Chapter 14

THE EMPLOYMENT RELATIONSHIP: JOB ATTACHMENT, WORK EFFORT, AND THE NATURE OF CONTRACTS

DONALD O. PARSONS*

The Ohio State University

1. Introduction

In the U.S. economy approximately 100 million workers are matched with market work activities on any given day. Millions more are matched with nonmarket activities of various kinds, including child rearing, home production of a variety of goods and services, and schooling. Economic efficiency requires that (1) specific individuals and activities be appropriately matched and that (2) the individuals, once matched, undertake the activity with an appropriate level of effort or intensity. In this chapter I focus on the economic forces that influence and define important aspects of these elements of the employment relationship in a market economy.

Historically the employment relationship in the United States has been a simple one. In the last century most individuals worked for themselves or in small firms, employees of the railways being the most notable exception. The individual undertook his chosen activity at an effort level that he judged appropriate and received income according to the market evaluation of the resulting good or service. Although simple, the outcome could also be harsh since family income was heavily dependent on earnings and earnings insurance was unavailable, except through public and private charity and of course the family.

The institutional structure of the employment relationship has, however, been transformed in this century. The nature of the workplace has changed radically.[1] The share of the workforce that was self-employed or unpaid family workers declined from almost 50 percent (47.08 percent) in 1900 to less than 10 percent (9.22 percent) in 1978. See Table 14.1.

*Support for this chapter was provided in part by the National Institute on Aging. The comments of John Garen, Thomas Kniesner, Howard Marvel, Randy Olsen, and Timothy Perri are gratefully acknowledged.

[1] One of the major transformations, the growth in trade unionism, is considered at length in Chapter 18 by Henry Farber in this Handbook.

Handbook of Labor Economics, Volume II, Edited by O. Ashenfelter and R. Layard

Table 14.1
Self-employed and unpaid family workers as a share of total employees, 1900–1978.

Year	Self-employment share (%)
1978	9.22
1970	10.21
1960	16.25
1950	20.50
1940	26.83
1930	29.44
1920	32.35
1910	38.57
1900	47.08

Source: 1900–1960, S. Lebergott, *Manpower in Economic Growth*. New York: McGraw-Hill, 1964, p. 513; 1970, 1978, *Statistical Abstract of the United States, 1979*, p. 403.

The average size of workplace in the non-self-employed sectors of the economy moreover has increased rapidly over this period. In the manufacturing sector, for example, the proportion of employment in small workplaces with less than twenty workers declined from 14.4 percent in 1909 to 6.5 percent in 1977 (Table 14.2). Over the same time period, the proportion employed in establishments of 1000 or more workers doubled from 15.3 percent to 27.5 percent.

Table 14.2
Share of manufacturing employment by establishment size in the United States, 1909–1977.

	Establishment size (employment)					
	1–19 (%)	20–49 (%)	50–249 (%)	250–499 (%)	500–999 (%)	1000+ (%)
1977	6.5	8.8	28.1	15.6	13.5	27.5
1972	6.2	8.7	27.8	15.4	13.1	28.7
1967	5.6	8.3	26.0	14.5	12.8	32.8
1963	7.2	9.1	26.6	14.2	12.4	30.5
1958	7.8	9.4	26.0	14.0	12.3	30.5
1954	7.6	8.7	25.0	13.5	12.6	32.6
1947	7.2	8.7	24.7	13.5	13.1	32.8
1939	9.5	9.7	29.4	16.1	13.0	22.3
1929	9.9	9.2	28.1	15.1	13.3	24.4
1919	10.4	9.1	27.2	13.8	13.2	26.4
1914	13.1	10.6	30.0	15.3	13.2	17.8
1909	14.4	11.6	30.8	15.2	12.7	15.3

Source: Various Censuses of Manufacturing. The employment size intervals for 1939 and before include next highest integer, e.g. 1–20. The 1939 data are reported in the 1940 census.

As the size of the workplace grew, the need for explicit and implicit employment contracts to define and regulate the employment relationship grew correspondingly. At the same time, however, the information necessary for efficient employment contracts became for the most part more expensive. Information on employer and employee circumstances and activities that might in a small workplace be free or quite inexpensive, a byproduct of other productive activities, may be observable in a large firm only at prohibitive cost to all parties to the contract. Much of the recent literature on the employment relationship has stressed the interplay of efficiency objectives and the limitations imposed on the form of employment contracts by information costs.

The developments in each of the three areas reviewed below (work effort, specific human capital, and earnings insurance) have in many respects been independent but, as we shall see, the underlying approach in each is remarkably similar. Each focuses on the potentially distorting effect of contract enforceability and asymmetric information on efficient employment relationships. In each area, moreover, two broad questions have formed the bases for the analyses:

(1) among homogeneous workers, what is the optimal employment "contract" (service agreement and compensation package)? and

(2) among heterogeneous workers with identical observable traits, how can the provisions of the optimal contract be altered to secure a more appropriate match of worker and firm?

In the specific human capital literature, for example, the obvious returns to reduced job mobility suggest that backloaded compensation packages such as nonvested pensions are likely to be optimal among homogeneous workers. As it happens, if apparently identical workers differ in their mobility propensities, this same compensation scheme may have important self-selection effects as well. With a backloaded compensation package, the expected value of the package will be highest for the workers with the lowest self-perceived mobility propensity. Since contract considerations will play an important part in much of the discussion that follows, Section 2 is devoted to a brief review of several major themes in contract theory.

Contracts, whether implicit or explicit, can do no more than make feasible what would otherwise be mutually agreeable joint activities. The employment relationship is primarily formed by more basic considerations. In this chapter three aspects of the employment relationship will be reviewed in detail: (1) the supply of work effort by the employee, (2) the investment in employer–employee match specific skills, and (3) the provision of earnings insurance by the employer. The supply and demand for work effort of individual workers is an obvious and crucial factor in the employment relationship. Employers have preferences about the intensity with which employees undertake their tasks, preferences which may be quite at variance with those of the employees. It is essential that the employee be motivated to undertake the assigned task at the mutually agreed intensity. This may require no more than a handshake, although the large array of

employment incentive devices suggests otherwise. More typically some combination of incentives and employer monitoring may be required. The forces that determine the form of the optimal compensation package under various economic circumstances are developed in Section 3. Worker reliability is not likely to be homogeneously distributed over the workforce so that the job matching of earnest workers and vulnerable firms, as well as appropriate effort incentives once matched, may be important in the optimal employment relationship. The impact of worker heterogeneity on employment contracts is therefore also considered in this section.

More than optimal work effort is required of an efficient economy. Efficient job mobility is important as well if the economy is to respond to the rapid changes in product demand that characterize market economies. In the U.S. manufacturing sector between 1972 and 1977, for example, employment increased by 50 percent or more in several industries, including X-ray appliances and tubes (155 percent), fluid meters (80 percent), and oil field equipment (63 percent), while declining by 50 percent or more in others such as ammunition (−63 percent) and wool yard (−51 percent).[2] The rate of turnover in the economy, in part required to accommodate these changes, is large. In the manufacturing sector, for example, the average *monthly* employee separation rate fluctuated between 3.8 percent and 4.9 percent during the 1970s, suggesting that a 50 percent turnover rate over the course of a year is not unusual.[3]

Perfect fluidity among jobs is not likely to be optimal, however. Indeed, a large part of the workforce in the United States secures long time employment with a single firm. Two recent studies [Akerlof and Main (1981) and Hall (1982)] have attempted to estimate job duration in the British and U.S. economies. Hall (1982, p. 720) estimates that eventual completed tenure for all U.S. workers with a job in 1978 has a median of 7.7 years and that 28 percent are currently with a job that will last 20 years or more.

Fruitful models of the employment relationship must explain the individual incidence of job attachment and job turnover (and unemployment) as well as aggregate levels since job turnover probabilities are not uniform across individuals and groups. In particular, turnover and unemployment are concentrated among the inexperienced and among the poorly educated. Mincer and Jovanovic (1979) report two-year job separation probabilities that vary from more than 70 percent for males who have been working for less than a decade and who have less than one year on their current job to 5 to 6 percent for individuals who have

[2] One important dimension of job mobility, job search, will not be reviewed systematically in this chapter since the topic is reviewed in Mortensen (Chapter 15 in this Handbook). For an earlier theoretical and empirical survey of job search, see Parsons (1977); see also the still excellent theoretical survey by Lippman and McCall (1976).

[3] U.S. Bureau of the Census (1977).

Table 14.3
The ten-year retention rate within the firm among older males, by race and education, 1966–1976.

Schooling attainment (in years)	Race	
	White (%)	Black (%)
0–8	35	32
9–11	44	47
12	49	51
13–15	44	61
16+	55	71
Total	44	39

Source: National Longitudinal Surveys. The sample was limited to males 45 to 51 years of age in 1966 (55–61 in 1976). Sample size is 2006: 1454 whites and 552 blacks.

been working 40 to 44 years and have been with their current firm for more than 10 years.

Ten-year longitudinal data from the National Longitudinal Surveys similarly indicates that prevalence of job stability and illustrates as well systematic differences by skill. In Table 14.3 I report the ten-year retention rate among male respondents who were 45 to 51 years of age in 1966. For the group as a whole 44 percent of the surviving whites and 39 percent of the surviving blacks were with the same firm ten years later. The retention rate rises from 35 percent to 55 percent among whites as education increases from 0–8 years of schooling to 16 or more years of schooling and from 32 percent to 71 percent among blacks for the same schooling increase.

One factor that alters the economic value of job attachment is on-the-job learning, particularly learning specific to the firm, e.g. attributes of its suppliers, capital, personnel, and customers. Obviously a job match that has value specific to the firm and the individual will reduce mobility in any sensible economic regime, although, with fluctuating market conditions, not necessarily to zero.

The measurement of these direct and indirect job specific investments in the work force is imprecise since most costs are indirect. Nonetheless the few management studies reported in the economics literature suggest that the investment costs are substantial. Mincer (1962, p. 62) cites an American Management Association study of California firms which reported hiring, training, and separation costs per worker of $1535 in 1982 dollars. Oi (1962, p. 546) reports fixed employment costs at International Harvester of $1418 per worker in 1982 dollars. The investment expenditures apparently rise rapidly with skill level. Oi, for

Table 14.4
Firm investment per employee, 1969 (in 1982 dollars).

Skill level	Firm investment ($)
Least skilled (e.g. materials handler)	911
Semi-skilled (e.g. maintenance mechanic)	5715
First-line supervisor	13353
Middle manager	53413
Top level manager	113503

Source: Parsons (1972). All dollar figures are readjusted to 1982 $ levels.

example, reports estimates of $470 for a common laborer, $44765 for a two-year progressive student and $69778 for a four-year apprentice, again all in 1982 dollars. Parsons (1972) discusses a study undertaken in 1969 by a manufacturing firm, R. G. Barry, that indicates that the firm's investment in employees ranged from $911 for the lowest skill category to $113503 for a top level manager (1982 dollars). The full results are reported in Table 14.4.

The observed patterns of turnover are consistent with the importance of preserving match specific skills and with the positive correlation of general and specific skills illustrated in Table 14.4. Whether these separation rates are fully efficient, however, depends very much on the nature of feasible contracts. Efficiency is not assured by the usual competitive assumptions since the unique value of the specific job match implies that this income generating asset lacks the labor market guarantees carried by more widely demanded skills. Contracting problems may induce inappropriate initial job matching as well as subsequent job attachment if workers are heterogeneous in their mobility propensities. Unobservable heterogeneity among workers may induce firms to introduce a variety of ancillary employment practices such as screening and self-selection devices when specific human capital investments are heavy.[4] The nature of contracting and its implications for job separation in the presence of specific human capital are developed in Section 4.

In Section 5 a second factor that may alter the stability of the employment match is considered, namely worker and owner preferences for stable incomes. The income of the self-employed is, of course, vulnerable to business cycle fluctuations and more idiosyncratic, firm-specific reversals.[5] Among employees the important role of job loss in the cyclical behavior of experienced worker

[4] U.S. Department of Labor (1980).

[5] Aggregation phenomena that may be important in cyclical unemployment, e.g. the labor market congestion problems that arise with a greater number of simultaneous layoffs [Parsons (1980)], will be ignored below.

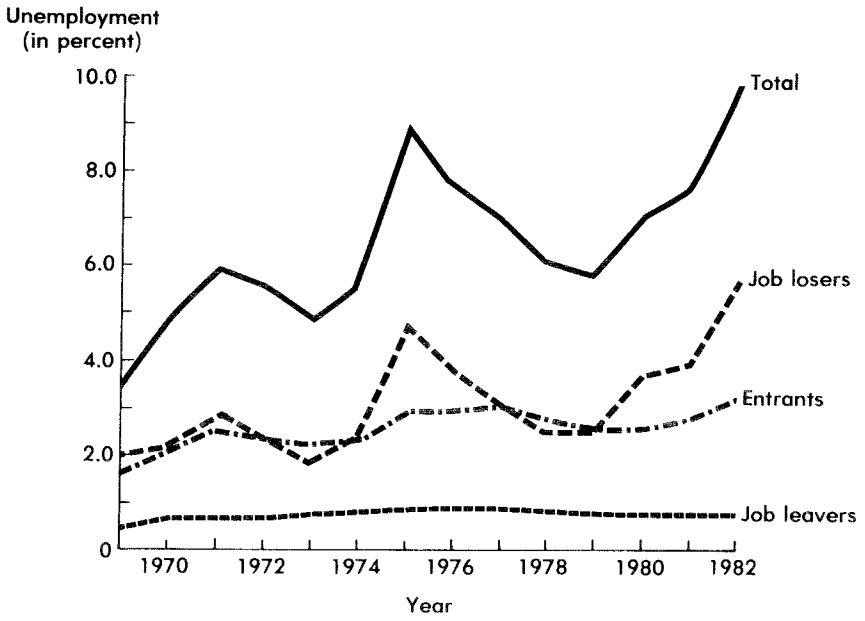


Figure 14.1. Unemployment rate by reason. *Source: Handbook of Labor Statistics*, 1980, p. 82.

unemployment and consequent earnings losses is apparent in Figure 14.1 in which the unemployment rate is separated into components by "cause" over the period 1969–1979. Clearly unemployment due to job loss is a powerful factor in the determination of total unemployment over the business cycle.

Section 5 will focus primarily on the attempts by individual workers to secure from the firm some form of earnings insurance. Although the individual may have alternative income sources, both public and private, standard insurance policies against the possibility of reduced earnings are not widely marketed. Presumably potential insurers perceive substantial moral hazard and adverse selection problems, induced by the insurer's inability to distinguish exogenous, random earnings losses from those due to choice or foreknowledge. The employing firm knows the nature of its own business conditions and the prospects of its employees better than a third party insurer and, if sufficiently large, may serve the function of insurer, either by offering direct cash payments to those laid off or by retaining them on payroll. The most direct application of this reasoning suggests that job attachments will be more secure (less likely to be broken) during periods of declining demand in large firms. Recently more complex models have been developed, combining aggregate nondiversifiable risk and severe informa-

tion problems of a particular sort, which yield contrary implications. Again these insurance models of the employment relationship are the focus of Section 5.

The review below will concentrate on theoretical models, in part because the review is already overlong and in part because empirical work has lagged substantially behind theoretical developments in this area. Unfortunately, given the scarcity and general primitiveness of existing employer data sets and the relative subtlety of many of the theoretical implications, this imbalance is likely to remain for some time.

2. The employment contract

A mutually advantageous exchange of services and payments for services between two parties may not be instantaneously and verifiably executed; either the service or payment exchange may be incompletely monitored or deferred in time. The reliability of performance by the contracting parties may therefore be a crucial characteristic of the economic environment. Indeed a theme that reoccurs throughout the following sections is the potentially important effect that contract enforceability may have on the employment relationship. In this section important elements of the current theory of contracts of particular relevance to the employment relationship will be reviewed.

The main function of any contract is to constrain behavior, behavior that in a broader context is nonoptimal but which at the moment the decision is made is attractive to some economic agent.[6] Life cycle reallocation of consumption by borrowing when young and repaying the loan when older, for example, may be optimal for a given individual; repayment of the loan as an isolated activity is surely unattractive to this same individual. In a sense, what follows can be viewed as an analysis of the limitations of contract relationships, most frequently as a consequence of information and enforcement difficulties.

The most prominent restrictions on employment contracts are legal, a reflection of the fact that human capital is embodied in individuals with certain "inalienable rights". Slavery and indentured servant contracts, for example, are not enforceable and long-term contracts that restrict individual mobility and behavior generally are viewed skeptically by the courts. While themselves not burdensome, these restrictions are only the most dramatic of a wide range of restrictions derived from the same philosophic base. Bankruptcy laws are a reflection of the same concerns, leading to the peculiar difficulty individuals have when young (and without collateral) of borrowing for schooling, apprenticeship

[6] Occasionally the term contract is used to mean repeated exchange *voluntarily* entered into in *each* period, e.g. Bull (1983). The standard notion of a contract as a "binding agreement" is lost with such a usage.

costs, and other employment expenditures. Sensitivity to this borrowing constraint is important since a frequently proposed solution to reliability problems is some form of bonding which may in fact not be feasible.

