it applies also to the exchange of services, that is, to the governance of personnel and its associated motivational problems. The principal-agent literature also explains governance structures for transactions, but it explains, in contrast, almost exclusively the exchange of services, that is, it deals with how to contract for labour. The focus is therefore almost exclusively on selection and motivation of agents, not on how to organize exchange of physical goods.[15]

The framework can be used to study the design of compensation contracts, a topic on which there is a large sociological literature (see, e.g., Edwards, 1986). It can also be used to analyze bureaucratic career systems, also a topic on which there is a large sociological literature (e.g., DiPrete, 1989). One of the areas to which it probably can most fruitfully be applied is the study of professions. The relationship between a client (principal) and a professional person (agent) is characterized by asymmetric information, as stressed by Parsons ([1939] 1954a), and often there is uncertainty with respect to the outcome of the professional treatment. The sociology literature (e.g., Parsons, [1939] 1954b; Abbott, 1983) has described the use of ethical codes that hopefully would inculcate the agents not to take unfair advantage of the principals, whereas the agency literature has studied the design of compensation contracts that align the interests of the two parties. In most situations a bit of both would do good, and, certainly, to rely exclusively on ethical codes to regulate such relationships would be a forlorn hope for institutional design, so a sensible analysis aimed at more normative statements would have to take the agency literature into account.

Appendix: A 'case' study of piece-rate contracts

Here, I present a 'case' study of piece-rate contracts. I first provide a brief description of a piece-rate contract and next the agency-theoretic analysis of this contract.[16] Then I go on to discuss how I believe the agency-theoretic analysis comes up with predictions, *not* prescriptions, that are out-of-line with observed practices. I describe what those practices are and I propose a couple of explanations for the practices, explanations that deal with fairness of arrangements and trade among agents.

In the agency literature one would analyze the piece-rate scheme in Figure 9.1 and derive predictions for how to choose the baseline wage $w_{s+p,j}$ and the piece rate β_j in, say, occupation or at task j. I focus on the piece rate β_j. The agency perspective states that the piece rate depends on several factors. It increases with the price of output. In many cases that is of little relevance, since there often is no price for the output as when an intermediary product is produced. Second, the piece rate decreases with the degree to which the agent finds it costly to expend effort, which is part of the agent's psychology or utility function. Third, the piece rate increases when it becomes easier to produce the product, referring to the rate at which effort is translated into output, which is part of the production technology.

I will focus only on the last prediction of the agency-theoretic framework: that the piece rate increases with the ease with which effort gets transformed into output. To discuss this aspect, I introduce the parameter α_j which denotes the ease with which effort gets translated into output in occupation or task j. It says how many units of output are on average produced per unit of effort. A high α_j indicates that it is easy to transform effort into outputs.

The last prediction is directly opposite to virtually all reports that exist about the setting of piece rates. Most well known is what happens when technological changes occur, that is, when α_j increases. In those cases the piece rate almost invariably decreases. It is usually even made explicit in union contracts or other employment contracts that the piece rate can be adjusted downward when α_j increases. This is exactly the opposite of the prediction from agency theory.

For example, in the British coal-mining industry it is part of the contract regulating the employment relationship that a piece rate can only be renegotiated if 'a persistent and continuing change in conditions' is encountered (Edwards and Heery, 1989, p. 84). In the U.S. steel industry similar provisions hold: 'The agreement should contain a provision that existing incentives should remain unchanged as long as the jobs or operations to which they apply remain unaltered' (Stieber, 1959, p. 194). In both cases, adjusting the piece rate meant adjusting it upward when it became harder to produce the output and downward when it became easier to produce the output. The piece rate β_j in occupation or task j moves in the opposite direction of the difficulty α_j with which output x_j is produced in occupation or task j.