A second contracting problem is that of information difficulties, particularly the asymmetric access to information that economic agents may have at a given time.[7] Certainly since much information (knowledge) relevant to the employment contract is generated (learned) as a byproduct of other activities the individuals and firms may undertake, information costs may differ across agents. Of particular consequence, the joint-product nature of information acquisition implies that unrelated third parties are likely to find information on employer and employee behavior relatively more costly to obtain than do the employer and employee themselves.

Insurance models make clear the importance of the information available to the contracting parties in determining equilibrium behavior and contract form. Theoretical models of the insurance industry, for example, have long recognized two information problems that, if sufficiently severe, make insurance contracting infeasible, namely moral hazard and adverse selection.[8] A contract may affect an insurer's loss experience adversely if losses are determined in part by the activities of the insured and if it is costly or impossible for the insurer to monitor these activities (moral hazard). Moreover, if the population is heterogeneous in the likelihood of loss and if this is known to the insured but not the insurer, the insurer may find his losses greater than would be anticipated by preinsurance loss rates because those with the greatest likelihood of loss will be the most eager to secure coverage (adverse selection). A third information problem, the inability to determine accurately and costlessly which state of nature has in fact occurred (imperfect state verification), has also recently received attention in the literature.[9] Parallels to the effort-monitoring issue in employment contracts should be self-evident. In one form or another, these same problems limit the range of feasible employment contracts and thereby alter the employment relationship.

Depending on the circumstances, the time dimension may be important in the information process. Information on a specific event may be relatively less

[7]Access to information is of course a loose but customary way of asserting that information acquisition may be costly. See Wachter, Williamson, and Harris (1975) for a detailed discussion of transactions costs underlying information collection. The term efficiency will often be used in the customary if misleading sense of optimal performance under the assumption that information acquisition is free.

[8]Important formal models of these problems include Spence and Zeckhauser (1971) and Pauly (1974); for a discussion of market equilibrium in the presence of adverse selection, see Rothschild and Stiglitz (1976). See also Akerlof (1970) for an early but still valuable discussion of the general problem of market problems that arise because of adverse selection.

[9]Townshend (1979) discusses an interesting model of state verification in which the agent who does not have access to critical information may purchase it. The agent will tend to do so optimally only if he suspects the informed agent's claim is quite wide of the mark; suspected minor violations will be ignored. Parsons (1984) develops a model with imperfect state verification.

expensive to collect at one point in time than at another. Generally one would expect information to be less costly if one were willing to wait longer for it, although even that process has obvious limits. This aspect of the process is not considered carefully in the literature; typically the time dependence of costs is simply characterized by its presence or absence at a particular time a decision must by assumption be made.

2.1. *Explicit contracts*

Contracts may be formal, explicit documents or less well-defined implicit agreements. Formal, explicit contracts have one major advantage, enforcement costs of contract performance are in part subsidized by the state. A wide range of penalties or damages can be imposed on a nonperforming party by the courts. In an uncertain world in which no individual is absolutely reliable, such a subsidized enforcement mechanism has a transparent appeal.

Explicit contracts, however, are quite demanding of information and may not be feasible. As a practical matter explicit contract contingencies must be limited to readily observable outcomes, a rather powerful restriction. Frequently only one of the two immediate parties to the relationship may know an important piece of information with any precision. The other contracting party, much less an uninvolved third party, may not have access to the information at any reasonable cost. Subtle questions such as whether a worker voluntarily quit because he felt he had prospects of a better job elsewhere or whether he was coerced to leave through an employer's manipulation of nonpecuniary job conditions are crucial to the efficient contracts considered below yet are not easily answered by a third party uninvolved in the case. This information verification problem surely limits the role that third parties can play in the enforcement of agreements and ultimately limits the exchanges that can be undertaken.

It is important to recall that performance bonding and the use of collateral make damage collection easier but do not eliminate the need for the courts or other contract enforcement mechanisms.[10] Someone, the courts or otherwise, must determine whether a bond or other form of collateral is to be forfeited.[11] Presumably this judgment requires observation of the agent's behavior and any appropriate contractual contingencies. The importance of institutional constraints, specifically the bankruptcy constraint on borrowing without collateral,

[10] See Benjamin (1978) for a discussion of the role of collateral in contract performance.

[11] Landes and Posner (1979) provide an insightful discussion of private "courts"; ultimately an agreement to abide by binding private arbitration if a dispute arises in the execution of a contract requires an enforcement mechanism to ensure that the binding arbitration agreement is itself honored by the parties.

must also be considered in assessing the effectiveness of explicit performance "bonding".[12]

Other informational requirements for explicit contracting may be important. Courts enforce contracts primarily through the imposition of damages. Many theories exist of how damages ought to be assessed in a situation of contract breach due to unforeseen contingencies but little empirical research exists on how damages are in fact imposed in such situations.[13] There is little reason to suppose court behavior is consistent. It is important that all contingencies which could lead to contract breach be foreseen and that the damages be prespecified if uncertainty about the outcome of the contract is to be eliminated. All possible states of the world can rarely be foreseen and precontracted, so that even a carefully considered, explicit contract will involve some degree of undesirable uncertainty.

2.2. *Implicit contracts*

Information costs may make third party enforcement of explicit contracts infeasible. The two parties to a contract may know whether or not satisfactory performance on a contract has been undertaken, yet find it prohibitively expensive to demonstrate that to another, unrelated individual. Almost surely, certain mutually beneficial exchanges will be hindered by this infeasibility. Private agreements or implicit contracts that are not enforceable in the courts may, in such circumstances, be attractive and, depending on circumstances, more or less efficient. The efficiency of the arrangement depends critically on the extent to which the agreement can be enforced by less formal mechanisms. The literature has focused on two possible economic enforcement mechanisms: subsequent, profitable relationships between the two contracting parties (repeated exchange) and reputational effects of contract performance that might alter subsequent contracting by other individuals and firms with the parties involved.

Concern about future, potentially profitable exchange between the two parties may provide some assurance of contract performance. Information requirements are limited since only the parties to the contract need have access to the information, so such contracts may be feasible when explicit contracts are not. Simple game theory examples, for example the prisoner's dilemma, illustrate that the enforcement power of such relationships may be more apparent than real [Luce and Raiffa (1957)]. Specifically if the relationship has a finite, known end

[12] See Kennan (1979) and MacDonald (1982) for discussions of borrowing constraints on bonding of specific human capital investments and Eaton and White (1982) for a similar discussion on effort bonding.

[13] Useful discussions of economic models of contract damages can be found in Barton (1972) and Shavell (1980).

point, it may be equivalent in a behavioral sense to a single period model. Since the last period is equivalent to a single period, the party coerced to perform would default in the last period since the contract is no longer multi-period, which in turn suggests that the next-to-last period is the "final" contract period from an economic viewpoint. This process leads to the conclusion that such a contract is not viable in any period. Infinite relationships avoid this problem, most plausibly those with random end points [Telser (1980)]. The use of approximate solution algorithms suggests that repeated exchanges may be self-enforcing if the number of exchange periods, although finite, is sufficiently large [Radner (1981)]. The introduction of some uncertainty in the expected behavior of the other party may also lead to more cooperative solutions [Kreps and Wilson (1982)].

The termination of any subsequent profitable relationships between the two contracting parties is only a special case, though perhaps an important one, of indirect, economic damages imposed on an individual who breaches an implicit contract. Other parties, somehow made aware of the contract breach, may alter their behavior in a manner which is adverse to the defaulter. These third party effects will be labelled reputational effects.

Reputational models were to my knowledge first developed extensively in the advertising and product quality literature. Nerlove and Arrow (1962) explored the dynamics of reputation development through advertising. Gould (1970) developed models of information dissemination more explicitly within the Nerlove–Arrow framework. Obviously the extent and speed with which information spreads to interested economic agents is crucial to the contract enforcement function of reputation. Nelson (1974) first linked the reputational process with product quality, essentially arguing that the building of brand identity (reputation) through advertising creates a performance bond of sorts such that heavy advertising and product quality will be positively linked. See Klein and Leffler (1981) for a general, informal discussion of this idea. The heart of the process is the notion that reputation is a bond that an individual posts for good performance, a bond not in the traditional sense of forfeiture of a tradeable asset but rather in the sense that poor performance will reduce the individual's wealth or asset holdings. Kotowitz and Mathewson (1979) and Schmalensee (1978) attempt to model this process formally (using the Nerlove–Arrow framework and a Markov process, respectively) and find not surprisingly that such reputational enforcement mechanism need not function desirably in all circumstances. If, for example, consumers rely heavily on the notion that advertised brands are high quality products, they may be profitably fooled by a cunning advertiser [Schmalensee (1978)]. Kihlstrom and Riordan (1984) note that such notions would not persist in the long run among rational consumers and consider the demand and cost structures that would lead to reputational equilibria. See also Shapiro (1982).

Presumably employment contracts enforced by reputational capital may be similarly secure under some circumstances yet vulnerable to breach under others, although the circumstances have yet to be well specified.[14] One might conjecture that, for given market size, reputation effects will be more important for large firms than small. A dissatisfied employee of a large firm is more likely to communicate directly or indirectly with a potential employee of that firm than is a dissatisfied employee of a smaller firm. Memory of an incident at a better known firm is also more likely to be persistent among uninvolved third parties. Empirical evidence for such size effects can be found in the noncontracted inflation adjustments that firms have made in retiree benefits in recent years. Allen, Clark and Summer (1984) report that such adjustments are positively and systematically related to firm (pension plan) size.[15]

A similar line of reasoning would suggest that individual employees will be less bound by reputational effects than will firms.[16] Inexpensive information on a specific individual is unlikely to be readily available to the firm and reputationally enforced reliability correspondingly less powerful. If sufficiently important, of course, information on any individual's past performance can be secured, so reputational effects may be important for sensitive jobs such as top management positions for which firms will be willing to invest substantial resources in search and information collection [Fama (1980), Perri (1983)]. Contract default may be expensive to such individuals and therefore implicit contracts relatively secure.

Although not directed explicitly at the reputational issue, empirical evidence does exist that past labor market behavior has effects on current labor market opportunities for both firms and workers. Abowd and Ashenfelter (1981), for example, estimate substantial compensating wage differentials for industries with histories of high past layoffs. Whether such informational effects hold at the individual firm level was not considered. Mincer and Jovanovic (1981) and Bartel and Borjas (1981) both present evidence that persistent job turnover by older workers has negative effects on the subsequent wages such individuals can secure. Much remains to be done of a theoretical and empirical nature on this issue.

I should stress that an "enforceable" implicit agreement does not require that both parties be reliable. Even if only one of the parties to an agreement, say the firm, is reliable, information costs and contract monitoring costs are much reduced and a variety of otherwise infeasible agreements may be undertaken. It only becomes necessary that appropriate information be known to the reliable

[14] Holmstrom (1981) and Charmichael (1984) provide brief discussions of reputational issues in the labor market.

[15] See also Hatch (1982).

[16] Firms may be more reliable because the end pointed problem that limits the repeated-exchange enforcement technique should be less severe than it would be for an individual. John Garen made this point to me.

party at the time of contract execution, not the unreliable party nor any uninvolved third party. The performance of the less reliable party in such circumstances may be more explicitly bonded, with the bond held by the reliable party. If the appropriate information is not costlessly available to the reliable party, some form of more costly performance monitoring may be necessary.

Some individuals not only appear to be more reliable, they are in fact more reliable, whether because of some innate differences in ethical standards or because of ancillary social pressures that make breaching a contract costly. To the extent such traits are observable, such individuals will be differentially attractive for situations in which monitoring and enforcement possibilities are weak. It is surely not inadvertent that many of the world's more repressive dictators have assigned brothers to head the army or police. Most firms have long since passed the stage in which family members can staff all responsible positions, but other selection devices may achieve something of the same goal. Firms may screen workers for reliability (in work effort or job attachment) both through initial statistical discrimination according to observable correlates of reliability and through trial periods of employment. The compensation schedule may also be structured to induce self-selection if the worker but not the firm is aware of the worker's reliability. Such mechanisms may make possible some attractive but otherwise infeasible exchanges.

3. Work effort and compensation

The employer and employee must negotiate a variety of terms for the provision of labor services. At least two dimensions of work commitment are fundamental: the number of hours the individual will spend at the work place and the intensity with which he undertakes the assigned tasks during that period. Traditionally factor demand theory, indeed demand theory in general, has assumed that the quantity and quality of services purchased can be perfectly and costlessly observed by the buyer. In the labor market, the time commitment is observable at relatively low cost in most circumstances. Even here exceptions come to mind, including the work hours of individuals who perform work away from any central work place, for example many types of salesmen. The quality of the input, in the present context work intensity or effort, typically involves greater measurement difficulty.[17] Compensation structures designed to minimize these difficulties are the focus of this section.

[17]Effort supply functions will be ignored below. See Pencavel (1977) for a discussion of such models analogous to traditional work hours supply models.

3.1. The compensation of homogeneous workers

The employee of course knows the level of intensity he is undertaking without incurring any substantial monitoring costs. Were he a reliable agent, his compensation could be based on his reported effort level. The models discussed in this section assume that the employer cannot systematically rely on these reports but instead must undertake some direct metering of worker activities and base the worker's compensation upon that metering.

The firm faces the choice of metering and rewarding input quantity and quality or output quantity and quality. Consider, for example, a value production function of the form:

$$V(q, Q) = f(h, H; \theta), \tag{3.1}$$

where V is the value of output, a function of the quantity (q) and quality (Q) of output, and $f(\cdot)$ is the production function with the quantity (h) and quality (H) of work effort and a random element (θ) as arguments.

If no randomness exists in the production technology, an incentive system based on output measures would be equivalent to one based on input measures. The system chosen would depend solely on observability considerations. Output quantity and quality need not be any easier to measure than input attributes. The output of white-collar tasks is often difficult to measure as is the output of many low-skilled tasks such as general maintenance. In large-scale production processes, individual output is often not even well defined conceptually. If one set of quantity/quality pairs, either for input or output, is observable at little cost, however, it will presumably form the basis for an incentive contract in this circumstance.

If workers and employers are both risk neutral, the introduction of randomness in the production process would not alter the certainty conclusion in any significant way. If, however, the standard assumption is maintained that workers are risk averse and firms risk neutral, the two methods are no longer equivalent. The worker presumably knows with certainty the hours and effort he provides to the firm and would, other things equal, prefer to be paid a certain wage based on that effort. If output alone is observable, the wage payment must in part be based on output, however random, if the worker is to be given an incentive to expend effort. Such a piece rate system will unfortunately introduce unwanted variability into the worker's earnings. The optimal compensation scheme must balance these considerations.

In the analysis that follows, I consider in turn the case in which there is no production function randomness and the case in which there is significant randomness in "production". I will label these cases the production worker

model and the managerial model since productivity risk is not likely to be a significant concern for most forms of production work.

I should note before beginning that the role of firm size on the optimal compensation scheme is not transparent. In small firms with nonstochastic production processes, the owner or manager is likely to observe both input intensity and output quality and the particular payment scheme may be a matter of indifference. If substantial risk does exist in the production process, the small firm owner may be less willing to absorb the risk entirely so that risk sharing may be more prevalent than in large firms with diversified owners.

3.1.1. The compensation of production workers

Large differences exist in the compensation structure of production workers. In a sample of 58 manufacturing industries subject to BLS nationwide occupational surveys between 1963 and 1968, six industries, primarily in apparel and footwear, reported 70 percent or more of their production workers were paid on an incentive basis. Two, Work Clothing and Men's and Boys' Shirts, reported that more than 80 percent of their production workers were paid on an incentive basis. Conversely seven industries reported that 2 percent or less of their workforce were paid on an incentive basis. In cigarette manufacturing and petroleum refining the percentage was less than one half a percent. See Table 14.5.

Table 14.5
Percent of production workers paid under incentive wage plans in selected manufacturing industries, 1963–68.[a]

Work Clothing	82%
Men's and Boys' Shirts	81
Men's and Boys' Suits and Coats	74
Footwear, except Rubber	70
Women's Hosiery	70
Children's Hosiery	70
Flour and Other Grain Products	2
Motor Vehicles	2
Paints and Varnishes	1
Fertilizer	1
Synthetic Fibers, Noncellulose	1
Cigarettes	0
Petroleum Refining	0

Source: Stelluto (1969).

[a]For precise industry definition and dates, see source.

What could account for these large differences in compensation systems? The industries at the two extremes suggest the possibility that firm size may be important although it is not altogether clear why smaller firms (as would typically be found in the apparel industry) would be more likely to use incentive pay than would larger firms (such as those that characterize petroleum refining). Indeed, one might expect that the greater information decentralization of large firms would lead to greater reliance on incentive pay.

Alchian and Demsetz (1962) propose two factors that may affect the relative efficiency of (1) output oriented reward systems such as incentive pay and (2) input oriented systems with direct supervision of effort. The principal of these is the difficulty of designing incentive pay when workers are organized in teams. If a large number of workers are needed to operate a single blast furnace, for example, aggregate output is a poor measure of performance by one single individual of the team; individual shirking remains a problem and direct monitoring of effort is likely to be necessary. A second argument with many of the same empirical implications is that employer monitoring of the worker's treatment of the firm's capital may be important in any case, the more so the greater the amount of capital and its complexity. If output is the sole reward criterion, the worker has an obvious incentive to use the cooperating capital in a nonoptimal way much as a renter of a house may treat it differently than would the home owner. If indeed the employer is already directly monitoring the worker's treatment of the capital equipment, the marginal cost of direct supervision of worker effort is surely reduced as well.

Both arguments by Alchian and Demsetz suggest that the prevalence of incentive pay should decrease with the capital intensity of the industry; conversely direct supervision of effort should increase. A multivariate analysis of the

Table 14.6
Determinants of the percentage of production workers paid on an incentive basis in selected manufacturing industries, 1963–68.[a]

	Coefficient	Standard error	*t*-statistic
Constant	56.107	8.318	6.74
Emp500[b]	0.217	0.100	2.18
$\log(K/L)$[c]	−17.317	2.632	−6.58
$\bar{R}^2$	0.43		

[a] The dependent variable is the percentage of production workers paid on an incentive basis. The sample size is 55 manufacturing industries, primarily at the 4-digit SIC level.

[b] The percentage of workers in establishments with 250 or more workers.

[c] The log of the capital-to-labor ratio.

complete set of incentive pay data collected by Stelluto (1969) strongly confirms these conjectures. An ordinary least squares regression of the percentage of production workers paid on an incentive basis on establishment size (the percentage of workers in establishments with 250 or more workers) and the log of the capital to labor ratio was undertaken and the results reported in Table 14.6. The coefficient of the log of the capital-to-labor ratio is negative and strongly statistically significant with a t-statistic of -6.58. The coefficient implies that a 10 percent increase in capital intensity will induce a 1.7 percentage point decline in production workers on incentive pay. Controlling for capital intensity, firm size has a positive coefficient significant at the 5 percent level. Apparently incentive pay probabilities do increase with firm size once the dominating effect of capital intensity is removed from the data. The effect is not large, however. A one percentage point increase in the prevalence of large establishments induces a 0.2 percentage point increase in the prevalence of incentive pay for production workers.

3.1.2. Models of effort monitoring and compensation

Much of the supervision process has not been carefully modelled. Several aspects of job supervision and compensation policies have been developed, however, and warrant discussion. In particular, economists have developed models that clarify the nature of the optimal employment relationship when input monitoring is incomplete, presumably because effort monitoring is costly. These models take two forms, occasional but precise monitoring and continuous but imperfect monitoring. The behavioral distinction is primarily in the introduction of risk to the reliable worker that imperfect monitoring introduces into the process.

The question of the optimal amount of sampling of behavior and of the corresponding penalty to impose for detected malfeasance has arisen naturally in the law and economics literature [Becker and Stigler (1974)].[18] Although the discussions focus on optimal compensation schemes for law enforcers subject to bribes, the extension to shrinking or malfeasance by a firm's employee is self-evident.

Becker and Stigler, for example, consider the question of the optimal employment contract of law enforcement officials who are subject to the temptation of bribes (b). Becker and Stigler conclude that a bond/severance pay compensation scheme will generally be efficient. If monitoring is costless and complete, the performance bond (B) must equal or exceed the magnitude of the bribe and performance is assured. If, however, monitoring is costly and perhaps incomplete, so that the probability of the detection of malfeasance and the forfeiture of the bond (p) is less than one, then performance is assured for all but risk-preferring

[18] See also Becker (1968).

individuals if

$$pB \geq b,$$

where p = detection probability, B = magnitude of bond, and b = magnitude of bribe. The expected return from accepting the bribe must be negative for bribery to be forestalled.

Any combination of detection magnitude and bond size that satisfies this inequality will secure complete compliance in this model [Harris and Raviv (1979) label complete compliance contracts "forcing contracts"]. Since performance monitoring normally requires the expenditure of real resources and posting of the bond does not, efficiency would seem to require an infinitesimally small level of monitoring activity and an infinitely large bond. Reasons why bonds and monitoring are finite are many. Capital difficulties facing workers may limit the size of the bond a worker is willing to post. The reliability of the firm in detecting (claiming) malfeasance will be strained with a sufficiently large bond. Risk aversion becomes a potentially important constraint on the equilibrium size of the bond when the possibility of misclassification arises since not even those workers making the contractually required effort are totally secure from bond forfeiture.

The Becker–Stigler model need involve no question of risk preferences since the monitoring is 100 percent accurate when undertaken; the reliably performing agent will never be disciplined. Harris and Raviv (1979) propose an alternative monitoring model in which the monitor is imperfect. They present an example of a market structure in which the optimal employment structure is quite similar to the Becker–Stigler bonding model, estimated effort above a critical level is paid a fixed wage, below the critical level a different, possibly zero wage (the worker is discharged). The worker's risk preference becomes central to the optimal level of the critical effort requirement and fixed wage in this situation.

Consider the example more specifically. Assume in particular that the value of production (V) is equal to the effort expended (H) so that

$$V = H,$$

and that the worker's utility function has the explicit form:

$$U = W^{\gamma} - H^{(1+\gamma)}/(1+\gamma),$$

where γ is a risk preference parameter in the interval $0 < \gamma < 1$. The individual therefore is risk averse with respect to wealth in this model and experiences increased disutility at an increasing rate with effort. Effort is unobservable but a "monitoring technology" exists that provides an unbiased estimate of effort, say $\hat{H}$, $E(H) = \hat{H}$, that is subject to a uniformly distributed random error over the

interval

$$[\hat{H} - \varepsilon, \hat{H} + \varepsilon].$$

Harris and Raviv assert that the optimal compensation contract in this case has the form:

$$W = \begin{cases} W^*, & \text{if } \hat{H} > H_c, \\ 0, & \text{otherwise,} \end{cases}$$

where $W^* = (1+\gamma)\varepsilon$ and $H_c = \varepsilon + \delta\gamma$ with $\delta = (2\varepsilon^{1-\gamma})^{-1/\gamma}$. The firm will pay a wage W^* if the monitoring index equals or exceeds a predetermined minimum enforcement level H_c and will discharge the worker without payment if the performance index falls short of that level. If the monitoring device is sufficiently accurate ($\varepsilon < 2^{-(1+\gamma)}$), the individual will choose an effort level at which dismissal probabilities will be zero, specifically the equilibrium level of effort will be $H^* = H_c + \varepsilon$.

If the monitoring index is subject to sufficiently large error ($\varepsilon \geq 2^{-(1+\gamma)}$), the dismissal probability will be positive. Specifically, with such monitoring error, equilibrium effort will be

$$H^* = (1+\gamma)\delta,$$

and the equilibrium dismissal rate (p^*)

$$p^* = 1 - (2^{1+\gamma}\varepsilon)^{-1/\gamma}.$$

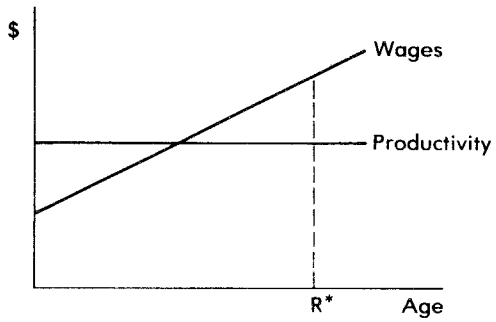
In this simple model, the impact of an increase in the imperfection of effort monitoring (an increase in ε) takes the following forms:

(1) the equilibrium wage if retained increases,
(2) the equilibrium level of effort declines, and
(3) the equilibrium dismissal rate (with no wage payment) increases.

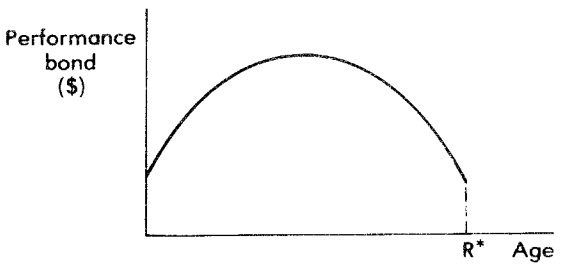
The model is somewhat peculiar in that the firm knows (from its knowledge of the environment assumed here) that all the workers, including those discharged, are performing faithfully, but it must nonetheless discharge workers according to its announced policy, presumably to maintain future credibility.

3.1.3. Performance monitoring and life cycle compensation

In the Becker–Stigler bonding model, variations in reliability with age may induce corresponding variations in the optimal bond and therefore variations in wage profiles over the life cycle. Difficulties in instantaneous monitoring of work



(a)



R*: compulsory retirement age

(b)

Figure 14.2. (a) Life cycle wages and productivity. (b) Life cycle (implicit) bonding.

effort suggest that a multiperiod compensation contract, perhaps even a lifetime compensation contract may be desirable. If the observation lag on effort is long, a deferred, contingent compensation contract may be optimal.

Lazear (1979, 1981) claims that job monitoring may induce a life cycle wage profile similar to the traditional human capital investment profile with wages below marginal productivity early in the life cycle and above marginal productivity later in the life cycle [see Figure 14.2(a)]. The back loaded wage payments provide a performance bond of sorts and Lazear claims that such a profile might therefore be induced by job monitoring considerations.

The Lazear model is an extension of the bonding model developed by Becker and Stigler (1974) discussed above in which the individual posts a bond of sufficient size to guarantee faithful performance of his duties and then receives

severance pay at the end of the period equal to the bond plus accrued interest if he is faithful. Becker and Stigler discuss a model of bonding and life cycle wage profiles but the specific model they develop has no interesting intertemporal linkage and reduces to a series of single period bonds. Lazear develops a more elaborate structure, indeed one sufficiently complex that he is unable to derive a solution or any properties of the solution.[19] In the absence of such results, he argues that the model may yield a wage profile as in Figure 14.2(a).

Consideration of the implied Becker–Stigler bond of such a wage profile makes the Lazear conjecture unlikely. What is the life cycle bond profile implied by a wage/productivity profile of the form illustrated in Figure 14.2(a)? In any period in which productivity exceeds wages, the individual is increasing the size of his performance bond. Conversely, if one ignores interest rate effects (assume $r = 0$), the individual is drawing down his bond in any period in which wages exceed productivity. The linear models of wages and productivity imply a life cycle pattern of bonding quite unlike what one might intuitively expect. In particular the performance bond behaves quadratically, increasing to a peak at midlife and declining after that point [see Figure 14.2(b)]. In such a regime it is the relatively stable age-intervals of midlife that bear the heaviest performance bond. Intuitively one might have expected just the opposite pattern of life cycle bonding with the heaviest bond for (1) young, immature workers with a disproportionately high predilection for shirking and malfeasance and perhaps for (2) older workers with their vulnerability to adverse health shocks. None of the additional considerations introduced by Lazear would seem to lead to the bonding scheme he discusses. A bonding profile of this sort, however, may be appropriate for the bonding of unilateral separations in the presence of specific human capital investments. See Section 4 below and particularly Kennan (1979).

3.1.4. Managerial compensation

The interest of economists in the behavioral consequences of (1) independent agent preferences and (2) incomplete information by the principal on the agent's activities has been a long-standing one, in large part because of its importance in the modelling of managerial behavior and its relationship to the theory of the firm. In the debate over owner versus managerial control, a variety of behavioral models of the firm have been constructed based on the notion that owners, for

[19]A two-period model of the Lazear structure is derived in Lazear and Moore (1984). The model is essentially a generalization of the Becker–Stigler bonding model with the possibility of contract breach by the firm and like the original model fails to provide a monitoring argument for life cycle compensation schemes. These remarks should not detract from the fundamental insight of the original Lazear (1979) paper that in a market economy quantity restrictions on purchase (in this case mandatory retirement) usually imply some pricing difficulty. See Parsons (1984a) for an adverse selection model of the demand for retirement aged workers.

Table 14.7
Incentive share of top management compensation in selected manufacturing and retail firms, 1960–63.[a]

	Total sample (%)	Extreme values deleted[b] (%)
Large manufacturing firms:		
Top executive	86.7	73.4
Top five executives	88.1	73.6
Small manufacturing firms:		
Top executive	91.1	46.0
Top five executives	81.5	50.4
Retail firms:		
Top executives	67.2	61.3
Top five executives	70.6	54.1

Source: Lewellan (1971).
[a]Incentive income was defined as the sum of after-tax stock-based remuneration, after-tax dividend income and absolute after-tax capital gains.
[b]Those with total compensation which deviated from the mean by more than two standard deviations were deleted.

whom profit maximization seems a thoroughly sensible assumption, are unable to monitor perfectly their agents, the managers. As a consequence, a variety of managerial objective functions have been proposed for the firm, e.g. sales maximization [Baumol (1967)] and perquisite maximization [Williamson (1963)]. Each assumes a considerable inability of the firm's owners to control managerial behavior.

The widespread practice of performance-related bonuses for top executives has been held to be a response to this control problem. In an important, early study, Lewellan (1971) reported that performance based rewards were the greater part of the compensation of top executives. In the 1960–63 period, for example, 70–90 percent of a top executive's wealth increments came from stock options and dividends and capital gains from company stock (see Table 14.7). Even when the sample is clipped at ± 2 standard deviations of average compensation to eliminate the handful of extremely large (and possibly unanticipated) capital gain winners and losers, the incentive share of income is in the range of 50–75 percent across the three groups (large manufacturers, small manufacturers, and retailers). In the clipped sample, at least, the incentive share is higher in large firms.

Randomness in the productivity process cannot be neglected in the managerial model. Shocks to firm performance are generated by a variety of factors outside the manager's control, e.g. business cycle fluctuations and more industry and firm-idiosyncratic fluctuations in consumer demand. This randomness adds a new dimension to the optimal compensation calculus. The manager's decisions will be completely efficient only if the manager bears all the risk of the enterprise

(collects all rents) and pays the owner of physical capital a fixed fee. Such a compensation policy, however, raises important questions of preferences for risk and optimal risk bearing if, as is frequently assumed, capital owners are risk neutral and employees risk averse since it implies that the employees bear all the risk of the enterprise. A tradeoff between productivity and efficient risk bearing arises, a tradeoff likely to be more severe the poorer the owner's ability to monitor the manager.

A number of fundamental considerations of optimal incentive systems can be illustrated in an example that did much to mold the early literature in this area, namely share cropping contracts between farmer and land owner, Cheung (1969) and Stiglitz (1974). The land owner is presumed (implicitly) to be able to measure *output* quantity and quality costlessly. Direct monitoring of work effort of the farmer by the land owner is assumed to be prohibitively expensive. The land owner must therefore design an output-based contract that will induce the farmer to work the land efficiently. An obvious incentive efficient contract would be a simple fee rental contract in which the farmer paid the land owner a fixed fee in return for the right to all crops grown on the land for a specified period of time.

Uncertainty of crop production due for example to variations in weather may, however, make the tenant farmer an inefficient residual claimant if he is risk averse. If the presumably wealthier landlord is risk neutral for gambles of this size, he would seem a logical insurance provider to the tenant, paying the tenant a fixed wage and "owning" the crop production himself. This insurance contract is of course the opposite of that proposed for pure incentive purposes. With these conflicting forces at work, the optimal linear contract can be shown to be one in which the farmer will be paid a fixed fee less than his alternative wages and will receive as well a share of the crop production. The fixed wage component of the compensation package will increase with the worker's risk aversion and will decrease with a greater need for effort incentives.

Consider this process more formally. A simple example in the spirit of Berhold (1971) and Stiglitz (1975) is useful in illustrating the interrelationship of the equilibrium employment contract and the economic behavior of the employer and employee. Assume a linear production function with an additive shock,

$$V = \mu H + \theta, \quad \mu > 0, \tag{3.1'}$$

where $V \equiv$ the value of the worker's output, $H \equiv$ the worker's effort level or intensity, and $\theta \equiv$ a random element with $E(\theta) = 0$. Assume further that:

(i) the employer is in a competitive industry and is risk neutral, and

(ii) the worker is risk averse.

More specifically assume that the worker has a utility function of the form

$$U = U(W - RH^2), \qquad U' > 0, \quad U'' < 0, \tag{3.2}$$

where $R \equiv$ a fixed disutility of work parameter and the disutility of work effort increases quadratically with effort and W of course is the wage payment. In this model income will not affect the worker's optimal intensity level.