So, what goes on here? I submit that there are two central processes operating. The first relates to fairness issues of the reward schemes. The idea is that the piece rate β_j usually is set so as to yield the same expected compensation per unit effort across tasks or occupations. The expected compensation per unit effort is given by the product of β_j and α_j. In that product, α_j gives the expected output per unit effort and β_j gives the compensation per unit of output. Hence, the product $\beta_j\alpha_j$, which gives the expected compensation per unit effort, gets equalized across tasks and occupations. To quote from the British coal-mining industry, 'Fairness, for instance, was meant to be achieved by basing incentive earnings on performance against supposedly objective measurements of effort. Pay was not linked directly to productivity or output, the primary concerns of management, as this would favor miners in those pits with the best geological conditions' (Edwards and Heery, 1989, pp. 45–46), with similar concerns prevalent a hundred years ago: '..., both parties would nominally agree that their object is to fix such a price per ton as would give roughly the normal earnings current in the district, ...' (Rowe, 1923, pp. 50–51). The story for the U.S. steel industry is the same (see, e.g., Stieber, 1989, ch. 11). Many more examples can be added (e.g., Bensman, 1985, pp. 81–85, 161; Gartman, 1986, p. 215). To sum up, fairness issues drive rates to be set so as to equalize the expected return per unit effort $\beta_j\alpha_j$ across tasks. One of the most interesting arrangements deviced to achieve such fairness when rates were not always fair, was found in the British coal mining industry in the early 1800's. Jobs within a mine were allocated between workers by a lottery system every two or four weeks (Jaffe, 1991, pp. 84, 99, 117, 119). This had the effect of 'fairly allocating easy and difficult working position,' or to 'equalize chances of working difficult places' and therefore 'to deal with inequalities of employment.'

The second thing that goes on is related to the first. In most settings relying on output-based compensation there is scope for trade among workers. The trades take the form of social rewards, of helping each other with tasks, sometimes sabotaging the tasks of others, and so forth.[17] Now, if $\beta_j\alpha_j$ is not set so as to equalize the expected compensation per unit effort across tasks, it is opened up for a shop-floor internal market among workers in trading services (e.g., Burawoy, 1979, p. 55). Consider two workers, in occupations j and g with expected compensations per unit effort of $\beta_j\alpha_j$, and $\beta_g\alpha_g$ respectively, and consider the case where $\beta_j\alpha_j$ is greater than $\beta_g\alpha_g$, as it according to agency theory would be if α_j is greater than α_g. If this is the case, it might pay for the worker in occupation g to abolish her tasks, just receive the baseline wage $w_{s+p,g}$, and instead help the worker in occupation j produce her output. The reason is that the return per unit effort is higher in occupation j. The worker in that occupation will in turn share some of her extra earnings with the worker who provides the help. The way this functions is that the worker providing the help usually assists in controlling the random factors ε_{ij} at the point of production in occupation or task j, that is, keeping these at their minimum.

The possibilities of such shop-floor trades can never be totally ruled out, but a system that encourages these, such as following the recommendations from agency theory would, is susceptible to become unruly. Moreover, it is likely to become an unfair system, because some workers will usually be in a better position to engage in such trades than other workers. So, I submit that the need to keep informal shop-floor trade under control, also necessitates equalizing $\beta_j \alpha_j$ across tasks and occupations.

References

Abbott, Andrew (1983), 'Professional Ethics', *American Journal of Sociology* 88(5), pp. 855–885.

Abegglen, James C., and George Stalk, Jr. (1985), *Kaisha. The Japanese Corporation,* Basic Books, New York.

Alchian, Armen A., and Harold Demsetz (1972), 'Production, Information Costs, and Economic Organization', *American Economic Review* 62(5), pp. 777–795.

Anderson, Erin (1985), 'The Salesperson as Outside Agent: A Transaction Cost Analyses', *Marketing Science* 4(3), pp. 234–254.

Aoki, Masahiko (1988), *Information, Incentives, and Bargaining in the Japanese Economy,* Cambridge University Press, New York.

Arrow, Kenneth J. (1985), 'The Economics of Agency', pp. 37–51 (ch. 2) in John W. Pratt and Richard J. Zeckhauser (Eds.), *Principals and Agents: The Structure of Business,* Harvard Business School Press, MA: Boston.

Baron, James N. (1988), 'The Employment Relation as a Social Relation', *Journal of the Japanese and International Economies* 2, pp. 492–525.

Bearle, A. and G. Means (1932), *The Modern Corporation and Private Property,* Macmillan, New York.

Bensman, David (1985), *The Practice of Solidarity. American Hat Finishers in the Nineteenth Century,* University of Illinois Press , Urbana, IL.

Berghahn, Volker, and Detlev Karsten (1987), *Industrial Relations in West Germany,* Berg Publishers Limited, Oxford.

Bernstein, Paul (1980), *Workplace Democratization: Internal Dynamics,* Transaction Books, New Brunswick, NJ.

Burawoy, Michael (1979), *Manufacturing Consent,* University of Chicago Press, Chicago.

Cohen, Lizabeth (1990), *Making a New Deal. Industrial Workers in Chicago, 1919–1939,* Cambridge University Press, New York.

Crozier, Michel (1967), *The Bureaucratic Phenomenon,* University of Chicago Press, Chicago, Il.

DiPrete, Thomas L. (1989), *The Bureaucratic Labor Market: The Case of the Federal Civil Service,* Plenum, New York.