The compensation and work effort agreement that will result from this market situation will depend on the information available to the two parties. If all desired information is available to both parties and freedom to contract is complete, the risk neutral firm would absorb all production risk and would pay the worker a wage based on the worker's expected production, $W = \mu H$. Confronting this wage gradient, the worker would choose to expend effort H^*, where

$$H^* = \mu / 2R, \tag{3.3}$$

that is effort will rise proportionately with productivity and decrease inversely with the disutility of effort parameter R. The worker's earnings will be W^*, where

$$W^* = \mu H^* = \mu^2 / 2R. \tag{3.4}$$

This efficient earnings–effort pair would be the one chosen for example by the risk neutral self-employed worker.

Consider now a more restricted information environment in which the employer cannot observe work effort directly but can costlessly measure the value of output V. Since eq. (3.1′) can be rewritten as

$$H = \frac{1}{\mu} V - \frac{1}{\mu} \theta, \tag{3.5}$$

it is clear that a perfect effort monitor could be constructed if θ as well as V were known, assuming of course that the employer knew his production function parameter μ. In particular, if the production process is subject to no random element, $\theta \equiv 0$, the monitoring of effort by output would be perfect. Indeed, a compensation scheme of payment according to product would be fully efficient,

$$W = \mu H = \mu \left(\frac{1}{\mu} \right) V = V. \tag{3.6}$$

Payment strictly according to product would yield the efficient level of effort in this model even with an unobservable θ with positive variance since the optimal effort is independent of θ. Such a contract would, however, not generally induce a risk averse worker to accept employment if competitive firms with equivalent technology exist. Such firms could offer different contracts that have equivalent (expected) profit potential yet offer the risk averse worker greater expected utility.

The nonoptimality of strict payment by value of product can be demonstrated by deriving the optimal compensation for a class of compensation rules that

include strict payment by product as a special case, namely the class of linear wage functions,

$$W = \alpha_0 + \alpha_1 V. \tag{3.7}$$

Clearly strict payment by value is the special case $\alpha_0 = 0$ and $\alpha_1 = 1$.

As a first step in the derivation of the optimal linear compensation function, consider the worker's optimal choice of effort when faced with an arbitrary linear compensation schedule as in eq. (3.7). The precise nature of the worker's decision will in general depend on whether he knows the realized value of the random element θ before he decides on the appropriate effort level or not. If he does, the worker is assumed to maximize his utility as in the standard labor supply problem:

$$\max_H U = \max_H U(W - RH^2) = \max_H U(\alpha_0 + \alpha_1 \mu H + \alpha_1 \theta - RH^2). \tag{3.8}$$

If θ is not known to the worker prior to the effort decision, it will be assumed that he maximizes his expected utility:

$$\max_H EU = \max_H \int U(\alpha_0 + \alpha_1 \mu H + \alpha_1 \theta - RH^2) \cdot f(\theta)\, d\theta. \tag{3.9}$$

In this simple model with an additive production error term, the effort level is independent of the realization of θ so the distinction between the known and unknown error term is unimportant. In particular the solutions to (3.8) and (3.9) are the same:

$$H^* = \alpha_1 \frac{\mu}{2R}. \tag{3.10}$$

Only if $\alpha_1 = 1$ in the optimal contract will the equilibrium effort level be equal to the full information effort level [eq. (3.3) above].

Direct substitution of the optimal effort level (3.10) into the worker's expected utility function (3.9) provides a measure of the worker's preferences for compensation rules. Expected utility is appropriate here since the compensation regime is assumed to be determined prior to the employment match and any subsequent realization of productivity. Of course the worker cannot choose among all possible combinations of compensation rules. The employer must be able to break even (achieve zero pure profit) on the employment contract for it to be offered to the worker so expected productivity must equal expected wages for a contract (α_0, α_1) to be offered or

$$E(V - W) = E\left(\mu \frac{\alpha_1 \mu}{2R} + \theta - \left(\alpha_0 + \frac{(\alpha_1 \mu)^2}{2R} + \alpha_1 \theta\right)\right) = 0. \tag{3.11}$$

Since $E\theta = 0$, the zero profit constraint implies that the fixed payment (α_0) and piece rate (α_1) must bear the following relationship:

$$\alpha_0 = \alpha_1(1-\alpha_1)\frac{\mu^2}{2R}. \tag{3.12}$$

The direct substitution of (3.10) and (3.12) into (3.9) yields a single variable problem, say α_1, with a necessary condition for a maximum of

$$\frac{dEU}{d\alpha_1} = (1-\alpha_1)\left(\frac{\mu^2}{2R}\right)\int_a^b U'(\cdot)f(\theta)\,d\theta + \int_a^b U'(\cdot)\theta f(\theta)\,d\theta = 0. \tag{3.13}$$

Since $U' > 0$, the expected value of U' (the first integral) must be positive while the concavity of the utility function insures that the second integral is negative, marginal utility is higher at lower productivity realizations. Thus eq. (3.13) suggests that

$$\alpha_1 \leq 1$$

with strict inequality, unless

(1) the worker is risk neutral, in which case U' is a constant, or

(2) the distribution of θ is degenerate, that is there is no randomness.

In all other cases the optimal contract will involve risk sharing between employer and employee with the contract specifying a fixed payment and partial piece rate compensation schedule. Effort will be less than the complete information contract [compare (3.3) and (3.10) with $\alpha_1 < 1$].

This compensation schedule, it should be stressed, is only the optimal *linear* compensation schedule. Nonlinear compensation schemes may and generally will dominate these linear structures [e.g. Shavell (1979), Holmstrom (1979), and Harris and Raviv (1979)]. The analysis of nonlinear structures, however, requires a substantial increase in the sophistication of mathematical technique, specifically calculus of variation and optimal control, with few substantive implications to date.

Nonetheless examples exist of highly nonlinear optimal compensation schemes. The Harris and Raviv (1979) example, for instance, changes in no significant way if production risk is introduced as a linear additive term in the value of production function. Risk is already present in this structure through the imperfect monitor. Again an absolute standard is asserted to be optimal, estimated effort above a critical level is paid one wage, below that level a different, possibly zero wage (the worker is discharged). Obviously the compensation scheme is nonlinear.

Lazear and Rosen (1981) propose a similar discrete, output-oriented compensation scheme based on an ordinal measure, productivity ranking in a

multiagent situation. Consider, for example, a situation of such complexity that an employer can judge which of two workers has performed more productively, but is unable to assign a cardinal value to the output of each.[20] To provide effort incentives, the employer must offer compensation based on the observable rank order, prizes in the terminology of contests. Lazear and Rosen explore the optimal prize structure in the two agent case. An important new feature is required in the analysis. Since compensation of a given worker's effort is dependent on the effort of his coworker, an assumption must be made of the game aspect of the model, that is how each expects the other to react. In particular, Lazear and Rosen assume a Nash noncooperative game solution in the two agent problem, a rather unattractive assumption when collusion among workers would be terribly valuable to them, but useful for illustrative purposes.

Lazear and Rosen then compare the efficiency of the rank order compensation system with the linear piece rate model and find that circumstances exist under which the rank order contest would be preferred to the linear piece rate system. This result may seem striking since the rank order compensation scheme ignores information used in the piece rate system, namely the actual value of output; the superiority of the rank order contest, when it occurs, must be due to the linearity constraint on piece rates. One could imagine, for example, that the rank order system would approximate the optimal absolute standard model in the Harris and Raviv example better than would the optimal *linear* piece rate model, although the comparison is between two inefficient structures and not in itself terribly interesting.

Lazear and Rosen consider another information structure that would make a rank order contest potentially attractive even if absolute measures of output were available, namely a random shock that is common to all agents. Indeed, Green and Stokey (1983), Holmstrom (1982), and Nalebuff and Stiglitz (1983) demonstrate under varying conditions that a component of the random element common to all agents is necessary if a rank order contest is ever to be optimal when cardinal output measures are available. The reason is intuitively appealing, namely that the rank of the worker's effort, unlike its absolute measure, is independent of the common random shock. Green and Stokey demonstrate in a model similar to that developed above that, with a sufficiently large number of agents (and prizes), the rank order model can approximate arbitrarily closely a nonlinear compensation scheme *with the common element observable*, i.e. the first best solution. Nalebuff and Stiglitz (1983) argue that this conclusion is dependent on the particular form of the production process, the additive error term, and is not in general true. Generally, some cardinal measure of the performance of

[20] Medoff and Abraham (1980, 1981) report several instances of relative performance monitoring among managers.

other agents, for example average output, will be usefully included in the employment contract if a component of the random shock is common to all agents and, as always, if collusion among workers is ruled out.

Indirect measures of environmental shocks may be available that are independent of the actions of the agents themselves and are therefore not vulnerable to collusion. The sales performance of an individual firm's sales staff in a region, for example, is likely to be correlated with aggregate demand measures such as income growth in the area; the crop yield of one's tenants is likely to be correlated with (observable) periods of moisture and temperature, independent of work effort. These measures may not capture industry-specific random effects, but other publicly available statistics may and would be sensibly included in an optimal contract in which individual productivity within the firm is subject to a common random element.

3.2. The compensation of heterogeneous workers

The importance of collecting high-quality information on individual workers, their productivity, reliability, etc. may itself justify long-term employment relationships if worker attributes are in part permanent and not simply a function of current incentives. If observationally equivalent workers systematically differ in the effort they will expend on an activity in a given incentive structure, the firm may adjust its compensation scheme over time as additional information on the reliability of each worker becomes known. In an excellent paper, Freeman (1977) develops a model of life cycle job mobility and compensation in an environment in which productivity assessments accumulate slowly over time. Although the model is based on a research firm example, the model applies equally well to productivity in general.

Freeman imagines a situation in which first period workers are observationally equivalent but in fact differ in the likelihood of making significant discoveries. Freeman assumes that the labor market is composed of two types of workers, high productivity researchers and low, and that the life cycle has two contract periods. After the initial period, workers are distinguished by their observable discoveries. As long as the discovery probability is not zero for the low-productivity individuals, the observed pool of successful discoverers will include a mix of individuals with high probability of success in the second period and low probability. The average probability of success in the second period will be higher among first period successes than among first period failures and the firm will optimally adjust its employment policy to reflect this additional information.

In an auction model of the second period, the firm would simply set wages equal to the expected productivity of each of the two observable groups: the discoverers and the nondiscoverers. If the expected productivity of the nondis-

coverers is less than the value of their time in alternative activities (work or leisure) they would separate from the firm in the second period.

In a long-term contract market, the first period worker may purchase insurance against being revealed as a low productivity worker (through the lack of significant first period discoveries). The second period wages of nondiscoverers may therefore be above auction market levels. The extent of this insurance is limited by the constraint that workers cannot be held to long-term contracts; successful discoverers must be paid the expected value of their second period productivity (or more) if they are to remain with the firm.

The compensation of unsuccessful discoverers need not be in the form of higher wages. Indeed, if an efficient separation policy is to be maintained, the compensation may instead take the form of severance pay or early pension rights. Freeman makes the interesting point that early pension rights to *all* employees is not generally efficient since the best workers may choose to accept the pension and seek work elsewhere unless their wages are raised above competitive levels by the amount of the severance payments, which itself would be inconsistent with the productivity insurance function.

Freeman considers the interesting question of how the research firm's employment policies would be altered if workers know *at the beginning of their careers* whether they are high productivity researchers or not, but that the firm does not. This asymmetric productivity foreknowledge alters dramatically the firm's optimal second period compensation structure in a manner quite similar to that for mobility [Salop and Salop (1976) and Nickell (1976)]. The difference in wage payments between discoverers and nondiscoverers in the second period becomes a self-selection mechanism. Low (or negative) wages among all workers in the first period and among nondiscoverers in the second period and correspondingly high wages among discoverers in the second would make employment in the firm attractive only to high productivity researchers. The asymmetry of information shifts the compensation package toward greater production rewards for the agent with the relevant information.[21]

4. Firm specific human capital

The incidence of job separation and consequently of unemployment is not uniform across the major demographic groups in the work force, but is systematically larger among the low skilled and the young. The strong, negative relationship between job separation and education was noted earlier (Table 14.2). Mincer

[21] Screening devices used to identify high productivity workers may themselves induce a variety of activities that would not be undertaken were information free. See Spence (1973) on the general question and Akerlof (1976), Miyazaki (1977), and Guasch (1983).

and Jovanovic (1981) report data from the National Longitudinal Surveys of Young Men that indicates that the annual separation rate drops by 90 percent from the first year to the sixth year with the firm (see Table 14.8 below).

The firm specific human capital hypothesis has been proposed as an explanation of these and other turnover patterns. As noted earlier (Table 14.3) firms (and worker) may have relatively heavy investments in their employees. The business firm, after all, is more than a physical plant and pieces of equipment. Imagine a new firm which is about to commence operation. A variety of human capital expenses must be incurred. Hiring costs (advertising, interviewing, etc.) will be generated if new workers are to be attracted to the firm. Screening costs will arise if worker quality is not immediately observable. Optimal job assignments are rarely obvious a priori. Allocating workers to jobs typically requires substantial trial and error with a corresponding loss of productivity. Over time a new worker will learn about his job, about the characteristics of other workers in the firm, and about the nature of the firm's markets or individual customers, as well as the reliability of suppliers of various factors. These investments are specific to the unique match between a firm and worker and will be lost if the match is broken.

Even among established firms employee separations (and the need to invest in replacements) are inevitable, whether through death or through economic turnover such as retirement, discharges, quits and layoffs. A successful, long-lived firm must establish employee compensation policies that economize on the rate of depreciation of this organizational capital. The employment policies undertaken by individual firms may have important implications for the performance of the economy through the balance struck between (1) the need for a fluid, mobile labor force to accommodate inevitable fluctuations in product demand and (2) the need for work force stability to protect heavy investments in firm specific capital.

4.1. Job attachment in a homogeneous work force

4.1.1. Specific human capital, job turnover, and feasible contracts

Large investments in a unique relationship between an individual and a single firm will almost surely reduce the efficient rate of separation between the two agents and is therefore likely to reduce actual separation rates in any but the most unfavorable contracting environments. Contracts, explicit and implicit, are important in the efficient resolution of this process since match-specific capital lacks the usual competitive labor market guarantees that more widely demanded skills carry. In this section, the interrelationship of firm specific investments, job attachment, and feasible contracts is specified and their implications for life cycle mobility are discussed. Finally, the impact of (mobility) heterogeneous workers

and imperfect information on the form of the firm's compensation package is developed for a market in which specific human capital investments are heavy.

The original theoretical literature in this area [Becker (1975), Oi (1962), Parsons (1972), Pencavel (1972), and Salop (1973)] focused on the effect of firm specific human capital on firm compensation policies and their consequences for firm layoff and quit experiences. More recently the literature has focused on a simple but expositionally powerful model of individual decision-making and randomness in job matching and separation [Mortensen (1978), Hashimoto and Yu (1980), and Hashimoto (1981)]. Contract effects on behavior are well illustrated at the individual level so this case forms an excellent introduction to the specific human capital investment process.

Single period efficiency requires that the individual be matched with the activity that is the highest valued use of his time. An employment contract that specifies complete worker immobility subsequent to specific human capital investment will typically not be efficient since worker productivities both inside the firm and out are likely to fluctuate randomly over time. The firm may experience product demand reversals, the worker sudden increases in the value of his general human capital in alternative firms. In both cases efficiency may require a reassignment of the individual to a different job.

Consider, for example, a simple model in which the worker and firm must undertake an organizational capital investment if the worker is to be an efficient employee.[22] Assume that the worker's productivity in the firm and his productivity in the labor market are subject to random shocks but that the investment must be undertaken prior to the discovery of these random effects. In particular assume that

$$V_i = \mu_i + \theta_i, \quad i = 0,1, \tag{4.1}$$

where V_i denotes productivity; μ_i a fixed parameter; and θ_i the random productivity element with $i = 0$ the specific firm in question and $i = 1$ all other firms. Assume that $E(\theta_i) = 0$ so that the μ_i reflect "permanent" productivity differences. Presumably $\mu_0 > \mu_1$ by an amount sufficient to make the investment in the employee profitable on average.

The efficient separation policy is transparent. The firm and the worker should agree to separate if and only if the worker's actual productivity in the firm is less than his productivity outside the firm,

$$V_0 < V_1$$

or

$$\mu_0 + \theta_0 < \mu_1 + \theta_1. \tag{4.2}$$

[22]Again this model is in the spirit of those developed by Mortensen (1978), Hashimoto and Yu (1979), and Hashimoto (1981).

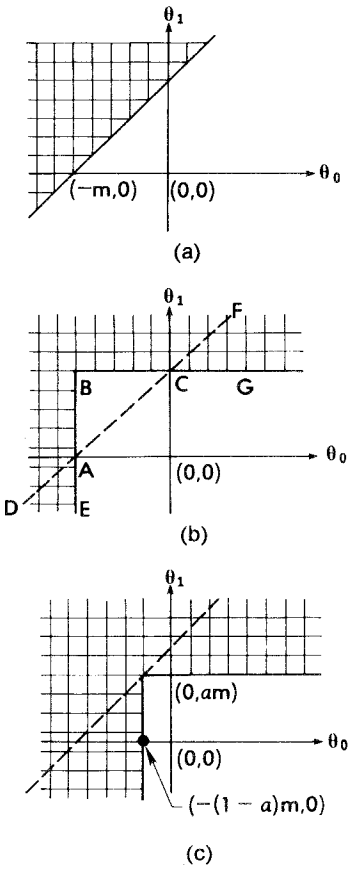


Figure 14.3. (a) The efficient job separation region. (b) The job separation region with bonding. (c) The job separation region with investment sharing.

Defining m as the difference between the expected productivities in the two activities ($m \equiv \mu_0 - \mu_1$), efficient separation will occur whenever

$$m < \theta_1 - \theta_0. \tag{4.3}$$

Unusually high productivity in the outside sector ($\theta_1 \gg 0$) and unusually depressed conditions within the firm ($-\theta_0 \gg 0$) will induce efficient separation.

Graphically the combination of inside and outside productivity shocks that will optimally induce job separations is represented by the cross-hatched area in Figure 14.3, part (a). The critical boundary of the separation–attachment region, of course, is determined by inequality (3.3). The optimal separation rate in this

environment will be determined by the frequency distribution of outcomes of θ_0 and θ_1. In a bivariate normal distribution with θ_0 and θ_1 independent and with $E\theta_0 = E\theta_1 = 0$, an increase in the variance of either variable will induce greater efficient separation rates. If, as seems plausible, θ_0 and θ_1 are positively correlated so that productivity in the economy and in the firm fluctuate together (as might occur in a business cycle) then efficient separations will be less. If they are perfectly (positively) correlated no efficient separations may occur.

The efficient rate of job separation depends completely on the difference in expected productivity between the two activities, m, and of course on the size of random shocks, θ_0 and θ_1. The expected margin in productivities is of course not exogenous to the system but is determined by the size of the investment cost, say c. In the case of permanent occupational choice (no separations) in a market with risk neutral economic agents, the investment must provide an expected return (productivity margin) such that

$$\mu_0 - \mu_1 \geq c. \tag{4.4}$$

In a competitive product market with free entry, the relationship will hold as an equality. It is easily demonstrated in this model that larger specific human capital investments will generate lower equilibrium separation rates.

An important implication of this analysis is that no meaningful causal distinction exists between layoffs and quits. The separation rate is a function of the joint distribution of productivity shocks to the firm and to the economy. Job separation conditions are mutually agreed upon, based on complete information on the nature of these shocks. Turnover will occur only when job mobility is efficient because it is in the interest of both parties to agree to such a contract and information and contracting conditions are such that all desirable contracts are attainable.

The informational requirements of an efficient separation contract are large. Hashimoto (1981) proposes a "profit" sharing contract in which factor payments are determined as a share of the difference between current firm and alternative firm productivity. Defining the factor payments to the worker as w (wages) and to the firm as π (profits) the distribution rule would be

$$w = V_1 + \alpha(V_0 - V_1) = V_1 + \alpha m + \alpha(\theta_0 - \theta_1) \tag{4.5}$$

and

$$\pi = V_0 - w = (1-\alpha)(V_0 - V_1) = (1-\alpha)m + (1-\alpha)(\theta_0 - \theta_1), \tag{4.6}$$

where α is a predetermined sharing rule, $0 < \alpha < 1$. Clearly, under this rule each would choose to separate if and only if $m + (\theta_0 - \theta_1) < 0$ or, as in (3.3), $m < \theta_0 - \theta_1$, for then $w < V_1$ and $\pi < 0$.

For such an agreement to be explicitly contracted, an unrelated third party must be able to observe the random elements θ_0 and θ_1 as well as the permanent factors μ_0 and μ_1 and as a consequence m. Such observability is unlikely, even for the two parties actively engaged in the contract. If, as is plausible, only the firm has good information on θ_0, the random shock in current firm productivity, and only the worker on θ_1, the random element in the worker's outside alternatives, then the contract as stated is not feasible. Other mechanisms also yield the efficient separation rate but have the same information requirements. Mortensen (1978), for example, considers an "offer matching" option to the contract that would achieve efficiency in this model, the other party would be willing to bid up the return to the party that wishes to leave to the point that only efficient separations would occur. Clearly the bidding party needs information on the offer the other party receives if strategic games are to be avoided. Indeed in the more general model discussed by Mortensen that incorporates endogenous search activity, search intensity must be observable as well or offer matching will promote excessive search and separation.

An alternative compensation mechanism that requires less information is one that specifies fixed wages and a separation bond equal to the other's capital investment in the match.[23] In this compensation scheme each party's investment is protected. The worker, for example, will quit only if he receives an alternative wage bid greater than his contract wage by the amount of the firm's investment in him and the firm receives the forfeited bond as compensation. Assuming the worker's own investment is reflected in his wage (as must be the case in a competitive labor market), the worker will leave only if his alternative bid exceeds the total investment costs.

Formally the worker will quit when

$$V_1 > w + B_w, \tag{4.7}$$

where B_w denotes the bond the worker must forfeit upon departure and the firm will lay off the worker whenever

$$V_0 < w - B_F, \tag{4.8}$$

where B_F denotes the firm's bond it forfeits upon laying off the worker. Assuming the expected return margin m is equal to investment cost c [see eq. (4.4)] so that $m = c$, then

$$B_w = (1 - \alpha)m \tag{4.9}$$

[23] Mortensen (1978) formally considers such a structure.

and

$$B_F = \alpha m, \tag{4.10}$$

each bonds the other's share of the original investment. Recall α is the worker's share of the investment. If the worker is risk neutral and in a competitive labor market, he must be offered a fixed wage equal to the expected alternative wage plus the return on the worker's investment, or

$$w = EV_1 + \alpha m = \mu_1 + \alpha m. \tag{4.11}$$

This fixed compensation–bonding contract is not fully efficient because the returns to the investment are random.[24] The separation decision rules (4.7) and (4.8), following the substitution of the relationships in (4.9)–(4.11) reduce to the rules that the worker should quit if

$$\theta_1 > m, \tag{4.12}$$

that is, if he experiences an outside demand shock greater than the full investment cost, and that the firm should lay off the worker if

$$\theta_0 < -m, \tag{4.13}$$

that is, if his productivity in the firm drops by more than the full investment cost. The separation region of the (θ_0, θ_1) plane under this set of decision rules is illustrated in part (b) of Figure 14.3. The boundary of the efficient separation rules is marked by the dashed line for comparison purposes.

In some circumstances [realizations of (θ_0, θ_1)], bonded individuals will separate from the firm when they should have remained matched (triangle ABC). In others they will remain with the firm even though they should optimally have separated (the regions ADE and CFG). The reason the fixed bond model is not fully efficient is that the total value of the match is a random variable and the fixed bond model gives the other party no opportunity to express actual as opposed to prospective match value.

The information requirements of the bonding scheme are, moreover, not zero, particularly if the scheme must be administered by a third party. The bonding administrator must be able to identify which party wishes to end the match if the appropriate bond is to be forfeited. As a practical issue this decision is more

[24] The bonding model is fully efficient in the discussion by Mortensen (1978). He assumes a Poisson arrival process for random shocks to current productivity and alternative activity and considers an instant in time during which no more than one shock in total can occur. He further assumes a fixed and known cost to the party *not* receiving the shock of breaking the job match. A fixed value bond is efficient in this situation.

difficult than one might expect (unless one were familiar with the problems that confront unemployment insurance administrators in systems that distinguish by eligibility or benefit level between layoffs and quits). Both parties can alter other, less observable aspects of contract performance to induce the other to break the match. A firm can make a worker's life quite intolerable by varying the non-pecuniary returns of the job. The worker on the other hand can induce the firm to break the match by an appropriate display of sloth or ineptitude.

The information problem required for bonding is eliminated if one of the two parties is reliable, for both parties presumably know whether or not they are themselves responsible for inducing the separation. The reliable party could therefore hold the bond of the less reliable party, keeping the bond if that is appropriate or returning it, with or without its own forfeited bond as the agreement stipulates.

Becker (1975) proposes that nonvested pensions serve as such a bonding scheme, an idea worked out more formally by Kennan (1979). If the bond is substantial, of course, the worker may confront borrowing constraints that force the bond to be "posted" or accumulated over time with the relatively full efficiency of the bond secured only after some period with the firm. McDonald (1982) models the interrelationship of specific human capital and capital difficulties.

Earlier papers on specific human capital finance and turnover [Becker (1975), Oi (1962), and Parsons (1972)] explored a weak form of employment agreement, a fixed wage contract with no separation damages. No information problems arise in this structure since the payments are not contingent on the realizations of possibly unobservable market processes. In this compensation environment, the worker will quit whenever

$$V_1 - w > 0, \tag{4.14}$$

and the firm will lay off the worker whenever

$$w - V_0 > 0. \tag{4.15}$$

In a competitive labor market, the wage would be set as earlier, as the expected value of the alternative wage plus the worker's share of the investment costs ($w = \mu_1 + \alpha m$). The separation region for a given investment sharing parameter (α) is illustrated in Figure 14.3, part (c). Again the efficient separation boundary in the complete information case is indicated for comparison purposes by the dashed line. The fixed wage separation region is unambiguously larger than the efficient separation region.

Becker (1975) first proposed that there may be some optimal investment sharing (α) that will minimize this inefficient separation. The financing of and

returns to the investment will be shared between firm and worker since some investment by each in the relationship reduces the incentives for unilateral withdrawal from the relationship by either; the sharing is a form of mobility bond. The worker's investment in the activity provides the firm with a greater likelihood that the worker will not quit the firm after the investment period and thereby depreciate its investment in the worker's organizational capital. Similarly, the worker has greater security that he will not be subject to permanent layoff (and the depreciation of his organizational human capital) if the firm invested in him as well.

In the framework of the individual's quit decision [part (c)], the sharing should be undertaken such that the separation region *outside* the efficient separation region should include as little probability density of (θ_0, θ_1) as possible.[25] It is easily demonstrated that, if the variances of θ_0 and θ_1 are equal (the size of the random shocks to the individual's productivity inside the firm and outside are equal), then equal sharing of the investment will minimize the separation inefficiencies due to the nonobservability of θ_0 and θ_1 by both parties. As the variance of θ_0 increases relative to that of θ_1, the firm should optimally undertake an increasing share of the investment if inefficient separations are to be minimized.

The sharing hypothesis requires for its execution no information on realized values of (θ_0, θ_1) since a fixed wage contract underlies it. For the same reason, it will not be fully efficient. Hashimoto and Yu (1980) argue that, even though the realizations of θ_0 and θ_1 may not be observable, flexible wage contracts contingent on observables correlated with θ_0 and θ_1 will generally be desirable, that is will be more efficient than the fixed wage contract. As the correlation between the proxies for θ_0 and θ_1 and their corresponding true values increase, the contingent contract will approach the fully efficient separation behavior. Certainly it is easy to imagine aggregate observables such as vacancy rates, want ad lineage, etc. that proxy more or less well fluctuations in the worker's wage prospects. Similar data may also provide good indicators of firm demand conditions. Other potentially useful data include industry sales and, if the firm is a large, publicly held corporation, actual firm sales and profits.

Parsons (1972) develops formally Becker's sharing hypothesis in a two period framework in which the first period is devoted to specific training and the second to production. He reports a number of comparative static results. The employer's share of the investment will be higher (1) the less responsive worker quit rates to wage changes, (2) the more responsive trainee supply to wage variations, and (3) the higher the worker's rate of discount. The firm's layoff propensity will be greater the lower unemployment insurance costs and required severance pay-

[25] More precisely the expected loss should be minimized which involves the size of a specific misallocation as well as its probability.

ments, the greater the likelihood of eventual rehire, and the more insensitive quits are to wage variations. Independently Pencavel (1972) and Salop (1973) developed elegant optimal wage policies for the zero layoff environment and derived many of the same quit rate implications.

The best evidence to date on the specific human capital sharing hypothesis suggests that the Becker limited information model is valid. Quits (worker initiated job separations) are affected by the worker's share of such investments [Pencavel (1972), Parsons (1972)] and layoffs by the employer's share [Parsons (1972)]. Nonetheless, the measurement of specific human capital is uncertain and the measurement of worker/firm shares of the capital verges on the speculative. The worker's investment in specific human capital (and therefore his incentive to quit) is typically indexed by the difference between actual wages and predicted wages (or proxies for the latter such as schooling) and the firm's investment (and therefore its incentive to lay off the worker) by the difference between total human capital and wages. Clearly this aspect of the firm's organizational capital formation remains very much an open research question.[26]

4.1.2. Life cycle separation behavior

If job matching costs are incurred only at the time of hire, say as job search and mobility costs, then one would expect that separation rates would remain constant over the life of the job match. The wage–separation models discussed above [Pencavel (1972), Salop (1973)] have such a property. If, however, firm specific human capital investment occurs, then one would expect an individual's separation rate to fall over time. Mincer and Jovanovic (1981) provide persuasive evidence that separation rates decline sharply with length of current job tenure and that what appears to be age effects on separation behavior are primarily a reflection of the tenure dependence of separations. In Table 14.8 two-year separation probabilities of young males from the National Longitudinal Surveys, reported in Mincer and Jovanovic, illustrate the tenure dependence of separation probabilities. Whether the individual has zero-to-four years of total work experience or five-to-nine, the separation rates drop from more than 70 percent in the first year on the job to less than 10 percent in the fifth year and beyond. Similar, if lower, separation patterns hold for older males.

The precipitous decline in mobility in the first few years could well be due to factors other than specific human capital, at least in its simplest form. Intuitively one might suppose that the individuals are in part "job shopping", learning about attributes of the firm that are not readily observable prior to working with

[26] Elizabeth Peters (1983) presents an interesting study of efficient separation in a small firm, in this case the family. Peters found that the wide differences across states and in divorce laws affected the distribution of assets after the marriage dissolved but had no significant effect on the incidence of divorce.

Table 14.8
Job mobility among young males by experience and tenure, pooled 1967–73 (percent moving in a two-year period).

Total work experience (years)	Tenure (years with current job)					
	0–1	1–3	3–5	5–7	7–9	9–11
			Total separations			
0–4	0.73	0.58	0.28	0.07	0.12	0.04
5–9	0.77	0.60	0.38	0.08	0.07	0.06
			Quits			
0–4	0.48	0.41	0.20	0.06	0.08	0.04
5–9	0.48	0.42	0.26	0.06	0.05	0.07
			Layoffs			
0–4	0.26	0.17	0.08	0.01	0.05	–
5–9	0.28	0.18	0.11	0.03	0.02	–

Source: Mincer and Jovanovic (1981, p. 25). The basic data set is the younger male cohort of the National Longitudinal Surveys. The total sample size is 3443.

the firm, experience goods in Nelson's terminology [Nelson (1970)]. The experience attribute could be the nonpecuniary nature of the job, e.g. coworkers, bosses, and job stress [Johnson (1978), Viscusi (1980)] or the worker's productivity in the firm [Jovanovic (1979)]. Burdett (1978) proposes a pure search model in which equivalent workers search their way to better and better jobs over time. If jobs are guaranteed at a fixed wage for a lifetime, then workers will face an upward drift in wages and a decline in the quit rate over the life cycle; adversity is assumed away. Mortensen (Chapter 15 in this Handbook) reviews this model at some length.

The tenure patterns, however, hold for layoffs as well as quits (Table 14.3), suggesting that the phenomenon is more than simply a reflection of job shopping and that the job match does grow in value with time with the firm. Indeed, the one empirical, if crude attempt to unravel these various factors (Mincer and Jovanovic) indicates that perhaps a third of the tenure dependence phenomenon is due to worker heterogeneity (individuals with intrinsically lower separation probabilities have disproportionately longer job tenure, ceteris paribus). Further, they decomposed life cycle wage growth by source: one half is general on-the-job training, one-quarter the worker's share of specific human capital investments, and one-quarter job search related, primarily real productivity effects, i.e. the individual finds a more productive match (pure wage search factors do not appear to be a major factor in life cycle wage profiles).

Jovanovic (1979) develops an interesting model that combines alternative wage search (wage growth by job change) and specific human capital investment (wage growth within the firm). The model is sufficiently complex that it has not been

systematically exploited but suggests a number of important interactions. On the one hand, search and consequently mobility decline with greater amounts of accumulated specific human capital (Jovanovich assumes complete and efficient contracting here). On the other hand, the attractiveness of specific human capital investments may change with the optimal level of productive search. If the intrinsic value of a match is not immediately evident to a worker, specific investments may not be undertaken at once (if production technology offers a choice) but may instead be undertaken only after active job shopping is past.