Edwards, Christine, and Edmund Heery (1989), *Management Control and Union Power. A Study of Labour Relations in Coal-Mining,* Oxford University Press, Oxford, U.K..

Edwards, P.K. (1986), *Conflict at Work,* Oxford University Press, Oxford, U.K..

Eisenhardt, Kathleen M. (1985), 'Control: Organizational and Economic Approaches', *Management Science* 31(2), pp. 134–149.

Eisenhardt, Kathleen M. (1988), 'Agency- and Institutional-Theory Explanations: The Case of Retail Sales Compensation', *Academy of Management Journal* 31(3), pp. 488–511.

Eisenhardt, Kathleen M. (1989), 'Agency Theory: An Assessment and Review', *Academy of Management Review* 14(1), pp. 57–74.

Elbaum, Bernard (1984), 'The Making and Shaping of Job and Pay Structures in the Iron and Steel Industry', chapter 3 (pp. 71–107) in Paul Osterman (ed.), *Internal Labor Markets,* MIT Press, Cambridge, MA.

Fama, E., and Michael Jensen (1983), 'Separation of Ownership and Control', *Journal of Law and Economics* 26, pp. 301–325.

Gartman, David (1986), *Auto Slavery. The Labor Process in American Automobile Industry, 1897–1950,* Rutgers University Press, New Brunswick, NJ.

Gouldner, Alvin (1954), *Patterns of Industrial Bureaucracy,* The Free Press, Glencoe, NY.

Heimer, Carol A. (1985), *Reactive Risk and Rational Action. Managing Moral Hazard in Insurance Contracts,* University of California Press, Berkeley, CA.

Holmström, Bengt (1982), 'Moral Hazard in Teams', *Bell Journal of Economics* 3(2), pp. 324–40.

Holmström, Bengt, and Paul Milgrom (1991), 'Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design', *Journal of Law, Economics, and Organization* 7 (Special Issue), pp. 24–52.

Hughes, Everett C. (1958), *Men and Their Work,* The Free Press, Glencoe, NY.

Jaffe, James A. (1991), *The Struggle for Market Power. Industrial Relations in the British Coal Industry, 1800–1840,* Cambridge University Press, New York.

Jensen, Michael, and W. Meckling (1976), 'Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure', *Journal of Financial Economics* 3, pp. 305–360.

Laffont, Jean-Jacques (1989), *The Economics of Uncertainty and Information,* MIT Press, Cambridge, MA.

Lazear, Edward P. (1986), 'Salaries and Piece Rates', *Journal of Business* 59(3), pp. 406–31.

Le Roy Ladurie, Emmanuel (1974), *The Peasants of Languedoc,* University of Illinois Press, Urbana, IL.

Lysgaard, Sverre (1961), *Arbeiderkollektivet,* Universitetsforlaget, Oslo.

Meyer, John W., and Brian Rowan (1977), 'Institutionalized Organizations: Formal Structure as Myth and Ceremony', *American Journal of Sociology* 83(2), pp. 340–363.

Milgrom, Paul, and John Roberts (1992), *Economics, Organization, and Management,* Prentice Hall Englewood, Cliffs, NJ.

Moe, Terry (1984), 'The New Economics of Organization', *American Journal of Political Science* 28(4), pp. 739–777.

Nalbantian, Haig R. (1987), 'Incentive Compensation in Perspective', pp. 3–43 (ch. 1) in Nalbantian, Haig R. (Eds.), *Incentives, Cooperation, and Risk Sharing. Economic and Psychological Perspectives on Employment Contracts,* Rowman and Littlefield, Totowa, NJ.

Parsons, Talcott [1939] (1954a), 'The Professions and Social Structure', pp. 34–49 in Talcott Parsons, *Essays in Sociological Theory,* Free Press, New York.

Parsons, Talcott [1939] (1954b), 'A Sociologist Looks at the Legal Profession', pp. 370–85 in Talcott Parsons, *Essays in Sociological Theory,* Free Press, New York.

Perrow, Charles (3rd ed) (1986), *Complex Organizations. A Critical Essay*, Random House, New York.

Petersen, Trond (1991), 'Reward Systems and the Distribution of Wages', *Journal of Law, Economics, and Organization* 7 (Special Issue), pp. 130–158.

Petersen, Trond (1992a), 'Payment Systems and the Structure of Inequality: Conceptual Issues and an Analysis of Salespersons in Department Stores', *American Journal of Sociology* 98(1), pp. 67–104.

Petersen, Trond (1992b), 'Individual, Collective, and Systems Rationality in Work Groups: Dilemmas and Market-type Solutions', *American Journal of Sociology* 98(3), pp. 469–510.