4.2. Employment contracts with mobility – heterogeneous workers

If workers differ intrinsically in their job mobility propensities, efficiency would surely require, ceteris paribus, that low mobility workers be matched with jobs in which large, firm-specific investments are profitable. If low-mobility individuals are easily identifiable, by themselves and by firms, this mobility consideration could be easily handled within a competitive labor market, job mobility propensity would simply be another attribute of the individual worker. If, as seems plausible, the worker but not the firm is aware of the individual's mobility propensity, however, the question arises of whether the firm can design a compensation scheme that will cause the worker to reveal his perceived characteristics. In the insurance literature, for example, Rothschild and Stiglitz (1976) consider the question of whether high-risk individuals can be induced to reveal to the insurance firm their riskiness by manipulation of price and extent-of-coverage tie-ins, high-risk individuals being induced to purchase more complete coverage and to pay disproportionately heavier premiums.

Salop and Salop (1976) and Nickell (1976) independently proposed alteration of the tenure structure of wages as a self-screening device which would induce highly mobile individuals to identify themselves. In particular, if firms with heavy specific investments offer low wages in the first period and high wages in the second, the expected compensation stream should be relatively attractive to those with low mobility. One could presume that an appropriately steep wage gradient could separate the workers, if no other factors intervened. Indeed, Salop and Salop argue that the optimal wage pattern would be one in which the individual worker's first period wage would be entirely reduced by the amount of the specific investment and raised correspondingly in the second period (or future periods in a multiperiod model). The logic of the argument is transparent. In a world in which only the worker can observe his innate mobility trait, the worker must bear all the costs of the specific human capital investment for the incentives to be perfect. Obviously this solution ignores the firm/worker contracting problems discussed above. Moreover, as Nickell points out, such a solution ignores other plausible sources of heterogeneity. Specifically workers may sys-

tematically differ in access to borrowed funds. This source of heterogeneity is particularly significant since the optimal mobility-revealing strategy is to force the worker to make all specific investments. Random differences in access to capital would make the wage gradient an imperfect selection device, the more so the larger the randomness in discount rates and the less positively correlated these rates are with immobility traits. Nonetheless, these self-selection arguments provide an important, additional reason why workers may share in specific human capital financing.

5. Employment insurance contracts

Empirically wages appear to fluctuate little in response to relatively large fluctuations in product demand. Employment levels also fluctuate proportionately less than do sales and the employment adjustments that do occur are quite slow.[27] Nadiri and Rosen (1974), for example, report that in the manufacturing sector only about 40 percent of the adjustment in production worker levels to a decline in sales occurs in the first quarter, with the full adjustment requiring five to six quarters. The estimated adjustment among nonproduction workers is minimal even in the long run. The large procyclical fluctuations in average productivity are apparently a reflection of the "labor hoarding" that occurs during economic downturns.

As discussed above, the specific human capital hypothesis has predictions for both the existence and incidence of the observed labor hoarding consistent with these observations. A number of papers in the macroeconomic literature, especially Baily (1974), Gordon (1974), and Azariadis (1975), have proposed an alternative explanation, that employment stability is directly purchased from the firm as a part of the employee's compensation package. Among experienced workers, job layoffs seem to be the primary force driving cyclical unemployment rates (see above Figure 14.1). Given the potentially large earnings losses generated by any significant period of unemployment, it is natural to imagine that layoff probabilities under various demand conditions are an important aspect of a worker's compensation.

Since the early literature had a strong macroeconomic orientation, it focused primarily on the implication of the worker's earnings security demands for wage "stickiness"; only recently have the full implications of the models for labor hoarding become fully appreciated [e.g. Akerlof and Myazaki (1980)]. In the same tradition, papers by Calvo and Phelps (1977), Hall and Lilien (1979), Grossman and Hart (1983), Azariadis (1983), Chari (1983), and Green and Kahn

[27] Beyond the Nadiri and Rosen work discussed below, see the earlier work by Brechling (1965) and Ball and St. Cyr (1966) and also Clark (1973).

(1983) consider employment contracts under more complex informational assumptions. In this section I explore the motivations and consequences of risk reallocation in employment contracts within the contract framework developed above.[28]

The literature discussed in this section focuses on unemployment as the alternative activity to working for the current firm. All studies assume that unemployment has some value, at the very least respite from work effort and related stress. Nonetheless, the unemployment activity is likely to have a value considerably less than customary wages if not less than current productivity. In the absence of earnings insurance, the loss of any substantial amount of labor income is likely to have severe, negative consequences on family well-being, since for the majority of families in the United States, earnings are the sole source of income of any consequence. Assets, particularly relatively liquid assets which can be cheaply converted into consumption, are few. The ability to borrow against future earnings to finance current consumption is severely restricted as well.

The traditional private insurance market for earnings security is primarily limited to life insurance; private disability insurance is not widespread and unemployment insurance essentially nonexistent. Reasons for the absence of a vigorous, private market for unemployment insurance are not difficult to enumerate:

(1) the difficulty of state verification,
(2) moral hazard, and
(3) adverse selection.

Each of these is a serious problem in the provision of unemployment insurance.

Consider the question of state verification; that is, did the insured event in fact occur? Particularly at the lower skill levels at which unemployment is a considerable risk, an unrelated, private third party may have considerable difficulty determining whether or not an individual is currently employed. Even the government tax collectors with their unique access to information and their tremendous array of coercive powers have significant problems monitoring the earning activity of this group. The difficulties for private insurers are surely more substantial.

Moral hazard and adverse selection problems compound an insurer's problems. Moral hazard problems of two kinds may arise: layoffs (entry into unemployment) may be excessive as firms and workers treat unemployment insurance as subsidized leisure [Feldstein (1975)]. The duration of unemployment may be increased as well if the laid off workers reduce the intensity of their search for alternate jobs. Adverse selection is also likely to be a problem since individuals

[28] Excellent reviews of this literature are available. See Azariadis (1981) for a fine discussion of the original rigid wage literature and Hart (1983) and Azariadis and Stiglitz (1983) for critical discussions of the more recent literature on employment contingent wage contracts.

are likely to have better prior information on their prospects for layoff and unemployment than are private insurers of those individuals.

The public provision of unemployment insurance eliminates some but not all of these insurance difficulties. Adverse selection problems are eliminated in a universal system. Access to information unavailable to private insurers reduces but does not eliminate the state verification problem. Still the moral hazard problems remain. With incomplete experience rating, firms and workers have an incentive to treat unemployment insurance as subsidized leisure; workers have a reduced incentive to seek alternative employment, the firm an increased incentive to put them on layoff status. No doubt as a consequence of these insurance difficulties, public unemployment insurance is incomplete. Many unemployed workers are not covered by the system. Of those that are, few receive payments that come close to replacing lost earnings, although they receive some compensation through additional leisure.[29]

The limited private and public availability of earnings insurance suggests that the employer himself may find it profitable to provide some level of protection to the worker against unemployment earnings losses. Certainly the employer has access to higher quality information, both on his own intentions and that of his employees, than do private insurance firms or even the government. One aspect of state verification, whether the individual has been laid off, is obviously known to the employer. Adverse selection is not likely to be a serious problem since again the firm is aware of future layoff probabilities. Some aspects of the problem do remain outside its control; the firm is unlikely to know for example whether the laid off worker is seeking alternative employment diligently or indeed whether he has already found alternative employment.

The ability of the firm to provide employment insurance is dependent on the information available to the agents. Under what conditions will insurance be a feasible part of the employment contract? How much will the worker's earnings fluctuations be moderated? Perhaps most importantly, how will employment behavior be affected in the optimal contract? The models discussed below are the first important steps in answering these questions.

Whatever its specific content, the earnings insurance "contract" will presumably be an implicit one. If the various contingencies were observable at reasonable cost to third parties as is necessary for explicit, legally enforceable contracts, the insurance function would be undertaken by specialized insurance firms. Employers, for example, typically provide life insurance with its rather fewer ambiguities through insurance firms and are not themselves the primary insurer. As with any implicit contract, the reliability of the agents as well as their differential access to information is an important factor in forming the agreement and in determining subsequent behavior.

[29]See Topel and Welch (1980) for a review of the literature on public unemployment programs.

The attitudes of the agents toward risk, only briefly considered in the specific human capital discussion, becomes crucial in the analysis of employment of earnings insurance contracts. Since the purpose of insurance is to reallocate risk bearing, the attitudes of the parties toward risk and the cost to each of diversifying this particular risk are crucial to the bargain. The most common assumption is that the worker is risk averse, that the employer behaves as though he is risk neutral, and that as a consequence the employer sells insurance to the worker within the constraints of the information set available to each. The risk neutrality of employers is based on either of two arguments, that capital owners as a group are wealthier and less concerned by a given wealth fluctuation or, more compellingly, that physical capital owners can diversity their portfolios more cheaply than can human capital owners.

However plausible, the assumed risk preferences of the two groups are based on empirical conjecture and indeed models have been developed and are discussed below which assume the converse. One should note that the diversification argument has implications that the simple difference-in-risk-preference argument does not. One would expect that diversification among owners and therefore the willingness to insure workers is more complete in large, publicly held firms. Secondly, aggregate risk due to business cycle fluctuations is not diversifiable and is therefore less likely to be insured by the firm than are random shocks to the firm or industry if the two agents' attitudes toward risk are similar.[30]

In the section that follows I consider first the random firm specific shocks to worker productivity that are most amenable to internal insurance. The nature of the employment contract under various assumptions of information availability and contract reliability is explored. The difficulties and behavioral changes that arise with aggregate, nondiversifiable risk are then considered.

5.1. Individual productivity risk

In this section I assume that the productivity of the individual in the current firm and elsewhere in the economy are subject to individual-specific random shocks that are unrelated to aggregate or business cycle fluctuations. In such an environment it is natural to assume that employers act as if they are risk neutral and therefore that they act as insurers of risk averse workers if information problems are not too severe. This assumption captures the intuitive notion that owners can more easily diversify their physical capital holdings than can workers diversify their human capital position in the firm.

[30] See Reagan and Stulz (1982) for a more extended discussion of labor contracts with aggregate and disaggregate risk.

5.1.1. *Job attachment with complete information*

Consider again the simple two activity model of the previous section:

$$V_i = \mu_i + \theta_i, \quad i = 0,1, \tag{5.1}$$

where V_i denotes productivity, μ_i a fixed productivity parameter, and θ_i the random shock to productivity with $i = 0$ the current firm and $i = 1$ the alternative activity. The alternative activity could be another job or unemployment (leisure, home productivity, job search, etc.). Assume that $E(\theta_0) = E(\theta_1) = 0$, so that $m = \mu_0 - \mu_1$ reflects the permanent difference in productivity between the two activities; in this case m need not be positive although it presumably would be whenever specific investments or mobility costs had been incurred.

The efficient separation policy in the risk neutral (expected productivity maximization) case is, as discussed above, to separate if productivity in the firm is less than productivity outside the firm. Symbolically:

$$V_0 < V_1$$

or

$$\theta_1 - \theta_0 > \mu_0 - \mu_1 \equiv m.$$

In a world of risk neutral competitive firms and complete information, earnings insurance will be provided at zero cost and the efficient separation rule is unaffected by worker risk aversion.

Not all the efficient separation contracts discussed in the previous section will be optimal, however, when workers are risk averse since different contract forms have different implications for earnings stability. In the profit sharing contract, for example, in which wage (W) is set according to the rule:

$$W = V_1 + \alpha(V_0 - V_1).$$

Wage payments to the worker have a nonzero variance, namely

$$V(W) = \alpha^2\sigma_0^2 + (1-\alpha^2)\sigma_1^2,$$

where $\sigma_0^2 = V(\theta_0)$ and $\sigma_1^2 = V(\theta_1)$ and by assumption $\text{cov}(\theta_0, \theta_1) \equiv 0$.

In the complete information case, risk neutral employers can (profitably) guarantee workers a fixed wage payment, independent of outcome. In the competitive case with equal expected value for either outcome ($\mu_0 = \mu_1 = \mu$) the payment would be μ. The worker would remain with the firm or separate as efficiency required ($V_0 \gtrless V_1$); he would receive a fixed wage $W = \mu$ if he remained

with the firm. If separation is efficient, appropriate side payments would be made from firm to worker if $V_1 < \mu$ and from worker to firm if $V_1 > \mu$.

The courts, of course, might refuse to enforce a requirement of severance payments from worker to firm, although the same transaction could perhaps be designed less transparently. For example, an implicit bond could be posted that would be forfeited to the firm upon a job separation at the worker's initiative. In a multi-period framework this could be executed through low entry wages and high subsequent compensation (seniority-based wage increases, noninvested pensions, etc.).

Conversely, the firm may have some limitation on its ability to pay severance benefits to workers it chooses to lay off, perhaps because of moral hazard problems due to reduced work effort by workers desiring to induce layoff and consequently receive severance benefits payments. Such an economic environment underlies the now famous set of papers on wage rigidity and employment contracts of Baily (1974), Gordon (1974), and Azariadis (1975).[31] The mechanism underlying these models can be illustrated in a special case of the two activity choice models. Label the second activity unemployment. Presumably the value of this activity (V_1) is a function of the value of leisure or job search time, government unemployment compensation payments, etc. Assume that the value of unemployment is fixed, $\sigma_1^2 = 0$ or $V_1 = \mu_1$, and, though plausible but not essential, that the value of unemployment is less than the mean value of productivity in the firm, $\mu_1 < \mu_0$. Assume further that V_0 is distributed over the interval $-\infty$ to ∞ according to the density function $f(V_0)$.

It is easy to demonstrate formally that, with complete information and no contract constraints, the risk neutral firm will offer the risk averse worker an employment contract with the following features:

(1) the firm will lay off the worker if and only if $V_0 < \mu_1$, that is the critical current productivity threshold for separation, say V_c, is $V_c = V_1 \equiv \mu_1$;

(2) the worker will be offered a wage (if retained) that is independent of the productivity realization V_0; and

(3) the severance payment, say K, will be such that the "worker" is equally well off whether employed or unemployed: he is fully insured.

The precise level of the wage–severance-pay package will depend on the value of compensation available elsewhere, that is, the compensation necessary to attract workers to this firm's contract prior to realization of actual productivity.

What now if the firm is unable to offer severance pay or (loosely) unemployment compensation? The effect on the firm's separation policy is unambiguous; the firm will enter into an employment contract with the worker which specifies that the separation rate will be less than efficiency considerations alone would dictate. The layoff threshold V_c will be strictly less than the alternative value of

[31]Again see the review of these and subsequent papers in Azariadis (1981).

the individual's time ($V_c < \mu_1$). The reason is transparent; because of the severance pay restrictions, the firm can insure the worker against adverse demand fluctuations only by retaining him on the payroll. As long as the real cost of doing so is not too great, that is V_0 is not too much less than V_1, such a contract is attractive. The result is a quite plausible motivation for "labor hoarding" independent of specific human capital considerations.

Ironically, the pioneering literature in this area, again Baily (1974), Gordon (1974), and Azariadis (1975), stressed the wage rigidity aspect of this structure; the "overemployment" implication was not stressed [Akerlof and Myazaki (1980)]. Although the overemployment implication is quite consistent with what we know of employment behavior, it has a rather obvious deficiency in explaining unemployment, an objective of the early studies.

5.1.2. *Asymmetric information*

The discussion to this point has assumed complete information, most critically on the realization of V_0 or alternatively θ_0, so that the firm and worker can write explicit contracts guaranteeing performance if they wish. What if the employer alone observes the true productivity realization? If the firm is reliable, perhaps motivated by reputational effects, the previous analysis is unchanged. If private unemployment insurance payments are not feasible, labor hoarding will result.

If the firm is not reliable, the labor hoarding contract is no longer viable. The worker would undertake *only* rigid wage contracts with an unreliable firm in this simple model since wages but not the realized V_0 are observable and therefore amenable to explicit contracts. Not only would the firm find it profitable to lay off workers whenever V_0 fell below $V_1 = \mu_1$, if it contracted for a rigid wage $W > \mu_1$, it would lay off workers profitably whenever $V_0 < W$. Rigid wage contracts with unreliable firms, if they existed at all, would be characterized by high wages and high layoff rates. The existence of such contracts in equilibrium is dependent on the shape of the distribution of realized productivities, V_0 or alternative θ_0.

The worker negotiating with the "unreliable" firm has somewhat less stark alternatives available than is suggested by this simple model. Even though the actual realization of θ_0 is not available for inclusion in the employment contract, presumably quantities correlated with θ_0 are observable to the worker and to third parties. Most obviously wage and employment prospects could be tied to the firm's profitability if it is a large, publicly held firm or perhaps to government statistics on industry sales if profitability or other firm specific data is not available. If these measures are relatively highly correlated with the firm's demand, then contracts may be feasible that approximate efficient separation and complete insurance. Similarly, if severance pay payments are restricted, the second-best contract of labor hoarding may at least be obtainable. The degree to

which firm demand can be predicted by public statistics is an important empirical issue which has not to my knowledge been pursued.

Recently researchers have proposed the firm's employment level itself as an indicator of product demand conditions, albeit an endogenous one, e.g. Calvo and Phelps (1977) and Hall and Lilien (1979).[32] At the individual level the discrete (zero–one) nature of employment and layoff in the models considered makes the employment variable useless as a contract performance monitor. However, if a work sharing arrangement is contracted, that is each individual's work hours are reduced proportionately and no one is laid off, more subtle contract relations can be designed. Alternatively, if a firm's work force can be viewed collectively, as might be appropriate for a trade union, the percentage of the work force laid off could be a sensible employment stipulation equivalent to hours reduction for an individual.

The basic intuition behind employment-contingent wage policies is that wage and employment options may be designed in such a way that the employer is induced to reveal accurately the realized state of product demand that he alone can observe. In particular, focusing on the wage attribute alone, if stipulated wages differ across states (and are presumably lowest in the worst states), the firm would always have an incentive to declare that the worst state has occurred. Clearly this makes contracts with wage adjustments across states infeasible. A wage–hours contract, however, that penalizes the firm for declaring a low state of demand by requiring it to employ the worker for less hours in that state may be feasible. The usefulness of such contracts requires that the hours reduction penalize the firm more than the worker, which suggests that the effect of hours variations on (1) worker productivity in the various states and (2) on the worker's utility (the marginal value of leisure) will be crucial to the desirability of such a contract.

The introduction of variable work hours adds an element to the choice set that requires additional discussion even in the complete information case. The unemployment activity and the work activity are no longer simple either/or alternatives but rather may be chosen in any linear combination that add to total time available; the classical labor supply analysis is appropriate. Consider an example in which only two states of the world are possible and that product in each of these two states is a linear function of work hours so

$$V = \mu^j h, \quad j = 1,2, \tag{5.2}$$

where μ^j is a positive productivity parameter and $h \equiv$ hours worked. Assume

[32] More complete reviews of this literature can be found in Hart (1983) and Azariadis and Stiglitz (1983).

$\mu^1 > \mu^2$, i.e. state 1 is the high demand state. Firm profits can be represented by

$$\pi = V - W, \tag{5.3}$$

and worker utility by

$$U = U(W, h), \quad \text{with } U_1 > 0 \text{ and } U_2 < 0. \tag{5.4}$$

In a competitive labor market (zero firm profits) without insurance, the individual would in equilibrium be at the points (W^j, h^j) if the wage rate is $W^j = \mu^j h^j$ in each state; see the somewhat unconventional representation of the labor–leisure graph in Figure 14.4(a).[33] The usual labor–leisure analysis suggests that the substitution effect of greater productivity will induce greater work hours but that the income effect, if leisure is a normal good (as it seems to be empirically), induces fewer work hours, so that the gross or total effect on equilibrium work hours is ambiguous. In the case illustrated, work hours are higher in the high productivity case ($h^1 > h^2$).

Consider now the insurance attribute of the decision process. The worker is quite clearly better off in the high productivity state in Figure 14.4(a) and the question arises whether the firm could not (at least in the full information case) profitably supply security to the worker. Clearly the risk neutral firm could provide such a contract. Assume for simplicity that each of the two states is equally probable and that the prior expected value of the two auction market contracts is sufficient to attract workers into this firm's employment. Assume further that the worker's preference function is such that the income effect on leisure is zero.[34] The employer could improve the worker's expected well-being without reducing his own by offering to transfer equal increments of income (ΔW) from the good state to the bad state until worker utility is equalized across the two states. The resulting full insurance contract is illustrated in Figure 14.4(b). This is the optimal contract in the full information case (the contract that would maximize expected profits, given an expected utility constraint, or conversely maximize expected utility, given an expected profit constraint).

Is this contract feasible if only the employer knows which state is realized? In this special case the answer is clearly yes. Consider again Figure 14.1(b). For a given state of demand parameterized by μ, isoprofit lines can be constructed of the form

$$\mathrm{d}\pi = \mu\,\mathrm{d}h - \mathrm{d}W = 0,$$

[33] At least as a labor economist if not a macroeconomist would view it. This representation appears for example in Azariadis and Stiglitz (1983).

[34] Formally this restricts the worker's utility function to those of the form $U = V(W - g(h))$, where presumably $g' > 0$ and $g'' > 0$.

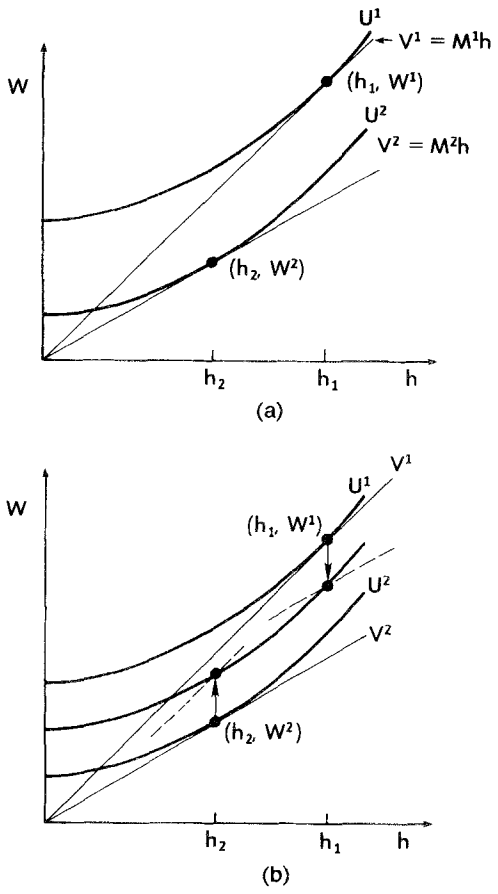


Figure 14.4. Employment-contingent wage contracts.

with the property that all combinations of h and W above and to the left of the line will be less profitable to the firm than those on the line for a given state of demand μ. The isoprofit line through $(h^1, W^1 - \Delta W)$ when $\mu = \mu^1$, denoted by the dashed line through that point, indicates that the employer has a profit incentive to reveal the true state in this case. He can reach no larger isoprofit line by claiming $(h^2, W^2 + \Delta W)$. Similarly, it is more profitable to report state 2 when state 2 in fact occurs; the isoprofit line through the state 2 contract point when $\mu = \mu^2$ indicates profits are greater than if state 1 is falsely announced.

The feasibility of the full insurance contract for the no-income effect utility function rests on the fact that insurance contracting brings the worker to the same indifference curve and that the indifference curve is convex to the hours

axis. In the more general case in which income effects are not zero, insurance motives will not lead to the same level of utility in both states and the analysis is more complex. Chari (1983) demonstrates in a model similar to the one discussed here that the "truth-telling" constraints, the requirement that the wage–employment induce the firm to reveal product demand accurately, will distort the wage–employment package if leisure is a normal good and that, moreover, employment will be too large in the low demand state in this framework. Again labor hoarding will occur. Green and Kahn (1983) consider an otherwise similar model with a continuously distributed random productivity shock and prove that in that case overemployment will occur if leisure is a normal good and conversely underemployment will occur if leisure is an inferior good. Since the empirical literature strongly supports the claim that leisure is a normal good, these models would seem to be quite consistent with the earlier work of Baily (1974), Gordon (1974), and Azariadis (1975).

5.2. *Aggregate productivity risk*

The simple characterization of the employer as risk neutral must be questioned in the case of aggregate business cycle risk since the firm owners can no longer diversity away any risk they bear in this job match; total wealth is reduced, not simply reallocated. The standard risk neutrality argument must instead be based on the empirical conjecture that risk aversion decreases with wealth and that owners are wealthier than workers. One would expect that workers would never be fully insured in this case but would enter into a coinsurance agreement with firm owners, with the sharing of the risk dependent on the relative risk aversion of each party. Indeed optimizing models of this sort are straightforward to construct [Reagan and Stulz (1983)].

Grossman and Hart (1983) and Azariadis (1983) in a framework similar to that discussed above demonstrate that in the situation in which (1) demand information is held solely by a risk averse, unreliable firm and (2) the worker's utility function is characterized by zero income effect on leisure, a wage–hours contract may lead to underemployment, not overemployment of workers. Indeed, if workers are risk neutral, a model similar to that discussed above not only may but will lead to underemployment.

It is perhaps not surprising that a model as counterintuitively constructed as this one will lead to counterintuitive implications. The contracting problems in the asymmetric information case are unusually severe. The firm's incentive to deceive, and therefore the need to impose employment penalties if a low demand state is reported, is a function of the difference in wage payments between the two states. If the risk neutral worker is to insure the firm's owners, however, the wage differential between the high productivity state and the low productivity

state must be *larger* than the auction market outcomes. The worker must accept less than his productivity in the bad state and more in the good state. The greater fluctuations of wages than productivity implied by this model will induce unusually severe employment penalties on the firm if it is to report the state of demand truthfully.

Empirically one might imagine that this structure is uninteresting; theoretically however it is important. It is a reminder that what appears to be a theoretical result is in fact based on implicit empirical propositions (in this case of observability of key variables, desire for risk bearing of economic agents, etc.). More importantly the recognition that employers may be risk averse to some extent explains why insurance within the firm may not be complete even in a situation of complete information and complete freedom of contracting. Additional empirical work on risk sharing in employment contracts is clearly needed.[35]

6. Conclusion

An efficient economy requires that individuals be appropriately matched with a job or other productive activity and that, once matched, the individual undertake the assigned activity with an appropriate level of intensity or effort. Extensive, if nonrigorous, empirics suggest that free markets perform these functions well in economies largely composed of self-employed individuals and small firms, although earnings insurance and other distributional issues may remain problems. The rapid increase in firm size in the United States and other industrialized countries in this century and the related decline in self-employment raises anew these critical efficiency questions. Does the introduction of an arms length relationship between employer and employee change the nature and quality of the economic outcomes?[36] How well do employment contracts, implicit and explicit, facilitate the efficient outcome?

[35] Much empirical evidence exists to support the notion that wages and employment do not respond in the short run as the simplest auction market model would suggest. Focused analyses of the earnings insurance function of the firm, however, are few. Hashimoto (1975) and Raisian provide evidence that suggests that wages are less rigid than commonly supposed. Brown (1982) explores the time limitation of wage rigidity in the U.S. manufacturing sector and finds that this sector approximates the auction market in the long run, although the long run is indeed long, approximately six years. Kniesner and Goldsmith (1984) provide a comprehensive review of this literature.

[36] Certainly compensation schedules appear to differ in large firms. Economists, for example, have known for some time that large firms pay an average higher wages for the same skills than do small firms [Mellow (1982)]. Garen (1982) has presented convincing, if not conclusive evidence that the large firm wage differential is the result of differential information on employees and as a consequence adverse selection among new hires in large firms. Parsons (1983) documents the large firm size effects on mandatory retirement and the provision of actuarially unfair pensions in large firms. He develops an adverse selection model that explains the relative lack of demand for retirement aged workers in such firms.

In the previous sections I have examined what I believe are three important aspects of the employment relationship. The limitations on contracts are central to the resolution of these questions. The limitations derive primarily from two sources. The fact that human capital is embodied in individuals whose liberty is independently valued has led to restrictions on the individual's freedom to contract about future earnings, both directly and indirectly through bankruptcy clauses. The availability of information (or at a deeper level the magnitude of information costs) is a second source of contract limitations. Obviously, for formal contracts to be enforced, all important information must be observable to a third party at some reasonable cost. Implicit contracts between the parties are less demanding of information availability since third party observability is not necessary. Other questions of contract reliability and enforcement arise when contract performance is no longer overseen by the courts. Agent reliability must be induced by other mechanisms, perhaps by the value of repeated exchange or by reputational considerations.

This chapter focuses on three important areas of the employment relationship: (1) the effort intensity of employees, (2) the acquisition and retention of firm specific human capital, and (3) employment and earnings stability as insurance. In each case the fundamental economic issues are outlined and then the likelihood and consequences of contracting limitations are explored. A common pattern has evolved, to some extent independently, in these three literatures. In the case of homogeneous workers, the crucial questions become: (1) Which agents, if any, have relatively inexpensive access to high quality performance information? (2) What is the contract reliability of the agent with this information access? (3) What form of contract (compensation package) will lead to accurate revelation of the information or appropriate agent behavior without explicit information revelation? The important supplementary question of the optimal *costly* information collection process has been less systematically pursued, although such considerations will surely reduce the most aberrant consequences of employment contracting difficulties.

The various literatures also explore optimal compensation schemes when workers are heterogeneous in one or another performance–reliability dimension. The effect on compensation practices may be great if the employee knows this characteristic of himself but the employer does not. The question then becomes: How can the compensation package be altered to induce the worker to reveal through his actions this private information?

A more detailed review of the various findings is not warranted at this point. I should mention, however, that these three substantive areas are only a subset of the interesting theoretical work on the employment relationship that has recently been undertaken. The reader is particularly referred to the theoretical literature on seniority [Gordon (1974), Grossman (1977, 1978), and Reagan (1984)] and on the independent effect of effort monitoring on hierarchies and firm size [e.g.

Williamson (1967), Alchian and Demsetz (1972), Calvo and Wellicz (1978), and Rosen (1983) and from a different perspective, Lucas (1978) and Oi (1983)].[37]

The empirical analysis of employment contracting has only begun. I suspect that much more empirical work is now necessary if progress in this area is not to degenerate into the relatively uninsightful enumeration of the theoretical possibilities. I suspect moreover that the collection of new data specifically focused on employment practices and policies will be essential if the empirical work is to illuminate employment contracting in any serious way.[38] Unfortunately, as empirical researchers in industrial organization have found, firms tend to be much more wary of providing information than individuals and families tend to be, so that large scale survey research may not be as rewarding as it has been in other areas of labor economics.

References

Abowd, John M. and Orley Ashenfelter (1981) "Anticipated unemployment, temporary layoffs, and compensating wage differentials", in: Sherwin Rosen, ed., *Studies in labor markets*. Chicago: University of Chicago Press, 141–170.

Abraham, Katharine G. and James L. Medoff (1983) "Length of service and the operation of internal labor markets", NBER Working Paper No. 1085.

Abraham, Katharine G. and James L. Medoff (1983) "Length of service, terminations, and the nature of the employment relationship", NBER Working Paper No. 1086.

Akerlof, George (1970) "The market for lemons: qualitative uncertainty and the market mechanism", *Quarterly Journal of Economics*, 84:488–500.

Akerlof, George (1976) "The economics of caste and of the rat race and other woeful tales", *Quarterly Journal of Economics*, 80:599–617.

Akerlof, George A. and Brian G. M. Main (1981) "An experience-weighted measure of employment and unemployment duration", *American Economic Review*, 71:1003–1011.

[37]The empirical analysis of economic behavior within and between firm job hierarchies shows considerable promise. Medoff and Abraham (1980, 1981) have found that within managerial and professional job grades job performance ratings and relative wage are inversely related across experience levels; relative performance is lower, relative wages higher for individuals within a job category the longest. Medoff and Abraham argue that if performance ratings are meant to be narrowly defined measures of actual productivity, this relationship is inconsistent with a simple auction model and suggests other forces are at work, plausibly productivity insurance. This interpretation is not valid if worker performance ratings are based on all tasks assigned, including training. If training is important, as it will be for newer workers, performance could be outstanding yet actual product and auction wages minimal. Indeed, Brown (1983) presents evidence that positive experience effects on wages end with the end of the (self-reported) training period. Exploring the interaction of the internal and external market Cho (1983) has found that the job search activities and the quit propensities of young workers are powerfully affected by their promotion prospects as well as more customary variables such as wages, education, etc.

[38]For recent promising analyses of employment behavior based on employer samples, see Abraham and Medoff (1983a, 1983b) and Bishop (1982). Wolpin (1977) and Lazear and Moore (1984) attempt to use behavioral comparisons of the self-employed and wage and salary earners to explore firm size effects. This is certainly a promising approach, although rather little is known of the decision to become a self-employed individual so that such comparisons must be approached with caution.