Petersen, Trond (1992c), 'Individual, Collective, and Systems Rationality in Work Groups: Dilemmas and Nonmarket Solutions', *Rationality and Society* 4(3), pp. 332–55.

Petersen, Trond (1993), 'The Economics of Organization: The Principal-Agent Relationship', *Acta Sociologica* 36(3), pp. 277–293.

Petersen, Trond (1994), 'Transaction Cost Economics', chapter 2 in Pål Foss (ed.), *Economic Approaches to Organizations and Institutions: An Introduction*, Dartmouth Publishing Company, Aldershot, UK.

Randall, Adrian (1991), *Before the Luddites. Custom, Community and Machinery in the English Woolen Industry 1776–1809,* Cambridge University Press, New York.

Rowe, J. W. F. (1923), *Wages in the Coal Industry*, P. S. King and Son, London.

Roy, Donald (1952), 'Quota Restriction and Goldbricking in a Machine Shop', *American Journal of Sociology* 57(5), pp. 427–442.

Russell, Raymond (1985), 'Employee Ownership and Internal Governance', *Journal of Economic Behavior and Organization* 6(3), pp. 217–241.

Shelton, John (1967), 'Allocative Efficiency v. "X-Efficiency": Comment', *American Economic Review* 57(5), pp. 1252–58.

Stark, David (1990), 'Work, Worth, and Distributive Justice', Program on Central and Eastern Europe, Working Paper Series, #5, Harvard University, Cambridge, MA. Translated into French 'La Valeur du Travail, et sa Retribution en Hongrie.' in *Actes des Recherche en Science Sociales* 85, pp. 3–19, 1990 (November).

Stieber, Jack (1959), *The Steel Industry Wage Structure*, Harvard University Press, Cambridge, MA.

Stinchcombe, Arthur L. (1965), 'Social Structure and Organizations', pp. 142–193 in James G. March (ed.), *Handbook of Organizations,* Rand-McNally, Chicago.

Stinchcombe, Arthur L. (1983), *Economic Sociology,* Academic Press, New York.

Thompson, E. P. (1963), *The Making of the English Working Class,* Vintage, New York.

Ward, Chris (1990), *Russia's Cotton Workers and the New Economic Policy. Shop-Floor Culture and State Policy 1921–1929,* Cambridge University Press, New York.

Whyte, William Foote (1955), *Money and Motivation*, Harper and Row, Publishers, New York.

Williamson, Oliver E. (1985), *The Economic Institutions of Capitalism*, The Free Press, New York.

Wolfson, Mark A. (1985), 'Empirical Evidence on Incentive Problems and Their Mitigation in Oil and Gas Tax Shelter Programs', pp. 101–125 (Chap. 5) in John W. Pratt and Richard J. Zeckhauser (Eds.), *Principals and Agents: The Structure of Business,* Harvard Business School Press, Boston, MA.

Notes to chapter 9

* This paper is based on a set of lectures given in the Department of Sociology, University of Oslo on April 30, 1991, on October 10, 1992, and on March 20, 1993.
I thank seminar participants, Pål Foss, and two reviewers for useful comments and discussions.
In particular, I thank Vemund Snartland for research assistance. I acknowledge research support from NORAS, grant archival number 2161001.
An earlier and slightly shorter version of this paper appeared as Petersen (1993), *Acta Sociologica* 36 (3), published by Scandinavian University Press.

** Trond Petersen , University of California, Berkeley and University of Oslo, is a sociologist who does research on reward systems in formal organizations and on quantitative methods of social research.

1 In Scandinavia, the most important work in this tradition is Lysgaard (1961), which unfortunately has yet not appeared in English. It was also a conceptually strong study.

2 The U.S. branch of labour history is in particular empiricist, or more traditional historical. A recent and outstanding study is Cohen (1990). A related subfield, particularly strong in Britain, now known as labour studies with an emphasis on the labour process, is conceptually stronger, but its origins and practices are more in sociology than in history. For a synthetic account, from someone whose research was in industrial relations but who discusses the historical studies, see Edwards (1986).

3 Other accounts are given in Moe (1984), Arrow (1985), Perrow (1986, ch. 7), Nalbantian (1987), and Eisenhardt (1989). In the economics literature there are several accounts. Milgrom and Roberts (1992, ch. 5–6) give a careful, relatively nontechnical introduction, also discussing empirical cases to which the theory can be applied. Baron (1988) offers some viewpoints from the perspective of sociology and social psychology.