Akerlof, George A. and Hajime Miyazaki (1980) "The implicit contract theory of unemployment meets the wage bill argument", *Review of Economic Studies*, 47:321–338.
Alchian, A. A. and H. Demsetz (1972) "Production, information costs, and economic organization", *American Economic Review*, 62:777–795.
Allen, Steven G., Robert L. Clark and Daniel A. Sumner (1984) "Post retirement adjustments of pension benefits", working paper, North Carolina State University.
Azariadis, Costas (1975) "Implicit contracts and underemployment equilibria", *Journal of Political Economy*, 83:1183–1202.
Azariadis, Costas (1981) "Implicit contracts and related topics: a survey", in: Zmira Horstein, Joseph Grice and Alfred Webb, eds., *Economics of the labor market*. Her Majesty's Stationery Office, 221–248.
Azariadis, Costas (1983) "Employment with asymmetric information", *Quarterly Journal of Economics*, 98(Supplement):157–172.
Azariadis, Costas and Joseph E. Stiglitz (1983) "Implicit contracts and fixed-price equilibria", *Quarterly Journal of Economics*, 98(Supplement):1–22.
Baily, M. N. (1974) "Wages and employment under uncertain demand", *Review of Economic Studies*, 41:37–50.
Ball, R. J. and E. B. A. St. Cyr (1966) "Short-run employment functions in British manufacturing", *Review of Economic Studies*, 33:179–207.
Bartel, Ann P. and George J. Borjas (1981) "Wage growth and job turnover: an empirical analysis", in: Sherwin Rosen, ed., *Studies in labor markets*. Chicago: University of Chicago Press, 65–84.
Barton, John H. (1972) "The economic basis of damages for breach of contract", *Journal of Legal Studies*, 1:277–304.
Baumol, William J. (1967) *Business behavior, value, and growth*. New York: Harcourt, Brace and World.
Becker, Gary (1975) *Human capital*. New York: Columbia University Press.
Becker, Gary (1968) "Crime and punishment: an economic analysis", *Journal of Political Economy*, 76:169–217.
Becker, G. and G. Stigler (1974) "Law enforcement and compensation of enforcers", *Journal of Legal Studies*, 3:1–18.
Benjamin, Daniel K. (1978) "The use of collateral to enforce debt contract", *Economic Inquiry*, 16:333–359.
Berhold, Marvin (1971) "A theory of linear profit-sharing incentives", *Quarterly Journal of Economics*, 85:460–482.
Bishop, John H. (1982) "The social payoff for occupationally specific training: the employers' point of view", National Center for Research in Vocational Education.
Brechling, F. (1965) "The relationship between output and employment in British manufacturing industries", *Review of Economic Studies*, 32:187–216.
Brown, James N. (1982) "How close to an auction is the labor market?", in: Ronald Ehrenberg, ed., *Research in labor economics*. Volume 5, 189–235.
Brown, James (1983) "Are those paid more really no more productive? Measuring the relative importance of tenure versus on-the-job training in explaining wage growth", Working Paper No. 169, Industrial Relations Section, Princeton University.
Bull, Clive (1983) "Implicit contracts in the absence of enforcement and risk aversion", *American Economic Review*, 73:658–671.
Burdett, K. (1978) "Theory of employee search: quit rates", *American Economic Review*, 68:212–220.
Calvo, Guillermo A. and Edmund S. Phelps (1977) "Employment contingent wage contracts", in: Karl Brunner and Allan H. Meltzer, eds., "Stabilization of the Domestic and International Economy", *Journal of Monetary Economics, Supplement*:160–168.
Calvo, Guillermo A. and S. Wellisz (1978) "Supervision, loss of control, and the optimum size of the firm", *Journal of Political Economy*, 86:943–952.
Chari, V. V. (1983) "Involuntary unemployment and implicit contracts", *Quarterly Journal of Economics*, 98(Supplement):107–122.
Charmichael, H. Lorne (1984) "Reputations in the labor market", *American Economic Review*, 74:713–725.

Cheung, S. (1969) "Transaction costs, risk aversion, and the choice of contractual arrangements", *Journal of Law and Economics*, 12:23–42.
Clark, C. Scott (1973) "Labor hoarding in durable goods industries", *American Economic Review*, 63:811–824.
Cho, Woo (1983) "Promotion prospects, job search and the quit behavior of employed youth", mimeo, The Ohio State University.
Eaton, B. Curtis and William D. White (1982) "Agent compensation and the limits of bonding", *Economic Inquiry*, 20:330–343.
Fama, Eugene F. (1980) "Agency problems and the theory of the firm", *Journal of Political Economy*, 88:288–307.
Feldstein, M. (1976) "Temporary layoffs in the theory of unemployment", *Journal of Political Economy*, 84:937–957.
Freeman, S. (1977) "Wage trends as performance displays productivity potential: a model and application to academic early retirement", *Bell Journal of Economics*, 8:419–443.
Garen, John (1982) "Worker heterogeneity, costly monitoring, wage rates, and firm size", mimeo, Wayne State University.
Gould, John P. (1970) "Diffusion processes and optimal advertising policy", in: Edmund S. Phelps et al., eds., *Microeconomic foundations of employment and inflation theory*. New York: W. W. Norton & Co.
Green, Jerry and Charles M. Kahn (1983) "Wage-employment contracts", *Quarterly Journal of Economics*, 98(Supplement):173–188.
Green, J. R. and N. L. Stokey (1983) "A comparison of tournaments and contracts", *Journal of Political Economy*, 91:349–364.
Gordon, D. F. (1974) "A neo-classical theory of Keynesian unemployment", *Economic Inquiry*, 12:431–459.
Grossman, H. I. (1977) "Risk shifting and reliability in labor markets", *Scandinavian Journal of Economics*, 79(2):187–209.
Grossman, H. I. (1978) "Risk shifting, layoffs, and seniority", *Journal of Monetary Economics*, 4:661–686.
Grossman, Sanford J. and Oliver D. Hart (1983) "Implicit contracts under asymmetric information", *Quarterly Journal of Economics*, 98(Supplement):123–156.
Guasch, J. Luis (1983) "Disturbing effects of contracts based on effort supervision", *Economic Letters*, 11:305–310.
Hall, R. E. (1982) "The importance of lifetime jobs in the U.S. economy", *American Economic Review*, 72:716–724.
Hall, R. E. and D. M. Lilien (1979) "Efficient wage bargains under uncertain supply and demand", *American Economic Review*, 69:868–879.
Harris, Milton and Bengt Holstrom (1982) "A theory of wage dynamics", *Review of Economic Studies*, 49:315–333.
Harris, M. and A. Raviv (1979) "Optimal incentive contracts with imperfect information", *Journal of Economic Theory*, 20:231–259.
Hart, Oliver D. (1983) "Optimal labour contracts under asymmetric information: an introduction", *Review of Economic Studies*, 50:3–35.
Hashimoto, M. (1975) "Wage reduction, job separation and specific human capital", *Economic Inquiry*, 13:485–504.
Hashimoto, M. (1981) "Firm-specific human capital as a shared investment", *American Economic Review*, 71:475–482.
Hashimoto, M. and B. T. Yu (1980) "Specific capital, employment contracts and wage rigidity", *Bell Journal of Economics*, 11:536–549.
Hatch, Sara P. (1982) "Pension plan provisions that directly affect retirement age choice", Urban Institute Project Report.
Holstrom, B. (1979) "Moral hazard and observability", *Bell Journal of Economics*, 10:74–91.
Holstrom, B. (1981) "Contractual models of the labor market", *American Economic Review Proceedings*, 71:308–313.
Holstrom, B. (1982) "Moral hazard in teams", *Bell Journal of Economics*, 13:324–340.

Jensen, M. C. and W. H. Mechling (1976) "Theory of the firm: managerial behavior, agency costs, and ownership structure", *Journal of Financial Economics*, 3:304–360.
Johnson, W. R. (1978) "A theory of job shopping", *Quarterly Journal of Economics*, 92:261–278.
Jovanovic, B. (1979) "Firm specific capital and turnover", *Journal of Political Economy*, 87:1246–1260.
Kennan, John (1979) "Bonding and the enforcement of labor contracts", *Economic Letters*, 3:61–66.
Kihlstrom, Richard E. and Michael H. Riordan (1984) "Advertising as a signal", *Journal of Political Economy*, 92:427–450.
Klein, Benjamin and Keith B. Leffler (1981) "The role of market forces in assuring contractual performance" *Journal of Political Economy*, 89:615–641.
Kniesner, Thomas J. and Arthur H. Goldsmith (1984) "Is the labor market an auction market, a contract market or something in between? A survey of the evidence for the United States", mimeo, University of North Carolina.
Kotowitz, Y. and F. Mathewson (1979) "Advertising, consumer information, and product quality", *Bell Journal*, 10:566–588.
Kreps, David M. and Robert Wilson (1982) "Reputation and imperfect information", *Journal of Economic Theory*, 27:253–279.
Landes, W. M. and R. A. Posner (1979) "Adjudication as a public good", *Journal of Legal Studies*, March:235–283.
Lazear, E. (1979) "Why is there mandatory retirement?", *Journal of Political Economy*, 87:1261–1284.
Lazear, E. (1981) "Agency, earnings profiles, productivity, and hours restrictions", *American Economic Review*, 71:606–620.
Lazear, Edward P. and Robert L. Moore (1984) "Incentives, productivity, and labor contracts", *Quarterly Journal of Economics*, 99:275–296.
Lazear, E. P. and S. Rosen (1981) "Rank-order tournaments as optimum labor contracts", *Journal of Political Economy*, 89:841–864.
Lebergott, Stanley (1964) *Manpower in economic growth: the American record since 1800*. New York: McGraw-Hill.
Lewellen, Wilbur G. (1971) *The ownership income of management*. New York: National Bureau of Economic Research.
Lippman, S. A. and J. J. McCall (1976a) "The economics of job search: a survey", *Economic Inquiry*, 14:155–189.
Lippman, S. A. and J. J. McCall (1976b) "The economics of job search: a survey", *Economic Inquiry*, 14:347–368.
Lucas, Robert E., Jr. (1978) "Adjustment costs and the theory of supply", *Bell Journal of Economics*, 9:508–523.
Medoff, James L. and Katharine G. Abraham (1980) "Experience, performance, and earnings", *The Quarterly Journal of Economics*, 94:703–735.
Medoff, James L. and Katharine G. Abraham (1981) "Are those paid more really more productive? The case of experience", *Journal of Human Resources*, 16:186–216.
Luce, R. D. and H. Raiffa (1957) *Games and decisions*. New York: Wiley.
MacDonald, Glenn M. (1982) "Specific investments and nonlabor income", *Bell Journal of Economics*, 13:225–233.
Mellow, Wesley (1982) "Employer size and wages", *Review of Economics and Statistics*, 64:495–501.
Mincer, Jacob (1962) "On-the-job training: costs, returns and some implications", *Journal of Political Economy*, 70(Supplement):50–79.
Mincer, J. and B. Jovanovic (1981) "Labor mobility and wages", in: S. Rosen, ed., *Studies in labor markets*. Chicago: University of Chicago Press, 21–64.
Miyazaki, Hajime (1977) "The rat race and internal labor markets", *Bell Journal of Economics*, 8:394–418.
Mortensen, D. (1978) "Specific capital and labor turnover", *Bell Journal of Economics*, 9:572–586.
Nadiri, M. Ishaq and Sherwin Rosen (1973) *A disequilibrium model of demand for factors of production*. New York: National Bureau of Economic Research.
Nalebuff, B. and J. E. Stiglitz (1983) "Prizes and incentives: towards a general theory of compensation and competition", *Bell Journal of Economics*, 14:21–43.
Nelson, Phillip (1974) "Advertising as information", *Journal of Political Economy*, 82:729–754.

Nerlove, Marc and K. J. Arrow (1962) "Optimal advertising policy under dynamic conditions", *Economica*, May:124–142.
Nickell, S. J. (1976) "Wage structures and quit rates", *International Economic Review*, 17:191–203.
Oi, W. Y. (1962) "Labor as a quasi-fixed factor", *Journal of Political Economy*, 70:538–555.
Oi, W. Y. (1983) "Heterogeneous firms and the organization of production", *Economic Inquiry*, 21:147–171.
Parsons, Donald O. (1972) "Specific human capital: an application to quit rates and layoff rates", *Journal of Political Economy*, 80:1120–1143.
Parsons, David O. (1977) "Models of labor market turnover: a theoretical and empirical survey", in: R. Ehrenberg, ed., *Research in labor economics*. Greenwich, Conn.: JAI, vol. 1.
Parsons, David O. (1980) "Unemployment, the allocation of labor, and optimal government intervention", *American Economic Review*, 70:626–635.
Parsons, David O. (1983) "The industrial demand for older workers", mimeo.
Parsons, David O. (1984) "Social insurance with imperfect state verification: income insurance for the disabled", Working paper #84-27, Economics Department, Ohio State University.
Pauly, M. V. (1974) "Overinsurance and public provision of insurance: the role of moral hazard and adverse selection", *Quarterly Journal of Economics*, 88:44–54.
Pencavel, John H. (1972) "Wages, specific training, and labor turnover in U.S. manufacturing industries", *International Economic Review*, 13:53–64.
Pencavel, John H. (1977) "Work effort, on-the-job screening, and alternative methods of remuneration", in: R. G. Ehrenberg, ed., *Research in labor economics*. Greenwich, Conn.: JAI, vol. 1.
Perri, Timothy J. (1983) "Explicit labor contracts, shirking, and turnover costs", *Atlantic Economic Review*, December:59–65.
Peters, Elizabeth (1984) "Marriage and divorce in a private contracting framework", mimeo, Center for Human Resource Research, Ohio State University.
Radner, Roy (1981) "Monitoring cooperative agreements in a repeated principal-agent relationship", *Econometrica*, 49:1127–1148.
Raisian, John (1983) "Contracts, job experience, and cyclical labor market adjustments", *Journal of Labor Economics*, 1:152–170.
Reagan, Patricia (1984) "Seniority in hierarchical labor markets", mimeo, Ohio State University.
Reagan, Patricia B. and René M. Stulz (1982) "Risk bearing, labor contracts, and capital markets", mimeo.
Rosen, Sherwin (1972) "Learning and experience in the labor market", *Journal of Human Resources*, 7:326–342.
Rosen, Sherwin (1982) "Authority, control, and the distribution of earnings", *Bell Journal*, 13:311–323.
Rothschild, M. and J. Stiglitz (1976) "Equilibrium in competitive insurance markets: an essay on the economics of imperfect information", in: Diamond and Rothschild, eds., *Uncertainty in economics*.
Salop, S. (1973) "Wage differentials in a dynamic theory of the firm", *Journal of Economic Theory*, 6:321–344.
Salop, J. and S. Salop (1976) "Self-selection and turnover in the labor market", *Quarterly Journal of Economics*, 90:619–628.
Schmalensee, Richard (1978) "Advertising and product quality", *Journal of Political Economy*, 86:485–503.
Shapiro, Carl (1982) "Consumer information, product quality, and seller reputation", *Bell Journal of Economics*, 10:55–73.
Shapiro, Carl (1980) "Damage measures for breach of contract", *Bell Journal of Economics*, 11:466–490.
Shavell, Steven (1979) "Risk sharing and incentives in the principal and agent relationship", *Bell Journal*, 10:55–73.
Spence, Michael (1973) "Job market signaling", *Quarterly Journal of Economics*, 87:355–374.
Spence, Michael and Richard Zeckhauser (1971) "Insurance, information, and individual action", *American Economic Review*, 61:380–387.
Stelluto, George L. (1969) "Report on incentive pay in manufacturing industries", *Monthly Labor Review*, 92:49–53.
Stiglitz, J. E. (1974) "Incentives and risk sharing in sharecropping", *Review of Economic Studies*, April:219–255.

Stiglitz, J. E. (1975) "Incentives, risk, and information: notes toward a theory of hierarchy", *Bell Journal of Economics*, 6:552–579.
Telser, Lester G. (1980) "A theory of self-enforcing agreements", *Journal of Business*, 53:27–44.
Topel, Robert and Finis Welch (1980) "Unemployment insurance: survey and extensions", *Economica*, 47:351–379.
Townsend, Robert M. (1979) "Optimal contracts and competitive markets with costly state verification", *Journal of Economic Theory*, 21:265–293.
U.S. Bureau of the Census (1977) *1977 Census of Manufacturers*. Washington: U.S. Government Printing Office.
U.S. Bureau of the Census (1979) *Statistical Abstract of the United States: 1979*. Washington: U.S. Government Printing Office.
U.S. Department of Labor (1980) *Handbook of Labor Statistics*. Washington: U.S. Government Printing Office.
Viscusi, W. K. (1980) "Self-selection, learning-induced quits, and the optimal wage structure", *International Economic Review*, 21:529–546.
Williamson, O. E. (1967) "Hierarchical control and optimal firm size", *Journal of Political Economy*, 75:123–138.
Williamson, O. E. (1963) "Managerial discretion and business behavior", *American Economic Review*, 53:1032–1057.
Williamson, Oliver E., Michael L. Wachter and Jeffrey E. Harris (1975) "Understanding the employment relation: the analysis of idiosyncratic exchange", *Bell Journal*, Spring:250–278.
Wolpin, Kenneth I. (1977) "Education and screening", *American Economic Review*, 67:949–958.