4 The branch of the principal-agent literature which relies less on mathematical modelling and is more empirical, termed positivist principal-agent literature by Eisenhardt (1989), does not rely heavily on the assumption of hyperrationality. Rather, it assumes agents to be guided by self-interest, but not necessarily to be omnipotent calculators (e.g., Jensen and Meckling 1976).

5 See, however, Stark (1990) on Hungarian work teams, where workers in fact rented machinery from management on after hours.

6 Shelton (1967), in a study of retail outlets that changed from a franchise system to a hired manager system or vice versa, showed that outlets run as franchises were more profitable, illustrating the powerful incentive effects of a system basing rewards exclusively on output and the weaker incentive effects of rewards that are independent of output.

7 Piece-rate contracts have been carefully studied in the formal principal-agent literature. Another relationship that has received much attention is that between shareholders of a company and its hired managers. Bearle and Means (1932) provide an early and important discussion of the so-called separation of ownership and control. Recent contributions using principal-agent theory include Jensen and Meckling (1976) and Fama and Jensen (1983).

8 In a classic article, Roy (1952) discusses how workers who receive a baseline wage plus a piece rate once they have reached a quota, referred to as 'making out', stop expending effort when they realize that they will not be able to make the quota. This phenomenon is called 'goldbricking' in the U.S., with similar names in other countries.

9 With a risk averse agent part of the moral hazard problem also arises because the outcome of the agent's action is uncertain and the reward depends partly on the outcome.

10 Williamson (1985, ch. 10) refers to this situation as a primitive team.

11 Similar screening over the long-term occurs in more standard market-type relationships as when producers build up differential reputations. See Wolfson (1985) for evidence from the oil industry.

12 Anderson (1985), in a study of salesforces for manufacturers of electronic items, shows similarly that when it is difficult to assess performance then there is higher reliance on an in-house salesforce paid on a salary and less reliance on outside salesrepresentatives remunerated on commission.

13 See also Petersen (1991), where a similar conclusion is reached based on an analysis of production workers in the U.S. nonferrous foundries industry.

14 Social rewards appear to be ubiquitous in employment relationships, independent of time and place, a phenomenon documented also by historians, not only in sociological and anthropological workplace ethnographies. Randall (1991, pp. 194–195) vividly describes how a weaver, named Teakle, who violated existing norms of work, was treated by other weavers, back in the year 1792 in Britain: 'Teakle was forced to sit astride the pole and was carried by the mob towards Owlpan where he was ducked in a mud hole.' Later in the evening, in the house where the weaver had sought refuge: 'Incensed, the crowd began to break down the door, smash the windows and also began to unroof the cottage. Those inside defended themselves as well as they could and after three hours of damage the mob left, threatening to return the following morning.'

15 Williamson (1985, pp. 50–52) briefly compares his framework to other frameworks, such as principal-agent theory. He notes that one of his central terms, opportunism, covers both 'moral hazard' and 'adverse selection', terms used in the principal-agent literature. Moreover, he claims that the analysis of moral hazard is too narrow in that literature.

16 This analysis draws on Petersen (1991). It is based on a particular model of the production technology and the preferences of the workers.

17 The industrial sociology literature is replete with discussions of how social rewards and trades among workers operate under output-related reward systems (e.g., Roy 1952; Whyte 1955; Burawoy 1979; Edwards 1986).

CHAPTER 10
ON THE RELEVANCE OF NEO-INSTITUTIONALISM TO THE PUBLIC SECTOR*

*Jan-Erik Lane** and Torgeir Nyen****

10.1 Introduction

The institutionalist paradigm became fashionable in the social sciences during the 1980s. This long known framework for social analysis which dates back to Montesquieu's emphasis on the importance of rules, i.e. laws and customs (1748) has been revitalized. The neo-institutionalist approach is well equipped to host new models, directing the conduct of social research towards a deeper understanding of the role of institutions for social systems. To sociology, the institutionalist perspective enters into a reorientation of social theory towards the interpretation of symbols in meaningful social interaction (Powell and DiMaggio, 1991). It has directed economic research to widen their models in order to explain how basic economic rules, e.g. the market mechanism, the private property regime or the joint stock company have emerged and are maintained. The study of international relations talk about so-called international regimes in order to account for how states reach cooperation (Krasner, 1983; Young, 1989).

Neo-institutionalism embraces a variety of new ideas, but three main institutionalist approaches can be identified, one sociological in tone and the others economic. The first approach originates in organization theory (Williams, 1990) whereas the second one has emerged from the new institutional economics (Williamson, 1986) and the third one from institutionalist economics (Coase, 1988; North, 1990).

In the neo-insitutionalist models there is a focus upon those institutions that are basic to society (Eggertson, 1990). Not taking the existence of institutions for granted,